

ADVANCED TOPICS IN SCIENCE AND TECHNOLOGY IN CHINA

Biao Wang

Mechanics of Advanced Functional Materials



ZHEJIANG UNIVERSITY PRESS
浙江大学出版社



Springer

ADVANCED TOPICS IN SCIENCE AND TECHNOLOGY IN CHINA

ADVANCED TOPICS IN SCIENCE AND TECHNOLOGY IN CHINA

Zhejiang University is one of the leading universities in China. In *Advanced Topics in Science and Technology in China*, Zhejiang University Press and Springer jointly publish monographs by Chinese scholars and professors, as well as invited authors and editors from abroad who are outstanding experts and scholars in their fields. This series will be of interest to researchers, lecturers, and graduate students alike.

Advanced Topics in Science and Technology in China aims to present the latest and most cutting-edge theories, techniques, and methodologies in various research areas in China. It covers all disciplines in the fields of natural science and technology, including but not limited to, computer science, materials science, life sciences, engineering, environmental sciences, mathematics, and physics.

Biao Wang

Mechanics of Advanced Functional Materials

With 170 figures



ZHEJIANG UNIVERSITY PRESS
浙江大学出版社



Springer

Author

Prof. Biao Wang
School of Physics and Engineering
Sun Yat-sen University
Guangzhou 510275, China
E-mail: wangbiao@mail.sysu.edu.cn

ISSN 1995-6819

e-ISSN 1995-6827

Advanced Topics in Science and Technology in China

ISBN 978-7-308-10025-0

Zhejiang University Press, Hangzhou

ISBN 978-3-642-33595-2

ISBN 978-3-642-33596-9 (eBook)

Springer Heidelberg New York Dordrecht London

Library of Congress Control Number: 2012947313

© Zhejiang University Press, Hangzhou and Springer-Verlag Berlin Heidelberg 2013

This work is subject to copyright. All rights are reserved by the Publishers, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed. Exempted from this legal reservation are brief excerpts in connection with reviews or scholarly analysis or material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work. Duplication of this publication or parts thereof is permitted only under the provisions of the Copyright Law of the Publishers' locations, in its current version, and permission for use must always be obtained from Springer. Permissions for use may be obtained through RightsLink at the Copyright Clearance Center. Violations are liable to prosecution under the respective Copyright Law.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publishers can accept any legal responsibility for any errors or omissions that may be made. The publishers make no warranty, express or implied, with respect to the material contained herein.

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

The author was interested in the piezoelectric solids, as a kind of functional materials, in the early 1990s. At that time the concept of smart materials and structures just emerged. It is very important to investigate the coupling effects of the elastic field and electric field on the material behaviors, especially when the material contains some inclusions and defects. Based on the research results, one can predict the effective properties of smart composites. With the progresses of the research activities related with the piezoelectric, ferroelectric, optoelectronic and other functional materials, it is getting more and more clear that, the strain or stress field distributions play a very important role in manipulating the electronic, magnetic and optical properties of the materials. Definitely at present stage no systematic theory has been established to deal with these problems.

The continuum mechanics conventionally studies the deformation, strength and the related mechanical properties of the materials. The initial task of mechanics was dealing with the driving force for particles motion, and then it was extended to become a continuum field theory to deal with the deformation. It should be extended to study the general driving force for the various functional behaviors of the materials. This book is an attempt to tackle mainly the following two problems: (1) to analyze the effect of stress and deformation on the functional properties of the materials, and (2) to establish the quantitative models related with the microstructural evolution. The general formulation will be developed from the detailed analyses of the separated examples.

Many of my students and former students had contributed to this book greatly. I would like to mention only a few of them. Dr. Yue Zheng, my former student, now a professor of School of Physics and Engineering, Sun Yat-sen University, Dr. Yangfan Hu, Mr. Xunlong Zhou, Mr. Xiaoyue Zhang, Dr. Xin Luo, Dr. Xiaoyan Lu, Mr. Weijing Chen, Dr. Kan Lai, Mr. Genghong Zhang, and Dr. Decai Ma for helping me to revise the format of the each chapter. I would like to express my sincere gratitude to all of them.

Finally, I would like to express my sincere gratitude to the National Natural Science Foundation of China for financial support through the key and general grants for many years.

Biao Wang
Guangzhou, China
March, 2012

Contents

1	Introduction	1
2	Basic Solutions of Elastic and Electric Fields of Piezoelectric Materials with Inclusions and Defects	5
2.1	The Coupled Differential Equations of Elastic and Electric Fields in Piezoelectric Solids.....	6
2.1.1	Thermodynamic Framework.....	6
2.1.2	Linear Constitutive Equations.....	8
2.1.3	The Equation of Equilibrium	11
2.1.4	The Basic Equations of a Static Electric Field	16
2.1.5	Differential Equations for Piezoelectric Materials	16
2.2	Boundary Conditions.....	17
2.3	Solution Methods for Two-Dimensional Problems	19
2.3.1	The Stroh Formalism for Piezoelectric Materials	19
2.3.2	The Lekhnitskii Formalism for Piezoelectric Materials	24
2.3.3	Conformal Transformation of the Core Function.....	29
2.4	Basic Solutions for Two-Dimensional Problems.....	30
2.4.1	Elliptical Cylindrical Inclusions in Piezoelectric Materials	30
2.4.2	Cracks	40
2.4.3	Dislocations and Line Charges.....	52
2.5	Solution Methods for Three-Dimensional Problems	53
2.5.1	Eigenstrains and Equivalent Inclusion Method.....	53
2.5.2	Method of Fourier Integrals	56
2.5.3	Method of Green's Function	57
2.6	Basic Solution for Three-Dimensional Problems	59
2.6.1	Ellipsoidal Inhomogeneous Inclusions	60
2.6.2	Flat Elliptical Cracks.....	74
2.6.3	Ellipsoidal Inhomogeneity Embedded in an Infinite Matrix when both Phases Undergo Eigenstrains	83
2.6.4	Green's Function.....	85

2.7	Remarks.....	89
	References	90
3	Micromechanics Models of Piezoelectric and Ferroelectric Composites	97
3.1	Background	98
3.2	Some Definitions	99
3.3	Effective Material Constants of Piezoelectric Composites.....	103
3.3.1	The Dilute Model.....	103
3.3.2	The Self-Consistent Model	107
3.3.3	The Mori-Tanaka Mean Field Model	108
3.3.4	The Differential Model	112
3.4	Energy Formulation of Ferroelectric Composites	113
3.4.1	Elastic Strain Energy Density for Ferroelectric Composites...	114
3.4.2	Intrinsic Free Energy Density for Ferroelectric Composites...	117
3.4.3	Total Free Energy for Ferroelectric Composites with Spherical Inclusions	121
3.5	Phase Diagrams	125
3.5.1	Total Free Energy for Ferroelectric Composites with Spherical Inclusions and Equiaxed Strains	125
3.5.2	Phase Diagrams and Total Polarizations	126
3.6	Remarks.....	138
	Appendix A: Radon Transform	140
	References	142
4	Determination of the Smallest Sizes of Ferroelectric Nanodomains	147
4.1	Introduction.....	147
4.2	Electric Fields in Ferroelectric Thin Film	149
4.2.1	General Expression of Electric Field of Ferroelectric Domain....	149
4.2.2	AFM-Induced Electric Field in Ferroelectric Thin Films	152
4.3	Energy Expressions	162
4.3.1	Energy Expression for 180° Domain in a Ferroelectric Film Covered with Top and Bottom Electrodes	162
4.3.2	Energy Expression for 180° Domain in Ferroelectric Film Induced by an AFM Tip without the Top Electrode.....	163
4.4	Driving Force and Evolution Equations of Domain Growth.....	166
4.5	Stability Analysis.....	168
4.6	Remarks.....	171
	Appendix B: Derivation of the Electric and Magnetic Field for a Growing 180° Domain.....	171
	References	177
5	Size and Surface Effects of Phase Transition on Nanoferroelectric Materials	179

5.1	Introduction and Overview of Ferroelectrics in Nanoscale Dimensions ..	179
5.1.1	Ferroelectric Thin Films in Nanoscale Dimensions	179
5.1.2	Ferroelectric Tunneling Junctions and Capacitors in Nanoscale Dimensions	181
5.1.3	Ferroelectric Multilayers in Nanoscale	183
5.1.4	Ferroelectric Nanowires and Nanotubes	184
5.1.5	Ferroelectric Nanograins or Nanoislands on Substrates.....	186
5.2	Thermodynamic Modeling and Stability Analysis of Ferroelectric Systems	188
5.2.1	Background of the Thermodynamic Modeling for Ferroelectrics..	188
5.2.2	Electrostatics for Ferroelectrics.....	190
5.2.3	Thermodynamics of Ferroelectrics	192
5.2.4	Stability Analysis on Critical Properties of Ferroelectric Systems	194
5.3	Ferroelectric Thin Films in Nanoscale	195
5.3.1	Thermodynamic Model for a Thick Ferroelectric Film	195
5.3.2	Size and Surface Effects on Ferroelectric Thin Films	197
5.3.3	The Evolution Equation and Stability of the Stationary States	198
5.3.4	Curie Temperature and Critical Thickness.....	209
5.3.5	Curie-Weiss Law of Ferroelectric Thin Film in Nanoscale.....	211
5.4	Critical Properties of Ferroelectric Capacitors or Tunnel Junctions....	218
5.4.1	The Thermodynamic Potential of the Ferroelectric Capacitors or Tunnel Junctions	219
5.4.2	The Evolution Equation and Stability of the Stationary States	222
5.4.3	Curie Temperature of the Ferroelectric Capacitors or Tunnel Junctions	223
5.4.4	Polarization as a Function of Thickness of the Ferroelectric Capacitors or Tunnel Junctions	225
5.4.5	Critical Thickness of the Ferroelectric Capacitors or Tunnel Junctions	227
5.4.6	Curie-Weiss Relation of the Ferroelectric Capacitors or Tunnel Junctions	229
5.5	Ferroelectric Superlattices in Nanoscale	231
5.5.1	The Free Energy Functional of Ferroelectric Superlattices.....	232
5.5.2	The Phase Transition Temperature of PTO/STO Superlattice ...	235
5.5.3	Polarization and Critical Thickness of PTO/STO Superlattice..	236
5.5.4	The Curie-Weiss-Type Relation of PTO/STO Superlattice	239
5.6	Ferroelectric Nanowires and Nanotubes	240
5.6.1	Surface Tension of Ferroelectric Nanowires and Nanotubes ..	240
5.6.2	Size and Surface Effects on Ferroelectric Nanowires	242
5.6.3	Ferroelectric Nanotubes	249
5.7	Ferroelectric Nanograins or Nanoislands	257
5.7.1	Free Energy of Ferroelectric Nanograins or Nanoislands	258

5.7.2	Stability of the Ferroelectric State and Transition Characteristics.....	259
5.7.3	Critical Properties of Nanograins or Nanoislands.....	261
5.8	Remarks.....	262
	References.....	264
6	Strain Engineering: Ferroelectric Films on Compliant Substrates	269
6.1	Background.....	269
6.2	Manipulation of Phase Transition Behavior of Ferroelectric Thin Films on Compliant Substrates	273
6.2.1	Free Energy Expressions.....	274
6.2.2	Evolution Equations.....	277
6.2.3	Manipulation of Ferroelectric Transition Temperature and Critical Thickness	278
6.2.4	Manipulation of the Order of Transition	282
6.3	Piezoelectric Bending Response and Switching Behavior of Ferroelectric Thin Film with Compliant Paraelectric Substrate.....	286
6.3.1	Model of Ferroelectric Thin Film with Compliant Paraelectric Substrate and the Energy Expressions.....	287
6.3.2	Solution of the Evolution Equation.....	290
6.3.3	The Stationary and Relative Bending Displacements of the Bilayer.....	291
6.3.4	Dynamic Piezoelectric and Bending Response of the Bilayer Under a Cyclic Electric Field	292
6.3.5	Examples and Discussions	293
6.4	Critical Thickness for Dislocation Generation in Piezoelectric Thin Films on Substrate	299
6.4.1	Elastic and Electric Fields in a Piezoelectric Semi-Infinite Space with a Dislocation.....	299
6.4.2	Critical Thickness for Dislocation Generation.....	304
6.4.3	Effect of Piezoelectric Behavior of the Materials on the Critical Thickness for Dislocation Formation	306
6.5	Critical Thickness of Dislocation Generation in Ferroelectric Thin Film on a Compliant Substrate	309
6.5.1	Mechanical Properties of the Problem	310
6.5.2	The Formation Energy and the Critical Thickness of Spontaneous Formation of Misfit Dislocation	312
6.6	Remarks.....	316
	References.....	317
7	Derivation of the Landau-Ginzburg Expansion Coefficients	321
7.1	Introduction.....	321
7.2	Fundamental of the Landau-Devonshire Theory	321

7.2.1	The History of the Landau Free Energy Theory	323
7.2.2	The Thermodynamic Functions of the Dielectrics and Phase Transition	323
7.2.3	The Expansion of the Free Energy and Phase Transition	324
7.3	Determination of Landau Free Energy Expansion Coefficients Based on Experimental Methods	331
7.3.1	The Experimental Observation of the Phase Transition Characteristics in Ferroelectrics	331
7.3.2	The Phenomenological Treatment of Devonshire Theory	335
7.3.3	The Elastic Gibbs Free Energy of PbTiO_3 and Its Coefficients ...	343
7.3.4	The Determination of the Expansion Coefficients from the First-Principle Calculation Based on the Effective Hamiltonian Method	345
7.4	Gradient Terms in the Landau-Devonshire-Ginzburg Free Energy Expansion	346
7.4.1	The Consideration of Spatial Non-uniform Distribution of the Order Parameters in the Landau Theory	346
7.4.2	The Critical Region and the Applicability of Landau Mean Field Theory	347
7.4.3	Determination of the Gradient Terms from the Lattice Dynamic Theory	349
7.4.4	The Extrapolation Length and the Gradient Coefficient	355
7.5	The Transverse Ising Model and Statistical Mechanics Approaches...	358
7.5.1	Phase Transition from the Transverse Ising Model	358
7.5.2	Relationship of the Parameters Between Landau Theory and the Transverse Ising Model	361
7.5.3	Determination of Landau-Ginzburg Free Energy Expansion Coefficients from Statistical Mechanics	363
7.6	Remarks	369
	References	370
8	Multiferroic Materials	377
8.1	Background	377
8.2	Coupling Mechanism of Multiferroic Materials	382
8.2.1	Single Phase Multiferroic Materials	382
8.2.2	Magnetoelectric Composites	401
8.3	Theories of Magnetoelectric Coupling Effect at Low Frequency	407
8.3.1	Energy Formulation for Multiferroic Composites	407
8.3.2	Phase Transition Behaviors in Layered Structures	410
8.3.3	Magnetoelectric Coupling Coefficients in Layered Structures	418
8.4	Magnetoelectric Coupling at Resonance Frequency	419
8.4.1	Magnetoelectric Coupling at Bending Modes	420
8.4.2	Magnetoelectric Coupling at Electromechanical Resonance	426

8.4.3	Magnetoelectric Coupling at Ferromagnetic Resonance.....	429
8.5	Remarks.....	431
	References	432
9	Dielectric Breakdown of Microelectronic and Nanoelectronic Devices ...	443
9.1	Introduction.....	443
9.2	Basic Concepts	448
9.2.1	MOS Structure	448
9.2.2	Different Tunneling Modes	450
9.2.3	Dielectric Breakdown Modes.....	452
9.2.4	Defect Generation	456
9.2.5	Basic Statistical Concepts of Dielectric Breakdown.....	458
9.2.6	Stress Induced Leakage Current.....	461
9.2.7	Holes Generation	462
9.2.8	Energetics of Defects	466
9.3	Mechanism Analysis of Tunneling Phenomena in Thin Oxide Film ...	471
9.3.1	Self-consistent Schrödinger's and Poisson's Equations	472
9.3.2	Transmission Coefficient	480
9.3.3	Tunneling Current Components	486
9.3.4	Microscopic Investigation of Defects from First-Principles Calculation	489
9.3.5	Manipulating Tunneling by Applied Strains	491
9.4	Degradation Models in Gate Oxide Films.....	493
9.4.1	Anode Hole Injection Model.....	494
9.4.2	Thermochemical Model	497
9.4.3	Anode Hydrogen Release Model	499
9.4.4	Thermal Breakdown Model	502
9.4.5	Mechanical-Stress-Induced Breakdown Model	503
9.4.6	Remarks	504
9.5	Statistical Models of Dielectric Breakdown.....	505
9.5.1	A Basic Statistical Model.....	505
9.5.2	A Three-Dimensional Statistical Model	507
9.5.3	Sphere and Cube Based Percolation Models.....	508
9.5.4	Combination of Percolation Model and Degradation Model ...	510
9.6	Damage of Dielectric Breakdown in MOSFET	511
9.6.1	Lateral Propagation of Breakdown Spot	512
9.6.2	Dielectric Breakdown-Induced Epitaxy	512
9.6.3	Dielectric Breakdown-Induced Migration	513
9.6.4	Meltdown and Regrowth of Silicided Poly-Si Gate.....	513
9.6.5	Damage in Substrate	514
9.7	Remarks.....	515
	References	516
	Index.....	525

Introduction

Why is mechanics related to advanced functional materials? It is a well-known fact that mechanics, especially solid mechanics, is directly related to structural materials. The focus of the subject is mainly concerned with displacement, deformation and strength problems. In fact, based on Newton's classical mechanics theory, mechanics gradually becomes a subject for studying only the mechanical motion of particles and its related phenomena. Since the widespread application of continuum mechanics in modern engineering, deformation, stiffness, strength and many mechanical properties have become the overwhelming priorities of modern solid mechanics. Advanced functional materials refer to those materials with some unique electronic, magnetic and optical properties. Therefore, it is common sense that functional materials are far from mechanics areas. An exceptional case is that of piezoelectric materials in which the electric field can induce deformation, whereas the applied stress field can create some electric charges.

Recently, we have found more and more examples of the properties of many functional materials being manipulated by applied mechanical stresses; for example, the phase transition temperature of ferroelectric materials can be tuned over a large range by applying a stress. Even the order of phase transition can also be manipulated. The dielectric material can become a semi-conductor under a stress or deformation. The energy band can also be regulated by a stress distribution. Especially in the case of nanoscale functional materials, the interaction between the mechanical deformation and some functionality is highly enhanced. This is one reason why this book emphasizes the mechanics of functional materials.

Following classical mechanics, the modern mechanics concept should be extended.

At the present stage, the scope of mechanics includes the basic laws, theorems and findings relating to the interaction of mechanical deformation with its driving mechanical force. The scope can be extended to cover the interaction between the **change** in the functional material behaviors with their generalized driving forces, an energy force using thermodynamics. In such a way, the mechanics concept and scope will be extended extensively. Of course, such a relationship is always the

key research goal in different fields, and the whole framework is still in its infancy.

Once we know the expression of the driving force for some changes, it becomes an easy task to establish the evolution equations of the system. For such a general problem, it is common to use thermodynamics to find the driving force as follows.

The first and second laws of thermodynamics can be expressed in the following form as

$$\begin{aligned} dU &= \delta Q + \delta W, \\ \delta Q &\leq TdS, \end{aligned} \quad (1.1)$$

where dU is the change in the internal energy of the system, δQ is the heat transferred to the system, δW is the work performed on the system and dS is the change in the entropy of the system, a state function. From Eq. (1.1), we can find: (1) the equilibrium criteria under different conditions; (2) evolution directions of the system under different conditions. For example, for an adiabatic system, $\delta Q = 0$, we know $dS \geq 0$ for any evolution process, and the equilibrium criterion is $dS = 0$, where the entropy has reached its maximum. In fact, the driving force for such a nonequilibrium process is dS . If we use some parameters $\mathbf{x} = \{x_1, x_2, \dots, x_n\}$ to describe the system, the corresponding driving force with respect to any change in the parameters is given by $\mathbf{f} = \left\{ \frac{\partial S}{\partial x_1}, \frac{\partial S}{\partial x_2}, \dots, \frac{\partial S}{\partial x_n} \right\}$.

Under the isothermal condition without any external work $\delta W = 0$, combining Eq. (1.1) gives

$$\begin{aligned} dF &\leq 0, \\ F &= U - TS, \end{aligned} \quad (1.2)$$

where F is the Helmholtz free energy of the system and the equilibrium criterion is that the Helmholtz free energy has reached its minimum. If some parameters $\mathbf{x} = \{x_1, x_2, \dots, x_n\}$ are used to describe the system, the corresponding driving force with respect to any change in the parameters can be derived in the same way as the above $\mathbf{f} = -\left\{ \frac{\partial F}{\partial x_1}, \frac{\partial F}{\partial x_2}, \dots, \frac{\partial F}{\partial x_n} \right\}$. One should notice that the positive driving force

is the reduction in the free energy. Therefore, a negative sign is added.

Under the isothermal condition, if some work is done during the evolution process, one can derive

$$\begin{aligned} dG &\leq 0, \\ G &= U - TS - W, \end{aligned} \quad (1.3)$$

where G is the Gibbs free energy. In such a case, the equilibrium criterion is that the Gibbs free energy has reached its minimum. If some parameters $\mathbf{x} = \{x_1, x_2, \dots, x_n\}$ are used to describe the system, the corresponding driving force with respect to any change in the parameters can be derived in the same way as the above $\mathbf{f} = -\{\frac{\partial G}{\partial x_1}, \frac{\partial G}{\partial x_2}, \dots, \frac{\partial G}{\partial x_n}\}$. As above, the positive driving force is the reduction in the free energy. Therefore, a negative sign is added.

Any irreversible process of a system can be described in terms of thermodynamic forces and thermodynamic flow as follows:

$$f_i = \frac{\partial S}{\partial x_i}, \quad j_i = \frac{dx_i}{dt}. \quad (1.4)$$

Under different conditions, one should replace S with the corresponding thermodynamic potentials. To establish the evolution equations of any parameter, the most simple form we can assume is that the thermodynamic flow is proportional to its thermodynamic force as follows,

$$f_i = K_i j_i. \quad (1.5)$$

Of course, if the system is far from its equilibrium state, the thermodynamic force may become very large and such a linear relation may not be valid. There is an idea to extend the research scope of mechanics to the generalized force and the generalized response. Although, at the present stage, it is in its infancy, many problems have been solved using the framework. In this work, we will pick some examples as illustration.

The book can be used as a reference work and is suitable for those readers who are interested in analyzing the effect of stress and deformations on the functional properties of materials, and it might be useful for those researchers who are interested in establishing the quantitative models related to microstructural evolution problems. Chapter 2 focuses on the analytical solutions of piezoelectric materials with inclusion and defects. The solution methods for both two-dimensional and three-dimensional problems are given and some basic solutions for coupled electric and elastic fields are listed. Chapter 3 shows the microstructural model to predict the effective constants of piezoelectric and ferroelectric composites. Chapter 4 is devoted to determining the smallest sizes of ferroelectric nanodomains, which is an important engineering problem. We will show how to determine the limiting sizes based on the evolution equations and stability analysis. Chapter 5 mainly focuses on studying the size and surface effects of the phase transition of nanoferroelectric materials, which is also based on evolution equations of the order parameters. Chapter 6 investigates how to manipulate ferroelectric properties using strain engineering. We will show the importance of mechanical deformation on the electric and magnetic properties.

Chapter 7 devotes a whole chapter to show how to determine the Landau-Ginzburg expansion coefficients, mainly concentrated on the coefficient of the gradient terms. Chapter 8 turns to multiferroic materials. Chapter 9 focuses on dielectric breakdown problems of microelectronic and nanoelectronic devices.

Basic Solutions of Elastic and Electric Fields of Piezoelectric Materials with Inclusions and Defects

The solution of elastic and electric fields in solids with heterostructures is the foundation of micromechanics for predicting the overall behavior of materials. For practical considerations, no purely homogeneous material exists. The effect of defects such as dislocations, cracks, and inhomogeneities has always been one of the most important issues in materials science. Even for an assumed homogeneous body, different distributions of phases or preloads will have a significant impact on the fields. For piezoelectric solids, the following three factors provide new challenges for deriving the solutions of the elastic and electric field in the materials: (1) the materials are intrinsically anisotropic; (2) the elastic field and electric field are coupled; (3) sometimes it is extremely difficult to determine the exact boundary conditions for the elastic and electric fields. In the early 1980's, researchers started to establish continuum models for the coupled elastic and electric field in piezoelectric materials. In the early 1990's, the research in this area was significantly enhanced, and some efficient solution methods for the coupled fields were proposed and some milestone analytical solutions were derived. In this chapter, we intend to summarize some basic solution approaches and list some basic important solutions. The solution methods and the basic solutions are classified into two groups: two-dimensional problems and three-dimensional problems. Solutions for some one-dimensional problems have already been well documented in Tiersten (1969). For two-dimensional and three-dimensional problems, the mathematical solution methods are quite different. Some basic solution methods for problems with different dimensions are introduced respectively. For two-dimensional problems, three basic solutions are chosen: the electroelastic fields for the infinity matrix containing one elliptical cylinder inclusion; the electroelastic fields for the infinity matrix containing one Griffith crack; and the electroelastic fields caused by a dislocation and line charge (also known as the 2-D Green's function). For three-dimensional problems, four

basic solutions are chosen: the electroelastic fields for the infinity matrix containing one ellipsoidal inclusion; the electroelastic fields for the infinity matrix containing a penny-shaped crack; the elastic fields for one ellipsoidal inclusion when both the matrix and the inclusion undergo eigenstrains; and the 3-D Green's function for anisotropic piezoelectric material (also known as electroelastic fields for a point defect and point charge). As examples, we arbitrarily choose the above familiar solutions in this book as an illustration; many important solutions have definitely been neglected. For solutions deduced for transversely isotropic piezoelectric materials, Ding and Chen's book (Ding and Chen, 2001) is recommended for readers who are interested.

Most methods and theories introduced in this chapter aim at treating piezoelectric solids with general anisotropy. At some specific points, we may reduce the complexity of the material constant tensor (say, to transverse isotropy) and obtain some explicit results.

Chapter 2 is arranged as follows: in Section 2.1, the basic set of equations concerning the elastic and electric fields for piezoelectric materials is constructed; in Section 2.2, the various kinds of boundary conditions for piezoelectric materials with heterostructures are introduced; in Section 2.3, solution methods for two-dimensional problems of piezoelectric bodies are introduced, focusing on two complex functional methods—the Stroh formalism and the Lekhnitskii formalism; in Section 2.4, solutions for several basic two-dimensional problems for piezoelectric bodies are presented; in Section 2.5, solution methods for three-dimensional problems of piezoelectric bodies are introduced; in Section 2.6, solutions for several basic three-dimensional problems for piezoelectric bodies are presented; a brief remark on Chapter 2 is provided in Section 2.7.

2.1 The Coupled Differential Equations of Elastic and Electric Fields in Piezoelectric Solids

The phenomenon of piezoelectricity was discovered by the brothers, Pierre Curie and Jacques Curie in α -quartz in 1880. The direct piezoelectric effect refers to the production of electric polarization and thus electric field when subject to external pressure, while the reverse piezoelectric effect refers to the production of strain/stress when an electric field is applied.

2.1.1 Thermodynamic Framework

According to the first law of thermodynamics, the infinitesimal change in internal energy dU (per unit volume) of a system equals the amount of work dW done on the system and the heat dQ that enters the system

$$dU = dW + dQ. \quad (2.1.1)$$

According to the second law of thermodynamics, we have

$$dQ \leq TdS, \quad (2.1.2)$$

where dS denotes the production of entropy, and the equal sign is for a reversible process. For a piezoelectric dielectric, dW can be separated into two parts

$$dW = dW_\sigma + dW_e, \quad (2.1.3)$$

where

$$dW_\sigma = \sigma_{ij} d\varepsilon_{ij} \quad (2.1.4)$$

denotes the mechanical work and

$$dW_e = E_i dD_i \quad (2.1.5)$$

denotes the electrical work. In Eqs. (2.1.4) and (2.1.5), σ_{ij} and ε_{ij} denote components of the stress and strain tensors, and E_i and D_i denote components of the electric field and electric displacement vectors. It should be noted that in Eqs. (2.1.4) and (2.1.5) and the whole range of this book, except for specified cases, the Einstein summation convention is applied.

For a reversible process Eq. (2.1.1) can thus be rewritten as

$$dU = \sigma_{ij} d\varepsilon_{ij} + E_i dD_i + TdS. \quad (2.1.6)$$

From the mathematical description of the internal energy in Eq. (2.1.6), it is noted that the internal energy of the system is described by the independent variables ε_{ij} , D_i and S , or called principal variables of the internal energy (Lines and Glass, 1977). In this case, other physical quantities appearing in the expression of the internal energy are dependent on these three. From Eq. (2.1.6), we have

$$\sigma_{ij} = \left(\frac{\partial U}{\partial \varepsilon_{ij}} \right)_{D_i, S}, \quad E_i = \left(\frac{\partial U}{\partial D_i} \right)_{\varepsilon_{ij}, S}, \quad T = \left(\frac{\partial U}{\partial S} \right)_{D_i, \varepsilon_{ij}}. \quad (2.1.7)$$

Eq. (2.1.7) provides a way to obtain the relations between different physical quantities once the exact expression of the internal energy U is known. The three expressions which can be obtained from Eq. (2.1.7) are called, respectively, the elastic, dielectric and calorimetric equations of state, generally called the

constitutive relations.

There are situations where the physical quantities other than ε_{ij} , D_i and S are more appropriate for choices as independent variables. The possible choices are restricted to the three conjugate pairs (σ_{ij} , ε_{ij}), (E_i , D_i), (T , S), from which different independent state variables can be chosen, then different constitutive relations can be established. As a result, eight possible energy expressions exist for the piezoelectric dielectrics. The other seven energy expressions are related to the internal energy by the following Legendre transformations:

$$\text{Helmholtz free energy:} \quad \psi = U - TS, \quad (2.1.8a)$$

$$\text{Enthalpy:} \quad H = U - E_i D_i - \sigma_{ij} \varepsilon_{ij}, \quad (2.1.8b)$$

$$\text{Elastic enthalpy:} \quad H_1 = U - \sigma_{ij} \varepsilon_{ij}, \quad (2.1.8c)$$

$$\text{Electric enthalpy:} \quad H_2 = U - E_i D_i, \quad (2.1.8d)$$

$$\text{Gibbs free energy:} \quad G = U - E_i D_i - \sigma_{ij} \varepsilon_{ij} - TS, \quad (2.1.8e)$$

$$\text{Elastic Gibbs free energy:} \quad G_1 = U - \sigma_{ij} \varepsilon_{ij} - TS, \quad (2.1.8f)$$

$$\text{Electric Gibbs free energy:} \quad G_2 = U - E_i D_i - TS. \quad (2.1.8g)$$

Substitution of Eq. (2.1.6) yields the differential forms of Eqs. (2.1.8a)–(2.1.8g) as

$$d\psi = \sigma_{ij} d\varepsilon_{ij} + E_i dD_i - SdT, \quad (2.1.9a)$$

$$dH = -\varepsilon_{ij} d\sigma_{ij} - D_i dE_i + TdS, \quad (2.1.9b)$$

$$dH_1 = -\varepsilon_{ij} d\sigma_{ij} + E_i dD_i + TdS, \quad (2.1.9c)$$

$$dH_2 = \sigma_{ij} d\varepsilon_{ij} - D_i dE_i + TdS, \quad (2.1.9d)$$

$$dG = -\varepsilon_{ij} d\sigma_{ij} - D_i dE_i - SdT, \quad (2.1.9e)$$

$$dG_1 = -\varepsilon_{ij} d\sigma_{ij} + E_i dD_i - SdT, \quad (2.1.9f)$$

$$dG_2 = \sigma_{ij} d\varepsilon_{ij} - D_i dE_i - SdT. \quad (2.1.9g)$$

2.1.2 *Linear Constitutive Equations*

In Eq. (2.1.7), we have provided the way to obtain constitutive relations based on the expression of the internal energy. To be consistent with most of the early works, the electric Gibbs free energy is used here. Considering the electric Gibbs free energy G_2 in Eq. (2.1.9g), for isothermal processes, Eq. (2.1.9g) becomes

$$dG_2 = \sigma_{ij} d\varepsilon_{ij} - D_i dE_i. \quad (2.1.10)$$

From Eq. (2.1.10), we have

$$\sigma_{ij} = \left(\frac{\partial G_2}{\partial \varepsilon_{ij}} \right)_{E_i}, \quad D_i = - \left(\frac{\partial G_2}{\partial E_i} \right)_{\varepsilon_{ij}}, \quad (2.1.11)$$

which determine σ_{ij} and D_i as a function of (ε_{ij}, E_i) . Now assume that G_2 takes the following form

$$G_2(\varepsilon_{ij}, E_i) = -\frac{1}{2} C_{ijkl}^E \varepsilon_{kl} \varepsilon_{ij} - e_{kij} E_k \varepsilon_{ij} - \frac{1}{2} \alpha_{ik}^E E_i E_k, \quad (2.1.12)$$

where it is shown that only quadratic terms appear. Substituting Eq. (2.1.12) into Eq. (2.1.11), we have

$$\begin{aligned} \sigma_{ij} &= C_{ijkl}^E \varepsilon_{kl} - e_{kij} E_k, \\ D_i &= e_{ikl} \varepsilon_{kl} + \alpha_{ik}^E E_k, \end{aligned} \quad (2.1.13)$$

where C_{ijkl}^E , e_{kij} and α_{ik}^E are called the elastic, piezoelectric and dielectric constants, respectively. A superscript E in C_{ijkl}^E indicates that, rigorously speaking, in the constitutive equations determined by G_2 , the values of elastic constants are determined under given values of the electric field. The same statement can be made for α_{ik}^E . Eq. (2.1.13) is called linear constitutive equations. For the constants given in Eq. (2.1.13), the following relations of symmetry hold:

$$\begin{aligned} C_{ijkl}^E &= C_{klij}^E = C_{jikl}^E = C_{ijlk}^E, \\ e_{kij} &= e_{kji}, \\ \alpha_{ik}^E &= \alpha_{ki}^E. \end{aligned} \quad (2.1.14)$$

Due to the symmetry relations, Eqs. (2.1.14), a compressed matrix notation can be made

$$C_{ijkl}^E = C_{pq}^E, \quad e_{kij} = e_{kp}, \quad (2.1.15)$$

where the transition relations between p, q and i, j, k, l are listed in Table 2.1 (IEEE Standards on Piezoelectricity, 1987).

Table 2.1 Matrix notation

ij or kl	p or q
11	1
22	2
33	3
23 or 32	4
13 or 31	5
12 or 21	6

Eqs. (2.1.13) can thus be transformed as

$$\begin{aligned}\sigma_p &= C_{pq}^E \varepsilon_q - e_{kp} E_k, \\ D_i &= e_{iq} \varepsilon_q + \alpha_{ik}^E E_k,\end{aligned}\tag{2.1.16}$$

where

$$\varepsilon_p = \begin{cases} \varepsilon_{ij}, & i = j \\ 2\varepsilon_{ij}, & i \neq j \end{cases}\tag{2.1.17}$$

2.1.2.1 Constitutive Equations in Different Forms

As shown in Eq. (2.1.9a)–(2.1.9g), there are in total seven possible energy descriptions. Every energy description will give a set of different yet interrelated constitutive equations. For the four free energy descriptions (temperature T rather than entropy S is selected as the principal variable), and for isothermal processes, the linear constitutive equations and corresponding energy expressions are listed in Table 2.2.

Table 2.2 Constitutive equations and their corresponding thermodynamic potentials

Independent variable set	Piezoelectric constitutive equations	Thermodynamic potentials
(ε_{ij}, D_i)	$\sigma_{ij} = C_{ijkl}^D \varepsilon_{kl} - h_{kij} D_k,$ $E_i = -h_{ikl} \varepsilon_{kl} + \beta_{ik}^E D_k$	Helmholtz free energy $\psi(\varepsilon_{ij}, D_i) = \frac{1}{2} C_{ijkl}^D \varepsilon_{kl} \varepsilon_{ij} - h_{kij} D_k \varepsilon_{ij} + \frac{1}{2} \beta_{ik}^E D_i D_k$
		Gibbs free energy $G(\sigma_{ij}, E_i) = -\frac{1}{2} s_{ijkl}^E \sigma_{kl} \sigma_{ij} - d_{kij} E_k \sigma_{ij} - \frac{1}{2} \alpha_{ik}^E E_i E_k$ ($G = \psi - \sigma_{ij} \varepsilon_{ij} - E_i D_i$)
(σ_{ij}, D_i)	$\varepsilon_{ij} = s_{ijkl}^D \sigma_{kl} + g_{kij} D_k,$ $E_i = -g_{ikl} \sigma_{kl} + \beta_{ik}^E D_k$	Elastic Gibbs free energy $G_1(\sigma_{ij}, D_i) = -\frac{1}{2} s_{ijkl}^D \sigma_{kl} \sigma_{ij} - g_{kij} D_k \sigma_{ij} - \frac{1}{2} \beta_{ik}^E D_i D_k$ ($G_1 = \psi - \sigma_{ij} \varepsilon_{ij}$)
		Electric Gibbs free energy $G_2(\varepsilon_{ij}, E_i) = -\frac{1}{2} C_{ijkl}^E \varepsilon_{kl} \varepsilon_{ij} - e_{kij} E_k \varepsilon_{ij} - \frac{1}{2} \alpha_{ik}^E E_i E_k$ ($G_2 = \psi - E_i D_i$)

Alternative constitutive equations and thermodynamic potentials expressed with the electric polarization P_i instead of D_i are presented with details in (Smith, 2005). Using the matrix notation in Table 2.1, the various coefficients presented in Table 2.2 are related by

$$\begin{aligned}
 C_{pr}^E s_{qr}^E &= \delta_{pq}, & C_{pr}^D s_{qr}^D &= \delta_{pq}, \\
 \alpha_{ik}^\varepsilon \beta_{ik}^\varepsilon &= \delta_{pq}, & \alpha_{ik}^\sigma \beta_{ik}^\sigma &= \delta_{pq}, \\
 C_{pq}^D - C_{pq}^E &= e_{kp} h_{kq}, & s_{pq}^D - s_{pq}^E &= -d_{kp} g_{kq}, \\
 \alpha_{ij}^\sigma - \alpha_{ij}^\varepsilon &= d_{iq} e_{jq}, & \beta_{ij}^\sigma - \beta_{ij}^\varepsilon &= -g_{iq} h_{jq}, \\
 e_{ip} &= d_{iq} C_{qp}^E, & d_{ip} &= \alpha_{ik}^\sigma g_{kp}, \\
 g_{ip} &= \beta_{ik}^\sigma d_{kp}, & h_{ip} &= g_{iq} C_{qp}^D.
 \end{aligned} \tag{2.1.18}$$

For transformation equations between s_{pq}^E and s_{ijkl}^E , d_{ip} and d_{ikl} , etc., one can refer to IEEE Standards on Piezoelectricity (1987).

2.1.2.2 Material Constant Matrices for Different Crystallographic Point Groups

Most piezoelectric materials of interest for technological applications are crystalline solids. Therefore, instead of a general expression based on anisotropic elasticity, we are going to provide specific information of the material constants for different crystallographic point groups that enable the existence of piezoelectricity. All the results are included in Table 2.3 (IEEE Standards on Piezoelectricity, 1987).

2.1.3 The Equation of Equilibrium

The continuum equation of motion for an elastic body is given by

$$\sigma_{ij,i} + f_j = \rho \ddot{u}_j, \tag{2.1.19}$$

where

$$\sigma_{ij,i} = \frac{\partial \sigma_{1j}}{\partial x_1} + \frac{\partial \sigma_{2j}}{\partial x_2} + \frac{\partial \sigma_{3j}}{\partial x_3} \quad (j = 1, 2, 3). \tag{2.1.20}$$

For static problems without body forces f_j , Eq. (2.1.19) becomes

$$\sigma_{ij,i} = 0. \tag{2.1.21}$$

Table 2.3 Material constants matrices for different crystallographic point groups

Crystal system	Point groups	Piezoelectric matrices	Elastic matrices	Dielectric matrices	Example materials
Triclinic	1	$e = \begin{bmatrix} e_{11} & e_{12} & e_{13} & e_{14} & e_{15} & e_{16} \\ e_{21} & e_{22} & e_{23} & e_{24} & e_{25} & e_{26} \\ e_{31} & e_{32} & e_{33} & e_{34} & e_{35} & e_{36} \end{bmatrix}$	$C^E = \begin{bmatrix} C_{11}^E & C_{12}^E & C_{13}^E & C_{14}^E & C_{15}^E & C_{16}^E \\ C_{12}^E & C_{22}^E & C_{23}^E & C_{24}^E & C_{25}^E & C_{26}^E \\ C_{13}^E & C_{23}^E & C_{33}^E & C_{34}^E & C_{35}^E & C_{36}^E \\ C_{14}^E & C_{24}^E & C_{34}^E & C_{44}^E & C_{45}^E & C_{46}^E \\ C_{15}^E & C_{25}^E & C_{35}^E & C_{45}^E & C_{55}^E & C_{56}^E \\ C_{16}^E & C_{26}^E & C_{36}^E & C_{46}^E & C_{56}^E & C_{66}^E \end{bmatrix}$ sym.	$u^d = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{22} & \alpha_{23} & \alpha_{33} \end{bmatrix}$ sym.	$\text{C}_3\text{H}_7\text{-O}_2\text{N}_2$
	2	$e = \begin{bmatrix} 0 & 0 & 0 & e_{14} & 0 & e_{16} \\ e_{21} & e_{22} & e_{23} & 0 & e_{25} & 0 \\ 0 & 0 & 0 & e_{34} & 0 & e_{36} \end{bmatrix}$	$C^E = \begin{bmatrix} C_{11}^E & C_{12}^E & C_{13}^E & 0 & C_{15}^E & 0 \\ C_{12}^E & C_{22}^E & C_{23}^E & 0 & C_{25}^E & 0 \\ C_{13}^E & C_{23}^E & C_{33}^E & 0 & C_{35}^E & 0 \\ C_{15}^E & C_{25}^E & C_{35}^E & C_{44}^E & C_{55}^E & 0 \\ C_{16}^E & 0 & 0 & 0 & 0 & C_{66}^E \end{bmatrix}$ sym.	$u^d = \begin{bmatrix} \alpha_{11} & 0 & \alpha_{13} \\ \alpha_{22} & \alpha_{23} & 0 \\ \alpha_{33} & \alpha_{35} & \alpha_{36} \end{bmatrix}$ sym.	$\text{C}_6\text{H}_4\text{N}_2\text{O}_6$
Monoclinic	m	$e = \begin{bmatrix} e_{11} & e_{12} & e_{13} & 0 & e_{15} & 0 \\ 0 & 0 & 0 & 0 & e_{24} & 0 & e_{26} \\ e_{31} & e_{32} & e_{33} & 0 & e_{35} & 0 \end{bmatrix}$	$C^E = \begin{bmatrix} C_{11}^E & C_{12}^E & C_{13}^E & 0 & C_{15}^E & 0 \\ C_{12}^E & C_{22}^E & C_{23}^E & 0 & C_{25}^E & 0 \\ C_{13}^E & C_{23}^E & C_{33}^E & 0 & C_{35}^E & 0 \\ C_{15}^E & C_{25}^E & C_{35}^E & C_{44}^E & C_{55}^E & 0 \\ C_{16}^E & 0 & 0 & 0 & 0 & C_{66}^E \end{bmatrix}$ sym.	$u^d = \begin{bmatrix} \alpha_{11} & 0 & \alpha_{13} \\ \alpha_{22} & \alpha_{23} & 0 \\ \alpha_{33} & \alpha_{35} & \alpha_{36} \end{bmatrix}$ sym.	$\text{LiH}_2(\text{SeO}_3)_2$
Orthorhombic	222	$e = \begin{bmatrix} 0 & 0 & 0 & e_{14} & 0 & 0 \\ 0 & 0 & 0 & 0 & e_{25} & 0 \\ 0 & 0 & 0 & 0 & 0 & e_{36} \end{bmatrix}$	$C^E = \begin{bmatrix} C_{11}^E & C_{12}^E & C_{13}^E & 0 & 0 & 0 \\ C_{12}^E & C_{22}^E & C_{23}^E & 0 & 0 & 0 \\ C_{13}^E & C_{23}^E & C_{33}^E & 0 & 0 & 0 \\ C_{14}^E & C_{24}^E & C_{34}^E & C_{44}^E & C_{45}^E & C_{46}^E \\ C_{15}^E & C_{25}^E & C_{35}^E & C_{45}^E & C_{55}^E & C_{56}^E \\ C_{16}^E & C_{26}^E & C_{36}^E & C_{46}^E & C_{56}^E & C_{66}^E \end{bmatrix}$ sym.	$u^d = \begin{bmatrix} \alpha_{11} & 0 & 0 \\ \alpha_{22} & \alpha_{23} & 0 \\ \alpha_{33} & \alpha_{35} & \alpha_{36} \end{bmatrix}$ sym.	Rochelle salt
	2mm	$e = \begin{bmatrix} 0 & 0 & 0 & 0 & e_{34} & 0 \\ 0 & 0 & 0 & 0 & e_{23} & 0 \\ e_{31} & e_{32} & e_{33} & 0 & 0 & 0 \end{bmatrix}$	$C^E = \begin{bmatrix} C_{11}^E & C_{12}^E & C_{13}^E & 0 & 0 & 0 \\ C_{12}^E & C_{22}^E & C_{23}^E & 0 & 0 & 0 \\ C_{13}^E & C_{23}^E & C_{33}^E & 0 & 0 & 0 \\ C_{14}^E & C_{24}^E & C_{34}^E & C_{44}^E & C_{45}^E & C_{46}^E \\ C_{15}^E & C_{25}^E & C_{35}^E & C_{45}^E & C_{55}^E & C_{56}^E \\ C_{16}^E & C_{26}^E & C_{36}^E & C_{46}^E & C_{56}^E & C_{66}^E \end{bmatrix}$ sym.	$u^d = \begin{bmatrix} \alpha_{11} & 0 & 0 \\ \alpha_{22} & \alpha_{23} & 0 \\ \alpha_{33} & \alpha_{35} & \alpha_{36} \end{bmatrix}$ sym.	$\text{Ba}_2\text{NaNb}_2\text{O}_{15}$

(To be continued)

TABLE 2.3)

Crystal system	Point groups	Piezoelectric matrices	Elastic matrices	Dielectric matrices	Example materials
Orthorhombic	4	$\mathbf{e} = \begin{bmatrix} 0 & 0 & 0 & e_{14} & e_{15} & 0 \\ 0 & 0 & 0 & 0 & -e_{15} & 0 \\ e_{31} & e_{32} & e_{33} & 0 & 0 & 0 \end{bmatrix}$	$\mathbf{C}^E = \begin{bmatrix} C_{11}^E & C_{12}^E & C_{13}^E & 0 & 0 & 0 \\ C_{12}^E & C_{22}^E & C_{23}^E & 0 & 0 & 0 \\ C_{13}^E & C_{23}^E & C_{33}^E & 0 & 0 & 0 \\ \text{sym.} & & & C_{44}^E & 0 & 0 \\ & & & C_{55}^E & 0 & 0 \\ & & & & C_{66}^E & 0 \end{bmatrix}$	$\mathbf{a}^E = \begin{bmatrix} \alpha_{11} & 0 & 0 \\ \text{sym.} & \alpha_{22} & 0 \\ & & \alpha_{33} \end{bmatrix}$	KSr ₂ NbO ₁₅
	422	$\mathbf{e} = \begin{bmatrix} 0 & 0 & 0 & e_{14} & 0 & 0 \\ 0 & 0 & 0 & 0 & -e_{14} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$	$\mathbf{C}^E = \begin{bmatrix} C_{11}^E & C_{12}^E & C_{13}^E & 0 & 0 & 0 \\ C_{12}^E & C_{22}^E & C_{23}^E & 0 & 0 & 0 \\ C_{13}^E & C_{23}^E & C_{33}^E & 0 & 0 & 0 \\ \text{sym.} & & & C_{44}^E & 0 & 0 \\ & & & C_{55}^E & 0 & 0 \\ & & & & C_{66}^E & 0 \end{bmatrix}$	$\mathbf{a}^E = \begin{bmatrix} \alpha_{11} & 0 & 0 \\ \text{sym.} & \alpha_{22} & 0 \\ & & \alpha_{33} \end{bmatrix}$	TeO ₂
	4mm	$\mathbf{e} = \begin{bmatrix} 0 & 0 & 0 & 0 & e_{15} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ e_{31} & e_{32} & e_{33} & 0 & 0 & 0 \end{bmatrix}$	$\mathbf{C}^E = \begin{bmatrix} C_{11}^E & C_{12}^E & C_{13}^E & 0 & 0 & 0 \\ C_{12}^E & C_{22}^E & C_{23}^E & 0 & 0 & 0 \\ C_{13}^E & C_{23}^E & C_{33}^E & 0 & 0 & 0 \\ \text{sym.} & & & C_{44}^E & 0 & 0 \\ & & & C_{55}^E & 0 & 0 \\ & & & & C_{66}^E & 0 \end{bmatrix}$	$\mathbf{a}^E = \begin{bmatrix} \alpha_{11} & 0 & 0 \\ \text{sym.} & \alpha_{22} & 0 \\ & & \alpha_{33} \end{bmatrix}$	BiTiO ₃
	$\bar{4}2m$	$\mathbf{e} = \begin{bmatrix} 0 & 0 & 0 & e_{14} & 0 & 0 \\ 0 & 0 & 0 & 0 & e_{14} & 0 \\ 0 & 0 & 0 & 0 & 0 & e'_{16} \end{bmatrix}$	$\mathbf{C}^E = \begin{bmatrix} C_{11}^E & C_{12}^E & C_{13}^E & 0 & 0 & 0 \\ C_{12}^E & C_{22}^E & C_{23}^E & 0 & 0 & 0 \\ C_{13}^E & C_{23}^E & C_{33}^E & 0 & 0 & 0 \\ \text{sym.} & & & C_{44}^E & 0 & 0 \\ & & & C_{55}^E & 0 & 0 \\ & & & & C_{66}^E & 0 \end{bmatrix}$	$\mathbf{a}^E = \begin{bmatrix} \alpha_{11} & 0 & 0 \\ \text{sym.} & \alpha_{22} & 0 \\ & & \alpha_{33} \end{bmatrix}$	NaIO ₃ ·3H ₂ O
Trigonal	3	$\mathbf{e} = \begin{bmatrix} e_{11} & -e_{11} & 0 & e_{14} & e_{15} & -e_{22} \\ -e'_{22} & e_{22} & 0 & e_{15} & -e_{14} & -e'_{11} \\ e_{31} & e'_{31} & e'_{33} & 0 & 0 & 0 \end{bmatrix}$	$\mathbf{C}^E = \begin{bmatrix} C_{11}^E & C_{12}^E & C_{13}^E & C_{14}^E & -C_{15}^E & 0 \\ C_{12}^E & C_{22}^E & C_{23}^E & -C_{14}^E & C_{15}^E & 0 \\ C_{13}^E & C_{23}^E & C_{33}^E & 0 & 0 & 0 \\ \text{sym.} & & & C_{44}^E & 0 & C_{45}^E \\ & & & C_{45}^E & C_{44}^E & C_{66}^E \\ & & & & C_{66}^E & C_{66}^E \end{bmatrix}$	$\mathbf{a}^E = \begin{bmatrix} \alpha_{11} & 0 & 0 \\ \text{sym.} & \alpha_{22} & 0 \\ & & \alpha_{33} \end{bmatrix}$	

where $C_{66}^E = \frac{1}{2}(C_{11}^E - C_{33}^E)$

To be continued

TABLE 2.3

Crystal system	Point group	Piezoelectric matrices	Elastic matrices	Dielectric matrices	Example materials
Trigonal	32	$\mathbf{e} = \begin{bmatrix} e_{11} & -e_{11} & 0 & e_{14} & 0 & 0 \\ 0 & 0 & 0 & 0 & -e_{14} & -e_{11} \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$	$\mathbf{C}^E = \begin{bmatrix} C_{11}^E & C_{12}^E & C_{13}^E & C_{14}^E & -C_{23}^E & 0 \\ C_{12}^E & C_{11}^E & C_{13}^E & -C_{14}^E & C_{23}^E & 0 \\ C_{13}^E & C_{13}^E & C_{33}^E & 0 & 0 & 0 \\ C_{14}^E & -C_{14}^E & 0 & C_{44}^E & 0 & 0 \\ -C_{23}^E & C_{23}^E & 0 & 0 & C_{55}^E & 0 \\ 0 & 0 & 0 & 0 & 0 & C_{66}^E \end{bmatrix}$	$\boldsymbol{\alpha}^E = \begin{bmatrix} \alpha_{11} & 0 & 0 \\ \text{sym.} & \alpha_{22} & 0 \\ & & \alpha_{33} \end{bmatrix}$	α -quartz
	3m	$\mathbf{e} = \begin{bmatrix} 0 & 0 & 0 & 0 & e_{15} & -e_{22} \\ -e_{22} & e_{22} & 0 & e_{15} & 0 & 0 \\ e_{31} & e_{31} & e_{33} & 0 & 0 & 0 \end{bmatrix}$	$\mathbf{C}^E = \begin{bmatrix} \text{sym.} \\ \text{sym.} \\ \text{sym.} \end{bmatrix}$ where $C_{66}^E = \frac{1}{2}(C_{11}^E - C_{12}^E)$	$\boldsymbol{\alpha}^E = \begin{bmatrix} \alpha_{11} & 0 & 0 \\ \text{sym.} & \alpha_{22} & 0 \\ & & \alpha_{33} \end{bmatrix}$	LiNbO ₃
Hexagonal	6	$\mathbf{e} = \begin{bmatrix} 0 & 0 & 0 & e_{14} & e_{15} & 0 \\ 0 & 0 & 0 & e_{15} & -e_{14} & 0 \\ e_{31} & e_{31} & e_{33} & 0 & 0 & 0 \end{bmatrix}$	$\mathbf{C}^E = \begin{bmatrix} C_{11}^E & C_{12}^E & C_{13}^E & 0 & 0 & 0 \\ C_{12}^E & C_{11}^E & C_{13}^E & 0 & 0 & 0 \\ C_{13}^E & C_{13}^E & C_{33}^E & 0 & 0 & 0 \\ 0 & 0 & 0 & C_{44}^E & 0 & 0 \\ 0 & 0 & 0 & 0 & C_{44}^E & 0 \\ \text{sym.} & \text{sym.} & \text{sym.} & 0 & 0 & C_{66}^E \end{bmatrix}$	$\boldsymbol{\alpha}^E = \begin{bmatrix} \alpha_{11} & 0 & 0 \\ \text{sym.} & \alpha_{22} & 0 \\ & & \alpha_{33} \end{bmatrix}$	LiIO ₃
	$\bar{6}$	$\mathbf{e} = \begin{bmatrix} e_{11} & -e_{11} & 0 & 0 & 0 & -e_{22} \\ -e_{22} & e_{22} & 0 & 0 & 0 & -e_{11} \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$	$\mathbf{C}^E = \begin{bmatrix} \text{sym.} \\ \text{sym.} \\ \text{sym.} \end{bmatrix}$ where $C_{66}^E = \frac{1}{2}(C_{11}^E - C_{12}^E)$		β -quartz
6mm	622	$\mathbf{e} = \begin{bmatrix} 0 & 0 & 0 & e_{13} & 0 & 0 \\ 0 & 0 & 0 & 0 & e_{14} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$			
	6mm	$\mathbf{e} = \begin{bmatrix} 0 & 0 & 0 & 0 & e_{15} & 0 \\ 0 & 0 & 0 & 0 & e_{25} & 0 \\ e_{31} & e_{31} & e_{33} & 0 & 0 & 0 \end{bmatrix}$			CdS

(To be continued)

(Table 2.3)

Crystal system	Point groups	Piezoelectric matrices	Elastic matrices	Dielectric matrices	Example materials
Hexagonal	$\bar{6}m2$	$\mathbf{e} = \begin{bmatrix} e_{11} & -e_{11} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -e_{11} \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$	$\mathbf{C}^E = \begin{bmatrix} C_{11}^E & C_{12}^E & C_{13}^E & 0 & 0 & 0 \\ C_{12}^E & C_{11}^E & C_{13}^E & 0 & 0 & 0 \\ C_{13}^E & C_{13}^E & C_{33}^E & 0 & 0 & 0 \\ \text{sym.} & & & C_{44}^E & 0 & 0 \\ & & & & C_{44}^E & 0 \\ & & & & & C_{66}^E \end{bmatrix}$	$\mathbf{a}^E = \begin{bmatrix} \alpha_{11} & 0 & 0 \\ \text{sym.} & \alpha_{22} & 0 \\ & & \alpha_{33} \end{bmatrix}$	
where $C_{66}^E = \frac{1}{2}(C_{11}^E - C_{12}^E)$					
Cubic	23 $\bar{4}3m$	$\mathbf{e} = \begin{bmatrix} 0 & 0 & 0 & e_{14} & 0 & 0 \\ 0 & 0 & 0 & 0 & e_{14} & 0 \\ 0 & 0 & 0 & 0 & 0 & e_{14} \end{bmatrix}$	$\mathbf{C}^E = \begin{bmatrix} C_{11}^E & C_{12}^E & C_{12}^E & 0 & 0 & 0 \\ C_{12}^E & C_{11}^E & C_{12}^E & 0 & 0 & 0 \\ C_{12}^E & C_{12}^E & C_{11}^E & 0 & 0 & 0 \\ \text{sym.} & & & C_{44}^E & 0 & 0 \\ & & & & C_{44}^E & 0 \\ & & & & & C_{44}^E \end{bmatrix}$	$\mathbf{a}^E = \begin{bmatrix} \alpha_{11} & 0 & 0 \\ \text{sym.} & \alpha_{11} & 0 \\ & & \alpha_{11} \end{bmatrix}$	Bi ₁₂ GeO ₂₀

For small deformation problems, the deformation compatibility equation is given by

$$\varepsilon_{ij} = \frac{1}{2}(\mathbf{u}_{i,j} + \mathbf{u}_{j,i}) . \quad (2.1.22)$$

2.1.4 *The Basic Equations of a Static Electric Field*

The two equations determining the quasistatic electric field in dielectrics are (Landau and Lifshitz, 1960; Tiersten, 1969)

$$\varphi_{,i} = -E_i , \quad (2.1.23)$$

and

$$D_{i,i} = 4\pi\rho_{ex} . \quad (2.1.24)$$

In Eq. (2.1.23) φ is introduced as the scalar potential of the electric field. ρ_{ex} in Eq. (2.1.24) denotes extraneous body charge density (Landau and Lifshitz, 1960). When the extraneous body charge is absent, Eq. (2.1.24) reduces to

$$D_{i,i} = 0 . \quad (2.1.25)$$

The two equations are obtained from an approximation of Poynting's theory of electromagnetic energy flux (Tiersten, 1969).

2.1.5 *Differential Equations for Piezoelectric Materials*

Substitution of Eqs. (2.1.13), (2.1.22) and (2.1.23) into Eqs. (2.1.19) and (2.1.24) gives

$$\begin{aligned} C_{ijkl}^E \mathbf{u}_{k,li} + e_{kij} \varphi_{,ki} + f_j &= \rho \ddot{\mathbf{u}}_j , \\ e_{ikl} \mathbf{u}_{k,li} - \alpha_{ik}^E \varphi_{,ki} &= 4\pi\rho_{ex} . \end{aligned} \quad (2.1.26)$$

For a static problem without body force or extraneous body charges, Eq. (2.1.26) reduces to

$$\begin{aligned} C_{ijkl}^E \mathbf{u}_{k,li} + e_{kij} \varphi_{,ki} &= 0 , \\ e_{ikl} \mathbf{u}_{k,li} - \alpha_{ik}^E \varphi_{,ki} &= 0 . \end{aligned} \quad (2.1.27)$$

Eq. (2.1.27) is widely seen in various literature dealing with mechanical and electrostatic problems of piezoelectric materials. Due to the similar structure of the relations between elastic quantities and electric quantities, a generalized equation formalism is constructed as an alternative expression of Eq. (2.1.27):

$$C_{iJKl}u_{K,li} = 0, \quad (2.1.28)$$

where uppercase subscripts vary from 1 to 4, and lowercase subscripts vary from 1 to 3. The additional physical quantities appearing in Eq. (2.1.28) are defined as

$$u_4 = \varphi, \quad (2.1.29)$$

and

$$C_{iJKl} = \begin{cases} C_{ijkl}^E, & J, K = 1, 2, 3 \\ e_{lij}, & J = 1, 2, 3; K = 4 \\ e_{ikl}, & J = 4; K = 1, 2, 3 \\ -\alpha_{ii}^E, & J, K = 4 \end{cases}. \quad (2.1.30)$$

The generalized stresses and strains associated with such a formulation are given by

$$\sigma_{iJ} = \begin{cases} \sigma_{ij}, & J = 1, 2, 3 \\ D_i, & J = 4 \end{cases}, \quad \varepsilon_{ij} = \begin{cases} \varepsilon_{ij}, & I = 1, 2, 3 \\ -E_j, & I = 4 \end{cases}. \quad (2.1.31)$$

By solving Eq. (2.1.27) or Eq. (2.1.28), together with various kinds of boundary conditions, the elastic and electric fields are obtained.

2.2 Boundary Conditions

The boundary conditions on an interface S separating two kinds of piezoelectric materials are given by

$$\sigma_{ij}^{(1)} n_j = \sigma_{ij}^{(2)} n_j, \quad (2.2.1)$$

$$u_i^{(1)} = u_i^{(2)}, \quad (2.2.2)$$

$$D_i^{(1)} n_i = D_i^{(2)} n_i, \quad (2.2.3)$$

$$\varphi^{(1)} = \varphi^{(2)}. \quad (2.2.4)$$

In Eqs. (2.2.1)–(2.2.4), variables with superscript (1) indicate physical quantities on one side of the interface S and variables with superscript (2) indicate physical quantities on the other side. n_j denotes the component of the unit normal to the interface S . For a traction-restricted surface, Eq. (2.2.1) becomes

$$\sigma_{ij}n_j = t_i^0, \quad (2.2.5)$$

where t_i^0 is given on the surface. For a displacement-restricted surface, Eq. (2.2.2) becomes

$$u_i = u_i^0, \quad (2.2.6)$$

where u_i^0 is given on the surface. It should be emphasized that for a specific type of surface, either Eq. (2.2.5) or Eq. (2.2.6) describes its mechanical condition. The validity of Eq. (2.2.3) requires that the surface charge density on S should vanish. For an interface S separating a dielectric and a conductor, Eqs. (2.2.3) and (2.2.4) become

$$D_i n_i = 4\pi\rho_s, \quad (2.2.7)$$

$$\varphi = \varphi_0, \quad (2.2.8)$$

where ρ_s is the charge density on the surface of the conductor and φ_0 is the electric potential of the conductor (Landau and Lifshitz, 1960). If the electric potential on the conductor is given as zero, Eq. (2.2.8) becomes

$$\varphi = 0. \quad (2.2.9)$$

Eq. (2.2.9) can also be interpreted as the electric boundary condition for short-circuited electrodes. The electric boundary condition for piezoelectric materials on the surface (or on an interface between the piezoelectric material and air) is of particular importance. Due to the fact that a dramatic change in the dielectric constants (about three orders of magnitude) occurs across the boundary, Deeg (1980) and Pak (1990a) assumed an approximate boundary condition for piezoelectric materials on the surface

$$D_i n_i = 0. \quad (2.2.10)$$

Eq. (2.2.10) is widely applied in theoretical studies of defects, such as cracks, voids, etc., in piezoelectric materials. For the validity of Eq. (2.2.10), it is assumed that no external charge resides on the crack surface. As summarized in Suo's work (Suo et al., 1992) on fracture mechanics for piezoelectric ceramics, the application of Eq. (2.2.10) to the analysis of defects may need further discussion in some specific case. For a small crack opening, McMeeking (1989) pointed out that for a Griffith crack simulated as an elliptic flaw with lower permittivity, the parameter

$\gamma = (\alpha_f / \alpha_m)(a/b)$ determined the validity of Eq. (2.2.10), where α_f and α_m denoted the permittivity of the flaw and the matrix, respectively, and a and b were the semi-axes of the ellipse ($a > b$). When γ is small enough, Eq. (2.2.10) can be applied. A quick glance at the definition of γ shows that even if the permittivity ratio is small, for a flat elliptic flaw ($a \gg b$), γ may not be small. When the crack is filled with conducting substances, the electric boundary condition becomes

$$\varphi^+ = \varphi^-, \quad (2.2.11)$$

where φ^+ and φ^- denote the electric potential on the two sides of the crack.

2.3 Solution Methods for Two-Dimensional Problems

For a piezoelectric body described by a rectangular Cartesian coordinate (x_1, x_2, x_3) , if the displacement fields u_i and the electric potential φ at any point of the body are independent of one component of the coordinate, e.g., x_3 , such a problem is called a two-dimensional problem. In anisotropic elasticity, two methods are mainly used for solving two-dimensional problems, namely the Stroh formalism and the Lekhnitskii formalism. A comprehensive review of the application of these two methods in anisotropic elasticity was given by Ting (1996). An extension of the use of these two methods into problems of piezoelectric materials is presented here.

2.3.1 The Stroh Formalism for Piezoelectric Materials

In 1958, Stroh constructed a complex functional method to solve the two-dimensional problems of cracks and dislocations in anisotropic elastic bodies. Later, he discussed the possibility of the application of the method to the problem of wave propagation in three-dimensional anisotropic elastic bodies (Stroh, 1962). Yet the essence of the Stroh formalism is two-dimensional restricted. In its application to the wave propagation problem, two characteristic directions are captured: the propagation direction and the oscillation direction. The various properties of the Stroh formalism were developed by Ting et al. (1988; 1992; 1994; 1996).

The Stroh formalism was first extended to treat elastic problems for piezoelectric materials by Barnett and Lothe (1975). Due to the mathematical elegance and technical effectiveness, the method is widely accepted in solving all

kinds of problems associated with piezoelectric materials (Deeg, 1980; Suo et al., 1992; Du et al., 1994; Liang et al., 1995; Liu et al., 1997a; b; Chung, 1995; Chung and Ting, 1996; Pak, 1992; Park and Sun, 1995; Lu et al., 2006). The construction of the Stroh formalism aims at treating the mechanical equilibrium equations independently. The solution obtained from the Stroh formalism satisfies the mechanical equilibrium equations for any configuration of boundaries spontaneously. The main part of the formalism is presented as follows.

For the coupled two-dimensional problem of the electroelastic field, suppose that the generalized displacement fields u_K , as introduced in Section 2.1.5, are merely functions of the components of coordinates x_1, x_2 . The expression of u_K can be introduced as

$$u_K = f_K(x_1, x_2) . \quad (2.3.1)$$

For the very first step of the Stroh formalism, we assume that

$$f_K(x_1, x_2) = a_K f(Z), \quad (2.3.2)$$

where

$$Z = x_1 + px_2 \quad (2.3.3)$$

and p, a_K are complex coefficients to be determined. When Eqs. (2.3.1) and (2.3.2) are substituted into Eq. (2.1.20), the following equation is obtained after manipulation

$$[\mathbf{Q} + p(\mathbf{R} + \mathbf{R}^T) + p^2\mathbf{T}]\mathbf{a}f'(Z) = 0, \quad (2.3.4)$$

where superscript T stands for matrix transpose, $f'(Z) = \frac{\partial f(Z)}{\partial Z}$ and $\mathbf{a} = [a_1, a_2, a_3, a_4]^T$.

$\mathbf{Q}, \mathbf{R}, \mathbf{T}$ are given respectively as

$$\mathbf{Q} = \begin{bmatrix} C_{1jk1}^E & e_{1j1} \\ e_{1k1}^T & -\alpha_{11}^\varepsilon \end{bmatrix}, \quad \mathbf{R} = \begin{bmatrix} C_{1jk2}^E & e_{2j1} \\ e_{1k2}^T & -\alpha_{12}^\varepsilon \end{bmatrix}, \quad \mathbf{T} = \begin{bmatrix} C_{2jk2}^E & e_{2j2} \\ e_{2k2}^T & -\alpha_{22}^\varepsilon \end{bmatrix}. \quad (2.3.5)$$

From Eq. (2.3.4) we have

$$\det[\mathbf{Q} + p(\mathbf{R} + \mathbf{R}^T) + p^2\mathbf{T}] = 0, \quad (2.3.6)$$

which determines four pairs of complex conjugate p , denoted by

$$p_{\alpha+4} = \bar{p}_\alpha, \quad \text{Im}(p_\alpha) > 0 \quad (\alpha = 1, 2, 3, 4), \quad (2.3.7)$$

where a top-imposed bar indicates the complex conjugate. Let us introduce the generalized stress potential vector as

$$\boldsymbol{\phi} = \mathbf{b}f(Z), \quad (2.3.8)$$

where

$$\sigma_{i1} = -\phi_{i,2}, \quad \sigma_{i2} = \phi_{i,1}, \quad D_1 = -\phi_{4,2}, \quad D_2 = \phi_{4,1}. \quad (2.3.9)$$

From Eq. (2.3.8) and the constitutive equations we can obtain that

$$\mathbf{b} = (\mathbf{R}^T + p\mathbf{T})\mathbf{a} = \frac{-1}{p}(\mathbf{Q} + p\mathbf{R})\mathbf{a}. \quad (2.3.10)$$

The generalized displacement vector $\mathbf{u}$ and the generalized stress potential vector $\boldsymbol{\phi}$ can then be expressed as

$$\mathbf{u} = 2 \operatorname{Re} \left\{ \sum_{\alpha=1}^4 \mathbf{a}_{\alpha} f_{\alpha}(Z_{\alpha}) \right\}, \quad \boldsymbol{\phi} = 2 \operatorname{Re} \left\{ \sum_{\alpha=1}^4 \mathbf{b}_{\alpha} f_{\alpha}(Z_{\alpha}) \right\}, \quad (2.3.11)$$

where $Z_{\alpha} = x_1 + p_{\alpha}x_2$ and

$$f_{\alpha}(Z_{\alpha}) = q_{\alpha}f(Z_{\alpha}) \quad (\text{not summed for } \alpha). \quad (2.3.12)$$

Eq. (2.3.10) can be rewritten as

$$\mathbf{u} = 2 \operatorname{Re} \{ \mathbf{A} < f(Z_{\alpha}) > \mathbf{q} \}, \quad \boldsymbol{\phi} = 2 \operatorname{Re} \{ \mathbf{B} < f(Z_{\alpha}) > \mathbf{q} \}, \quad (2.3.13)$$

where $\mathbf{q} = [q_1, q_2, q_3, q_4]^T$ are parameters to be determined by the boundary conditions and

$$\mathbf{A} = [\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \mathbf{a}_4], \quad \mathbf{B} = [\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3, \mathbf{b}_4], \quad (2.3.14)$$

$$< f(Z_{\alpha}) > = \begin{bmatrix} < f(Z_1) > & 0 & 0 & 0 \\ 0 & < f(Z_2) > & 0 & 0 \\ 0 & 0 & < f(Z_3) > & 0 \\ 0 & 0 & 0 & < f(Z_4) > \end{bmatrix}. \quad (2.3.15)$$

$\mathbf{a}_{\alpha}$ ($\alpha = 1, 2, 3, 4$) is the eigenvector corresponding to p_{α} in solving Eq. (2.3.13) and $\mathbf{b}_{\alpha}$ ($\alpha = 1, 2, 3, 4$) is obtained from Eq. (2.3.10) with $\mathbf{a}_{\alpha}$ ($\alpha = 1, 2, 3, 4$). Eq. (2.3.10) can be recast into a matrix form

$$\begin{bmatrix} -\mathbf{R}^T & \mathbf{I} \\ -\mathbf{Q} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix} = p \begin{bmatrix} \mathbf{T} & \mathbf{0} \\ \mathbf{R} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix}. \quad (2.3.16)$$

From Eq. (2.3.5) we learn that $\mathbf{T}^{-1}$ exists, thus leading to a standard eight-dimensional eigen equation

$$\mathbf{N}\xi = p\xi, \quad (2.3.17)$$

where

$$\mathbf{N} = \begin{bmatrix} \mathbf{N}_1 & \mathbf{N}_1 \\ \mathbf{N}_3 & \mathbf{N}_1^T \end{bmatrix}, \quad \xi = \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix}, \quad (2.3.18)$$

$$\mathbf{N}_1 = -\mathbf{T}^{-1}\mathbf{R}^T, \quad \mathbf{N}_2 = -\mathbf{T}^{-1}, \quad \mathbf{N}_3 = \mathbf{R}\mathbf{T}^{-1}\mathbf{R}^T - \mathbf{Q}. \quad (2.3.19)$$

Eq. (2.3.17) gives

$$\det(\mathbf{N} - p\mathbf{I}) = 0. \quad (2.3.20)$$

p solved from Eq. (2.3.20) is equivalent to those defined in Eq. (2.3.14). For every eigenvalue p_α there exists a right eigenvector ξ_α from Eq. (2.3.17). The normalization relation between the right and left eigenvectors generates the following equations

$$\begin{bmatrix} \mathbf{B}^T & \mathbf{A}^T \\ \bar{\mathbf{B}}^T & \bar{\mathbf{A}}^T \end{bmatrix} \begin{bmatrix} \mathbf{A} & \bar{\mathbf{A}} \\ \mathbf{B} & \bar{\mathbf{B}} \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}. \quad (2.3.21)$$

The expansion of Eq. (2.3.21) gives rise to the following identities

$$\mathbf{A}\mathbf{B}^T + \bar{\mathbf{A}}\bar{\mathbf{B}}^T = \mathbf{B}\mathbf{A}^T + \bar{\mathbf{B}}\bar{\mathbf{A}}^T = \mathbf{I}, \quad \mathbf{A}\mathbf{A}^T + \bar{\mathbf{A}}\bar{\mathbf{A}}^T = \mathbf{B}\mathbf{B}^T + \bar{\mathbf{B}}\bar{\mathbf{B}}^T = \mathbf{0}. \quad (2.3.22)$$

From Eq. (2.3.22), it is possible to define three real matrices as follows

$$\mathbf{S} = i(2\mathbf{A}\mathbf{B}^T - \mathbf{I}), \quad \mathbf{H} = i2\mathbf{A}\mathbf{A}^T, \quad \mathbf{L} = -i2\mathbf{B}\mathbf{B}^T, \quad (2.3.23)$$

where $\mathbf{H}$ and $\mathbf{L}$ are proved to be symmetric and nonsingular (Lothe and Barnett, 1976). Due to the useful properties of Eqs. (2.3.22) and (2.3.23), the solution form of Eq. (2.3.20) is often transformed to

$$\begin{aligned} \mathbf{u} &= 2 \operatorname{Re} \left\{ \mathbf{A} < f(Z_\alpha) > \mathbf{A}^T \right\} \mathbf{g} + 2 \operatorname{Re} \left\{ \mathbf{A} < f(Z_\alpha) > \mathbf{B}^T \right\} \mathbf{h}, \\ \phi &= 2 \operatorname{Re} \left\{ \mathbf{B} < f(Z_\alpha) > \mathbf{A}^T \right\} \mathbf{g} + 2 \operatorname{Re} \left\{ \mathbf{B} < f(Z_\alpha) > \mathbf{B}^T \right\} \mathbf{h}, \end{aligned} \quad (2.3.24)$$

where

$$\mathbf{q} = \mathbf{A}^T \mathbf{g} + \mathbf{B}^T \mathbf{h}. \quad (2.3.25)$$

In Eq. (2.3.25) the complex parameter vector $\mathbf{q}$ is replaced by two real vectors $\mathbf{g}$ and $\mathbf{h}$. Eq. (2.3.24) prevails for Eq. (2.3.20) in that the properties of $\mathbf{S}$, $\mathbf{H}$, $\mathbf{L}$ can be used when substituted into the boundary conditions. Eq. (2.3.23) expresses $\mathbf{S}$, $\mathbf{H}$, $\mathbf{L}$ in terms of $\mathbf{A}$ and $\mathbf{B}$, which depend on the solution of the eigen equation Eq. (2.3.24). Yet it was proved by Lothe and Barnett (1976) that $\mathbf{S}$, $\mathbf{H}$, $\mathbf{L}$ could be expressed in an integral form independent of the solution of the eigen equation Eq. (2.3.24). Denote

$$\mathbf{n}(\omega) = [\cos \omega, \sin \omega, 0]^T, \quad \mathbf{m}(\omega) = [-\sin \omega, \cos \omega, 0]^T, \quad (2.3.26)$$

in which ω ranges from 0 to 2π . Transforming Eq. (2.3.19) to an expression that varies with ω , we have

$$\begin{aligned} \mathbf{N}_1(\omega) &= -\mathbf{T}^{-1}(\omega)\mathbf{R}^T(\omega), \\ \mathbf{N}_2(\omega) &= -\mathbf{T}^{-1}(\omega), \\ \mathbf{N}_3(\omega) &= \mathbf{R}(\omega)\mathbf{T}^{-1}(\omega)\mathbf{R}^T(\omega) - \mathbf{Q}(\omega), \end{aligned} \quad (2.3.27)$$

where components of the tensors $\mathbf{Q}(\omega)$, $\mathbf{R}(\omega)$, $\mathbf{T}(\omega)$ are given by

$$\begin{aligned} Q_{JK}(\omega) &= n_i(\omega)C_{iJKl}n_l(\omega), \quad R_{JK}(\omega) = n_i(\omega)C_{iJKl}m_l(\omega), \\ T_{JK}(\omega) &= m_i(\omega)C_{iJKl}m_l(\omega). \end{aligned} \quad (2.3.28)$$

$\mathbf{S}$, $\mathbf{H}$, $\mathbf{L}$ can thus be expressed as

$$\mathbf{S} = \frac{1}{\pi} \int_0^\pi \mathbf{N}_1(\omega) d\omega, \quad \mathbf{H} = \frac{1}{\pi} \int_0^\pi \mathbf{N}_2(\omega) d\omega, \quad \mathbf{L} = -\frac{1}{\pi} \int_0^\pi \mathbf{N}_3(\omega) d\omega. \quad (2.3.29)$$

Eq. (2.3.29) provides an alternative to Eq. (2.3.23) for the Barnett-Lothe tensors $\mathbf{S}$, $\mathbf{H}$, $\mathbf{L}$. If we use the eight-dimensional standard eigen equation to solve the problem, it should be noted that an intrinsic relation exists between the introduced generalized stress potential vector and the generalized displacement vector. If we multiply the term $f'(Z)$ in Eq. (2.3.25) to both sides of Eq. (2.3.24), we have after manipulation

$$\mathbf{N} \begin{bmatrix} \mathbf{u}_{,1} \\ \boldsymbol{\phi}_{,1} \end{bmatrix} = \begin{bmatrix} \mathbf{u}_{,2} \\ \boldsymbol{\phi}_{,2} \end{bmatrix}. \quad (2.3.30)$$

2.3.2 The Lekhnitskii Formalism for Piezoelectric Materials

Lekhnitskii was among the first scholars who systematically treated the problem of elasticity for an anisotropic body (Lekhnitskii, 1968; 1981). A similarity with the Stroh formalism lies in that the Lekhnitskii formalism also started with the solution of the equilibrium equation and had introduced the core function $f(Z_\alpha)$. Yet the differences between these two methods are also obvious. Firstly, the Lekhnitskii formalism is constructed by use of the elastic stress instead of the displacement, and thus constructed by use of the compliance tensor s_{ijkl} rather than the stiffness tensor C_{ijkl} . Secondly, the Lekhnitskii formalism is not expressed in the matrix notation. The sextic eigen equation is lacking in the Lekhnitskii formalism (a modified Lekhnitskii formalism was constructed by Ting (1999) which provided the sextic eigen equation for anisotropic elasticity). Fang and Liu (2009) reconstructed the Lekhnitskii formalism on fracture mechanics of piezoelectric and ferroelectric solids. The main part of the Lekhnitskii formalism is presented here.

For a two-dimensional problem in a Cartesian coordinate, the constitutive equations for the strain and electric field are given by

$$\varepsilon_p = s_{pq}^D \sigma_q + g_{pk} D_k, \quad -E_i = g_{iq} \sigma_q - \beta_{ij}^\sigma D_j \quad (p, q = 1, 2, \dots, 6; i, j = 1, 2, 3). \quad (2.3.31)$$

Since the displacement u_i and the electric potential φ are independent of x_3 , we have

$$\varepsilon_3 = u_{3,3} = 0, \quad E_3 = \varphi_{,3} = 0. \quad (2.3.32)$$

Thus, from Eq. (2.3.31) we have

$$\begin{aligned} s_{33} \sigma_3 + g_{33} D_3 &= -(s_{31} \sigma_1 + s_{32} \sigma_2 + s_{34} \sigma_4 + s_{35} \sigma_5 + s_{36} \sigma_6 + g_{31} D_1 + g_{32} D_2), \\ g_{33} \sigma_3 - \beta_{33}^\sigma D_3 &= -(g_{31} \sigma_1 + g_{32} \sigma_2 + g_{34} \sigma_4 + g_{35} \sigma_5 + g_{36} \sigma_6 - \beta_{31}^\sigma D_1 - \beta_{32}^\sigma D_2), \end{aligned} \quad (2.3.33)$$

where superscripts are omitted for convenience. From Eq. (2.3.33) we can solve

$$\begin{aligned} \sigma_3 &= -\frac{1}{s_{33} \beta_{33} + g_{33}^2} [\sum_{q \neq 3} (\beta_{33} s_{3q} + g_{33} g_{3q}) \sigma_q + \sum_{i \neq 3} (\beta_{33} g_{3i} - g_{33} \beta_{3i}) D_i], \\ D_3 &= -\frac{1}{s_{33} \beta_{33} + g_{33}^2} [\sum_{q \neq 3} (g_{33} s_{3q} - s_{33} g_{3q}) \sigma_q + \sum_{i \neq 3} (s_{33} \beta_{3i} + g_{33} g_{3i}) D_i]. \end{aligned} \quad (2.3.34)$$

By substituting Eq. (2.3.34) into Eq. (2.3.31) and after manipulation, a reduced constitutive equation is obtained for two-dimensional problems

$$\varepsilon_p = s'_{pq} \sigma_q + g'_{pk} D_k, \quad -E_i = g'_{iq} \sigma_q - \beta'_{ij} D_j \quad (p, q = 1, 2, 4, 5, 6; i, j = 1, 2). \quad (2.3.35)$$

where

$$\begin{aligned}
 s'_{pq} &= s_{pq} - \frac{s_{p3}s_{3q}\beta_{33} + s_{p3}g_{3q}g_{33} + s_{3q}g_{3p}g_{33} - s_{33}g_{3p}g_{3q}}{s_{33}\beta_{33} + g_{33}^2}, \\
 g'_{iq} &= g_{iq} - \frac{g_{i3}s_{3q}\beta_{33} - s_{3q}g_{33}\beta_{i3} + g_{i3}g_{3q}g_{33} + s_{33}g_{3q}\beta_{i3}}{s_{33}\beta_{33} + g_{33}^2}, \\
 \beta'_{ij} &= \beta_{ij} + \frac{g_{i3}g_{j3}\beta_{33} - g_{i3}g_{33}\beta_{j3} - g_{j3}g_{33}\beta_{i3} + s_{33}\beta_{i3}\beta_{j3}}{s_{33}\beta_{33} + g_{33}^2}.
 \end{aligned} \tag{2.3.36}$$

The two-dimensional equilibrium equations are given by

$$\begin{aligned}
 \sigma_{1,1} + \sigma_{6,2} &= 0, \\
 \sigma_{6,1} + \sigma_{2,2} &= 0, \\
 \sigma_{5,1} + \sigma_{4,2} &= 0, \\
 D_{1,1} + D_{2,2} &= 0.
 \end{aligned} \tag{2.3.37}$$

Eq. (2.3.37) allows the induction of the following functions

$$\begin{aligned}
 \sigma_1 &= J_{,22}, \quad \sigma_2 = J_{,11}, \quad \sigma_6 = -J_{,12}, \\
 \sigma_5 &= H_{,2}, \quad \sigma_4 = -H_{,1}, \\
 D_1 &= G_{,2}, \quad D_2 = -G_{,1}.
 \end{aligned} \tag{2.3.38}$$

By considering the deformation compatibility equation, Eq. (2.1.22), for two-dimensional problems the strains are not independent of each other and satisfy

$$\begin{aligned}
 \varepsilon_{1,22} + \varepsilon_{2,11} - 2\varepsilon_{6,12} &= 0, \\
 \varepsilon_{5,2} - \varepsilon_{4,1} &= 0, \\
 E_{1,2} - E_{2,1} &= 0.
 \end{aligned} \tag{2.3.39}$$

Substitution of Eq. (2.3.35) into Eq. (2.3.39) and manipulations yields

$$\begin{aligned}
 L_A J + L_B G + L_C H &= 0, \\
 L_B J + L_D G + L_E H &= 0, \\
 L_C J + L_E G + L_F H &= 0,
 \end{aligned} \tag{2.3.40}$$

where

$$\begin{aligned}
 L_A &= s'_{22} \frac{\partial^4}{\partial x_1^4} - 2s'_{26} \frac{\partial^4}{\partial x_1^3 \partial x_2} + (2s'_{12} + s'_{66}) \frac{\partial^4}{\partial x_1^2 \partial x_2^2} - 2s'_{16} \frac{\partial^4}{\partial x_2^3 \partial x_1} + s'_{11} \frac{\partial^4}{\partial x_2^4}, \\
 L_B &= -g'_{22} \frac{\partial^3}{\partial x_1^3} + (g'_{12} + g'_{26}) \frac{\partial^3}{\partial x_1^2 \partial x_2} - (g'_{21} + g'_{16}) \frac{\partial^3}{\partial x_1 \partial x_2^2} + g'_{11} \frac{\partial^3}{\partial x_2^3}, \\
 L_C &= -s'_{24} \frac{\partial^3}{\partial x_1^3} + (s'_{25} + s'_{46}) \frac{\partial^3}{\partial x_1^2 \partial x_2} - (s'_{14} + s'_{56}) \frac{\partial^3}{\partial x_1 \partial x_2^2} + s'_{15} \frac{\partial^3}{\partial x_2^3}, \\
 L_D &= -\beta'_{22} \frac{\partial^2}{\partial x_1^2} + 2\beta'_{12} \frac{\partial^2}{\partial x_1 \partial x_2} - \beta'_{11} \frac{\partial^2}{\partial x_2^2}, \\
 L_E &= g'_{24} \frac{\partial^2}{\partial x_1^2} - (g'_{14} + g'_{25}) \frac{\partial^2}{\partial x_1 \partial x_2} + g'_{15} \frac{\partial^2}{\partial x_2^2}, \\
 L_F &= s'_{44} \frac{\partial^2}{\partial x_1^2} - 2s'_{45} \frac{\partial^2}{\partial x_1 \partial x_2} + s'_{55} \frac{\partial^2}{\partial x_2^2}.
 \end{aligned} \tag{2.3.41}$$

From Eq. (2.3.40) we have

$$\begin{vmatrix} L_A & L_B & L_C \\ L_B & L_D & L_E \\ L_C & L_E & L_F \end{vmatrix} J = 0. \tag{2.3.42}$$

By assuming that function J takes the same form as in the Stroh formalism, which means $J = J(Z)$, $Z = x_1 + \mu x_2$, Eq. (2.3.42) is transformed to an algebraic equation on μ , given by

$$\begin{vmatrix} l_a(\mu) & l_b(\mu) & l_c(\mu) \\ l_b(\mu) & l_d(\mu) & l_e(\mu) \\ l_c(\mu) & l_e(\mu) & l_f(\mu) \end{vmatrix} = 0, \tag{2.3.43}$$

where

$$\begin{aligned}
 l_a(\mu) &= s'_{11}\mu^4 - 2s'_{16}\mu^3 + (2s'_{12} + s'_{66})\mu^2 - 2s'_{26}\mu + s'_{22}, \\
 l_b(\mu) &= g'_{11}\mu^3 - (g'_{21} + g'_{16})\mu^2 - (g'_{12} + g'_{26})\mu - g'_{22}, \\
 l_c(\mu) &= s'_{15}\mu^3 - (s'_{14} + s'_{56})\mu^2 + (s'_{25} + s'_{46})\mu - s'_{24}, \\
 l_d(\mu) &= -\beta'_{11}\mu^2 + 2\beta'_{12}\mu - \beta'_{22}, \\
 l_e(\mu) &= g'_{15}\mu^2 - (g'_{14} + g'_{25})\mu + g'_{24}, \\
 l_f(\mu) &= s'_{55}\mu^2 - 2s'_{45}\mu + s'_{44}.
 \end{aligned} \tag{2.3.44}$$

Eq. (2.3.43) permits four pairs of conjugated roots of μ , just as p of the Stroh formalism. Thus the solution J takes the form

$$J = 2 \operatorname{Re} \sum_{\alpha=1}^4 J_{\alpha}(Z_{\alpha}), \quad (2.3.45)$$

where

$$Z_{\alpha} = x_1 + \mu_{\alpha} x_2, \quad \operatorname{Im}(\mu_{\alpha}) > 0 \quad (\alpha = 1, 2, 3, 4). \quad (2.3.46)$$

Thus, H and G are obtained from Eq. (2.3.40) as

$$\begin{aligned} G &= 2 \operatorname{Re} \sum_{\alpha=1}^4 \lambda_{\alpha} \frac{dJ_{\alpha}(Z)}{dZ} \bigg|_{Z=Z_{\alpha}}, \\ H &= 2 \operatorname{Re} \sum_{\alpha=1}^4 \delta_{\alpha} \frac{dJ_{\alpha}(Z)}{dZ} \bigg|_{Z=Z_{\alpha}}, \end{aligned} \quad (2.3.47)$$

where

$$\begin{aligned} \lambda_{\alpha} &= -\frac{l_b(\mu_{\alpha})l_f(\mu_{\alpha}) - l_c(\mu_{\alpha})l_e(\mu_{\alpha})}{l_d(\mu_{\alpha})l_f(\mu_{\alpha}) - l_e(\mu_{\alpha})l_e(\mu_{\alpha})}, \\ \delta_{\alpha} &= -\frac{l_c(\mu_{\alpha})l_d(\mu_{\alpha}) - l_b(\mu_{\alpha})l_e(\mu_{\alpha})}{l_d(\mu_{\alpha})l_f(\mu_{\alpha}) - l_e(\mu_{\alpha})l_e(\mu_{\alpha})}. \end{aligned} \quad (2.3.48)$$

The stresses are then obtained from Eq. (2.3.31) as

$$\begin{aligned} \sigma_{11} &= 2 \operatorname{Re} \sum_{\alpha=1}^4 \mu_{\alpha}^2 f'_{\alpha}(Z_{\alpha}), \\ \sigma_{22} &= 2 \operatorname{Re} \sum_{\alpha=1}^4 f'_{\alpha}(Z_{\alpha}), \\ \sigma_{12} &= -2 \operatorname{Re} \sum_{\alpha=1}^4 \mu_{\alpha} f'_{\alpha}(Z_{\alpha}), \\ \sigma_{13} &= 2 \operatorname{Re} \sum_{\alpha=1}^4 \delta_{\alpha} \mu_{\alpha} f'_{\alpha}(Z_{\alpha}), \\ \sigma_{23} &= -2 \operatorname{Re} \sum_{\alpha=1}^4 \delta_{\alpha} f'_{\alpha}(Z_{\alpha}), \\ D_1 &= 2 \operatorname{Re} \sum_{\alpha=1}^4 \lambda_{\alpha} \mu_{\alpha} f'_{\alpha}(Z_{\alpha}), \\ D_2 &= -2 \operatorname{Re} \sum_{\alpha=1}^4 \lambda_{\alpha} f'_{\alpha}(Z_{\alpha}), \end{aligned} \quad (2.3.49)$$

where

$$f_\alpha(Z_\alpha) = \frac{dJ_\alpha(Z)}{dZ}. \quad (2.3.50)$$

The displacements and the electric potential are solved as

$$\begin{aligned} u_1 &= 2 \operatorname{Re} \sum_{\alpha=1}^4 p_\alpha f_\alpha(Z_\alpha), \\ u_2 &= 2 \operatorname{Re} \sum_{\alpha=1}^4 q_\alpha f_\alpha(Z_\alpha), \\ u_3 &= 2 \operatorname{Re} \sum_{\alpha=1}^4 r_\alpha f_\alpha(Z_\alpha), \\ \varphi &= 2 \operatorname{Re} \sum_{\alpha=1}^4 s_\alpha f_\alpha(Z_\alpha), \end{aligned} \quad (2.3.51)$$

where

$$\begin{aligned} p_\alpha &= s'_{11}\mu_\alpha^2 + s'_{12} - s'_{16}\mu_\alpha + \delta_\alpha(s'_{15}\mu_\alpha - s'_{14}) + \lambda_\alpha(g'_{11}\mu_\alpha - g'_{21}), \\ q_\alpha &= s'_{12}\mu_\alpha + s'_{22} / \mu_\alpha - s'_{26} + \delta_\alpha(s'_{25} - s'_{24} / \mu_\alpha) + \lambda_\alpha(g'_{12} - g'_{21} / \mu_\alpha), \\ r_\alpha &= s'_{15}\mu_\alpha^2 + s'_{25} - s'_{56}\mu_\alpha + \delta_\alpha(s'_{55}\mu_\alpha - s'_{45}) + \lambda_\alpha(g'_{15}\mu_\alpha - g'_{25}), \\ s_\alpha &= g'_{11}\mu_\alpha^2 + g'_{12} - g'_{16}\mu_\alpha + \delta_\alpha(g'_{15}\mu_\alpha - g'_{14}) + \lambda_\alpha(\beta'_{12} - \beta'_{22}\mu_\alpha). \end{aligned} \quad (2.3.52)$$

By seeking a matrix form of Eq. (2.3.51), we have

$$\mathbf{u} = 2 \operatorname{Re}[\mathbf{A}\mathbf{f}(Z)], \quad (2.3.53)$$

where

$$\mathbf{A} = \begin{bmatrix} p_1 & p_2 & p_3 & p_4 \\ q_1 & q_2 & q_3 & q_4 \\ r_1 & r_2 & r_3 & r_4 \\ s_1 & s_2 & s_3 & s_4 \end{bmatrix}, \quad \mathbf{f}(Z) = \begin{bmatrix} f_1(Z_1) \\ f_2(Z_2) \\ f_3(Z_3) \\ f_4(Z_4) \end{bmatrix}, \quad \mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ \varphi \end{bmatrix}. \quad (2.3.54)$$

Introducing the stress potential function as in the Stroh formalism,

$$\boldsymbol{\phi} = 2 \operatorname{Re}[\mathbf{B}\mathbf{f}(Z)], \quad (2.3.55)$$

where

$$\mathbf{B} = \begin{bmatrix} -\mu_1 & -\mu_2 & -\mu_3 & -\mu_4 \\ 1 & 1 & 1 & 1 \\ -\delta_1 & -\delta_2 & -\delta_3 & -\delta_4 \\ -\lambda_1 & -\lambda_2 & -\lambda_3 & -\lambda_4 \end{bmatrix}. \quad (2.3.56)$$

By comparing Eqs. (2.3.53) and (2.3.55) to Eq. (2.3.24), the results obtained by using the Lekhnitskii formalism take the same form of the Stroh formalism.

2.3.3 Conformal Transformation of the Core Function

From Eq. (2.3.1) to Eq. (2.3.30), the Stroh formalism for piezoelectric materials is constructed without specifying the microstructure and boundary conditions for the materials. Eq. (2.3.24) was derived from the request for satisfaction of the static equilibrium equation for the two-dimensional problem. In Eq. (2.3.24), $\langle f(Z_\alpha) \rangle$ and $\mathbf{g}, \mathbf{h}$ are not specified. Similarly, in the Lekhnitskii formalism, the exact form of $J_\alpha(Z_\alpha)$ or $f_\alpha(Z_\alpha)$ is also not specified. While the determination of $\mathbf{g}, \mathbf{h}$ relies on the solution of the algebraic equations obtained from the boundary conditions at the interface or surface, an appropriate form of $\langle f(Z_\alpha) \rangle$ is vital, corresponding to specific configurations of the boundary, so that it is possible to obtain the algebraic equation. Here, we call $f(Z_\alpha)$ the core function for the Stroh formalism (or the Lekhnitskii formalism).

The analysis and selection of $f(Z_\alpha)$ can be treated by conformal transformation or conformal mapping, a very powerful method in functions of complex variables. By considering the notation $Z = x_1 + px_2$ in Eq. (2.3.24), for $p = p_{\text{re}} + ip_{\text{im}}$, $i = \sqrt{-1}$, the complex plane $f(Z) = Z$ is described by the variable set $(x_1 + p_{\text{re}}x_2, p_{\text{im}}x_2)$, which distinguishes itself from the real physical plane described by (x_1, x_2) . For a complicated boundary configuration described in the real physical plane, if we can find the right core function $f(Z_\alpha)$, it may reduce to a very simple configuration in the complex plane. And after solution in the complex plane, we can take an inverse transformation and find the corresponding solution in the real physical plane. If a unique inverse transformation exists, the transformation between the real physical plane and the complex plane is said to be conformal. Since this is not a book on mathematics, the details for all kinds of typical conformal transformations are not discussed here. Anyone who is interested in the subject is recommended to read the references to Courant and Hilbert (1953) and Liang et al. (1998).

2.4 **Basic Solutions for Two-Dimensional Problems**

Based on the Stroh formalism and Lekhnitskii formalism introduced in Section 2.3, some basic two-dimensional problems for piezoelectric materials are solved here. In Section 2.4.1, we consider an infinite piezoelectric matrix containing an elliptical cylindrical inclusion. Du et al. (1994), Liang et al. (1995), Chung (1995) and Chung and Ting (1996) are among the first researchers who extended the Stroh formalism to treat this problem. Qin et al. (1999) dealt with an infinite piezoelectric matrix containing holes and the effect of thermal load was included. Here the formulation of a general solution is based on Chung and Ting's work (Chung and Ting, 1996). In Section 2.4.2, the solution for a Griffith crack embedded in an infinite piezoelectric matrix is considered. The solution for the electroelastic fields is easily deduced from the results in Section 2.4.1, since the shape of a Griffith crack is considered as a limiting case with one of the half-axis approaching zero. However, the solution obtained by using the classical continuum mechanics gives unreasonable results for points very close to the crack tip. For this reason, the J -integral and the nonlocal theory for treating crack tip problems are briefly introduced. One should notice that crack problems are the most important issue for fracture mechanics and have received great attention from many researchers (Pak, 1990b; 1992; Wang, 1992a; Suo et al., 1992; McMeeking, 1989; Sosa, 1992; Eringen, 1978; 1979; Zhou et al., 1999; 2002a; 2002b; 2004; Zhou and Wang, 2003; 2006; Park and Sun, 1995; Xu and Rajapakse, 1999, 2000; Qin and Yu, 1997; Gao et al., 1997; Kuna, 2010; Fu and Zhang, 2000; Wang and Singh, 1997; Heyer et al., 1998; Fang and Liu, 2009). The discussion presented here covers only a part of these contents. In Section 2.4.3, the influence of line dislocations and line charges on the infinite piezoelectric matrix are treated. The solution presented is based on Barnett and Lothe's work (Barnett and Lothe, 1975). Similar problems have been discussed by Pak (1990b; 1992) and Chung and Ting (1995). The physical meaning of this solution can also be interpreted as the two-dimensional Green's function. Similar to the crack problems, due to its great physical significance Green's functions (including 2-D and 3-D) have been widely studied (Chen, 1993; 2000; Chen and Lin, 1993; Chen et al., 2005; Ding et al., 1996; 1997a; 1997b; 1997c; Dunn, 1994; Dunn and Wienecke, 1996; 1999; Mulier et al., 1998; Hou and Leung, 2009; Hou et al., 2007; Pan and Yuan, 2000; Qin, 1998; 1999a; 1999b).

2.4.1 *Elliptical Cylindrical Inclusions in Piezoelectric Materials*

2.4.1.1 **General Solution**

Consider an elliptical cylindrical inclusion embedded in an infinite piezoelectric matrix, with infinite length in the x_3 direction. The contour of the cross section of

the inclusion is defined by

$$\Gamma : \begin{cases} x_1 = a \cos \psi \\ x_2 = b \sin \psi \end{cases} \quad (2.4.1)$$

and plotted in Fig. 2.1. $\mathbf{n}$ and $\mathbf{m}$ are defined as unit vectors tangential and normal to Γ . If the angle between $\mathbf{n}$ and axis x_1 is denoted by ω , we have

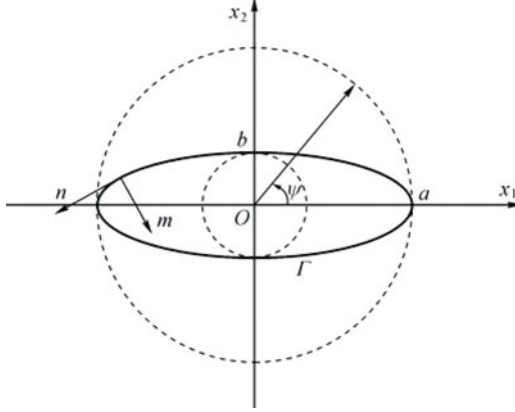


Fig. 2.1 Configuration of Γ in plane (x_1, x_2)

$$\mathbf{n} = [\cos \omega, \sin \omega, 0]^T, \quad \mathbf{m} = [-\sin \omega, \cos \omega, 0]^T. \quad (2.4.2)$$

The transformation equations between ω and ψ are given by

$$\cos \omega = \frac{-a}{\rho(\psi)} \sin \psi, \quad \sin \omega = \frac{b}{\rho(\psi)} \cos \psi, \quad (2.4.3)$$

where

$$\rho(\psi) = \sqrt{a^2 \sin^2 \psi + b^2 \cos^2 \psi}. \quad (2.4.4)$$

The solution for the elastic and electric fields of this problem takes the form of Eq. (2.3.24), where the core function $\langle f(Z_\alpha) \rangle$ is not specified. According to Eq. (2.3.24), the elastic and electric fields at the boundary Γ should take the form

$$u_k|_{\Gamma} = u_k(\cos \psi, \sin \psi). \quad (2.4.5)$$

It is natural to think that if $\langle f(Z_\alpha) \rangle$ is carefully chosen so that it can be expressed by $\cos \psi$ and $\sin \psi$ on Γ , the solution of the boundary value problem

will be very easy. Now consider Joukowski's transformation

$$Z_\alpha = c_\alpha \zeta_\alpha + d_\alpha \zeta_\alpha^{-1} \quad (\alpha = 1, 2, 3, 4), \quad (2.4.6)$$

where c_α and d_α are complex constants to be determined. In the complex Z_α -plane, the real and imaginary parts are expressed by

$$\operatorname{Re}(Z_\alpha) = a \cos \psi + \operatorname{Re}(p_\alpha) b \sin \psi, \quad \operatorname{Im}(Z_\alpha) = \operatorname{Im}(p_\alpha) b \sin \psi. \quad (2.4.7)$$

c_α and d_α are so chosen that in the ζ_α -plane, we have

$$\operatorname{Re}(\zeta_\alpha) = \cos \psi, \quad \operatorname{Im}(\zeta_\alpha) = \sin \psi. \quad (2.4.8)$$

Substitution of Eqs. (2.4.7) and (2.4.8) into Eq. (2.4.6) yields

$$c_\alpha = \frac{a - ip_\alpha b}{2}, \quad d_\alpha = \frac{a + ip_\alpha b}{2}, \quad (2.4.9)$$

thus we have

$$\zeta_\alpha = \frac{Z_\alpha + \sqrt{Z_\alpha^2 - (a^2 + p_\alpha^2 b^2)}}{a - ip_\alpha b}. \quad (2.4.10)$$

For a disturbance caused by the existence of a heterostructure, when $x_1 \rightarrow \infty$ or $x_2 \rightarrow \infty$, the generalized stress and strain fields should vanish. Thus, Eq. (2.4.10) is not suitable for application as the core function for such a disturbance. Since the generalized strain fields are first-order derivatives of the generalized displacement fields, two types of core functions are chosen for problems with an elliptic boundary (Chung and Ting, 1996). The first type is

$$f(Z_\alpha) = \mathbf{In}(\zeta_\alpha), \quad (2.4.11)$$

and the second type is

$$f(Z_\alpha) = \sum_{k=1}^{\infty} \zeta_\alpha^{-k}. \quad (2.4.12)$$

The generalized displacement vector and the generalized stress potential vector expressed with the first type of core functions are expressed as

$$\begin{aligned} \mathbf{u}^I &= 2 \operatorname{Re} \left\{ \mathbf{A} < \mathbf{In}(\zeta_\alpha) > \mathbf{A}^T \right\} \mathbf{g}_0 + 2 \operatorname{Re} \left\{ \mathbf{A} < \mathbf{In}(\zeta_\alpha) > \mathbf{B}^T \right\} \mathbf{h}_0, \\ \phi^I &= 2 \operatorname{Re} \left\{ \mathbf{B} < \mathbf{In}(\zeta_\alpha) > \mathbf{A}^T \right\} \mathbf{g}_0 + 2 \operatorname{Re} \left\{ \mathbf{B} < \mathbf{In}(\zeta_\alpha) > \mathbf{B}^T \right\} \mathbf{h}_0. \end{aligned} \quad (2.4.13)$$

The generalized displacement vector and the generalized stress potential vector expressed with the second type of core functions are expressed as

$$\begin{aligned} \mathbf{u}^{\text{II}} &= 2 \sum_{k=1}^{\infty} \text{Re} \left\{ \mathbf{A} < \zeta_{\alpha}^{-k} > \mathbf{A}^{\text{T}} \right\} \mathbf{g}_k + 2 \sum_{k=1}^{\infty} \text{Re} \left\{ \mathbf{A} < \zeta_{\alpha}^{-k} > \mathbf{B}^{\text{T}} \right\} \mathbf{h}_k, \\ \phi^{\text{II}} &= 2 \sum_{k=1}^{\infty} \text{Re} \left\{ \mathbf{B} < \zeta_{\alpha}^{-k} > \mathbf{A}^{\text{T}} \right\} \mathbf{g}_k + 2 \sum_{k=1}^{\infty} \text{Re} \left\{ \mathbf{B} < \zeta_{\alpha}^{-k} > \mathbf{B}^{\text{T}} \right\} \mathbf{h}_k. \end{aligned} \quad (2.4.14)$$

Since $\mathbf{In}(\zeta_{\alpha})$ is a multivalued function, we mark a cut along the line $\psi = 0$ and permit a discontinuity along the positive x_1 axis. From Eq. (2.4.8) it is easy to find that on the boundary Γ

$$< \mathbf{In}(\zeta_{\alpha}) > \big|_{\Gamma} = i\psi \mathbf{I}, \quad < \zeta_{\alpha}^{-k} > = [\cos(k\psi) - i \sin(k\psi)] \mathbf{I}. \quad (2.4.15)$$

Eq. (2.4.15) describes the basic characteristics of the two types of core functions on the boundary. Substituting Eq. (2.4.15) into Eq. (2.4.13) and Eq. (2.4.14), we have

$$\mathbf{u}^{\text{I}} \big|_{\Gamma} = (\mathbf{H}\mathbf{g}_0 + \mathbf{S}\mathbf{h}_0)\psi, \quad \phi^{\text{I}} \big|_{\Gamma} = (\mathbf{S}^{\text{T}}\mathbf{g}_0 - \mathbf{L}\mathbf{h}_0)\psi, \quad (2.4.16)$$

and

$$\begin{aligned} \mathbf{u}^{\text{II}} &= \sum_{k=1}^{\infty} [\cos(k\psi)\mathbf{h}_k - \sin(k\psi)(\mathbf{H}\mathbf{g}_k + \mathbf{S}\mathbf{h}_k)], \\ \phi^{\text{II}} &= \sum_{k=1}^{\infty} [\cos(k\psi)\mathbf{g}_k - \sin(k\psi)(\mathbf{S}^{\text{T}}\mathbf{g}_k - \mathbf{L}\mathbf{h}_k)]. \end{aligned} \quad (2.4.17)$$

If the solution is written in the form of Eq. (2.4.16) and Eq. (2.4.17), the generalized stress vector $\mathbf{t}_m$ and the generalized hoop stress vector $\mathbf{t}_n$ are then obtained at the elliptic cylindrical boundary. The analysis of $\mathbf{t}_m$ and $\mathbf{t}_n$ is helpful for the determination of the stress concentration at the boundary. The components of $\mathbf{t}_m$ and $\mathbf{t}_n$ are given by

$$\mathbf{t}_m = [\sigma_{m1}, \sigma_{m2}, \sigma_{m3}, D_m]^{\text{T}}, \quad \mathbf{t}_n = [\sigma_{n1}, \sigma_{n1}, \sigma_{n1}, -D_n]^{\text{T}}, \quad (2.4.18)$$

where σ_{m1} is the component of the stress defined on the surface perpendicular to the normal vector $\mathbf{m}$ in x_1 direction, and $D_m = \mathbf{D} \cdot \mathbf{m}$. As illustrated in Fig. 2.2 and Fig. 2.3, $\mathbf{t}_m$ and $\mathbf{t}_n$ are defined on different surfaces of the volume element at the boundary. Yet, since the electric displacement $\mathbf{D}$ is a physical quantity defined in the body but not on the surface, it takes the same value and direction on different surfaces of the volume element.

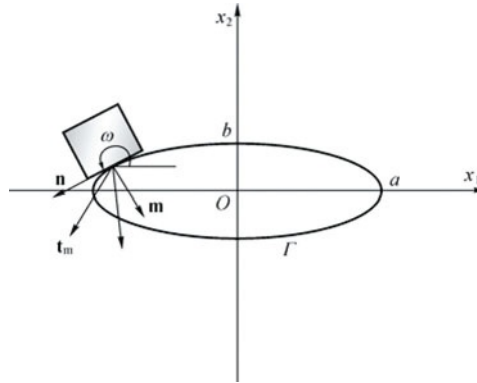


Fig. 2.2 Generalized stress vector $\mathbf{t}_m$ along Γ

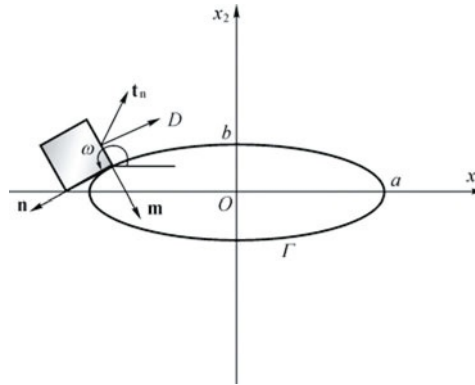


Fig. 2.3 Generalized hoop stress vector $\mathbf{t}_n$ along Γ

Using Eq. (2.3.9) and Eq. (2.4.18), after manipulation we have

$$\mathbf{t}_m = \phi_n = \frac{\partial \phi}{\partial \psi} \Big|_r, \quad \mathbf{t}_n = \phi_m = -\phi_{,1} \sin \omega + \phi_{,2} \cos \omega. \quad (2.4.19)$$

Alternatively, $\mathbf{t}_n$ can be expressed as (Chung, 1995; Ting and Yan, 1991)

$$\mathbf{t}_n = \mathbf{N}_3(\omega) \mathbf{u}_{,n} + \mathbf{N}_1^T(\omega) \mathbf{t}_m. \quad (2.4.20)$$

Substituting Eq. (2.4.16) and Eq. (2.4.17), we have

$$\begin{aligned} \mathbf{t}_m^I &= \frac{1}{\rho(\psi)} (\mathbf{S}^T \mathbf{g}_0 - \mathbf{L} \mathbf{h}_0), \\ \mathbf{t}_m^{II} &= \frac{-1}{\rho(\psi)} \sum_{k=1}^{\infty} k [\cos(k\psi) \mathbf{g}_k - \sin(k\psi) (\mathbf{S}^T \mathbf{g}_k - \mathbf{L} \mathbf{h}_k)]. \end{aligned} \quad (2.4.21)$$

and

$$\begin{aligned} \mathbf{t}_n^I &= \frac{1}{\rho(\psi)} [\mathbf{N}_3(\omega)(\mathbf{H}\mathbf{g}_0 + \mathbf{S}\mathbf{h}_0) + \mathbf{N}_1^T(\omega)(\mathbf{S}^T\mathbf{g}_0 - \mathbf{L}\mathbf{h}_0)], \\ \mathbf{t}_n^{II} &= -\frac{\mathbf{N}_3(\omega)}{\rho(\psi)} \sum_{k=1}^{\infty} k[\sin(k\psi)\mathbf{h}_k + \cos(k\psi)(\mathbf{H}\mathbf{g}_k + \mathbf{S}\mathbf{h}_k)] \\ &\quad - \frac{\mathbf{N}_1^T(\omega)}{\rho(\psi)} \sum_{k=1}^{\infty} k[\sin(k\psi)\mathbf{g}_k + \cos(k\psi)(\mathbf{S}^T\mathbf{g}_k - \mathbf{L}\mathbf{h}_k)]. \end{aligned} \quad (2.4.22)$$

2.4.1.2 Inhomogeneities

Assume that the infinite matrix is subjected to uniform generalized strain fields ε_{KI}^∞ (or stress fields σ_{KI}^∞) at infinity, and an elliptic piezoelectric inhomogeneity with boundary Γ is embedded in the matrix. Let the material constants for the matrix and the inhomogeneity be denoted, respectively, by C_{iJKI} and C_{iJKI}^* . The generalized displacement vector and the generalized stress potential vector for the matrix are given by

$$\begin{aligned} \mathbf{u} &= \mathbf{u}^\infty + 2 \operatorname{Re} \left\{ \mathbf{A} \langle f(Z_\alpha) \rangle \mathbf{A}^T \right\} \mathbf{g} + 2 \operatorname{Re} \left\{ \mathbf{A} \langle f(Z_\alpha) \rangle \mathbf{B}^T \right\} \mathbf{h}, \\ \boldsymbol{\phi} &= \boldsymbol{\phi}^\infty + 2 \operatorname{Re} \left\{ \mathbf{B} \langle f(Z_\alpha) \rangle \mathbf{A}^T \right\} \mathbf{g} + 2 \operatorname{Re} \left\{ \mathbf{B} \langle f(Z_\alpha) \rangle \mathbf{B}^T \right\} \mathbf{h}, \end{aligned} \quad (2.4.23)$$

where $\mathbf{u}^\infty$ and $\boldsymbol{\phi}^\infty$ are given by

$$\mathbf{u}^\infty = x_1 \boldsymbol{\varepsilon}_1^\infty + x_2 \boldsymbol{\varepsilon}_2^\infty, \quad \boldsymbol{\phi}^\infty = x_1 \mathbf{t}_2^\infty - x_2 \mathbf{t}_1^\infty, \quad (2.4.24)$$

and

$$\boldsymbol{\varepsilon}_1^\infty = [\varepsilon_{11}^\infty, 0, 2\varepsilon_{13}^\infty, -E_1^\infty]^T, \quad \boldsymbol{\varepsilon}_2^\infty = [2\varepsilon_{12}^\infty, \varepsilon_{22}^\infty, 2\varepsilon_{23}^\infty, -E_2^\infty]^T, \quad (2.4.25)$$

$$\mathbf{t}_1^\infty = [\sigma_{11}^\infty, \sigma_{12}^\infty, \sigma_{13}^\infty, D_1^\infty]^T, \quad \mathbf{t}_2^\infty = [\sigma_{21}^\infty, \sigma_{22}^\infty, \sigma_{23}^\infty, D_2^\infty]^T. \quad (2.4.26)$$

It was proved by Wang (1992a) that, for an ellipsoidal shaped inhomogeneity, when the applied field at infinity is uniform, the generalized elastic stress and strain inside the inhomogeneity are also uniform. Thus the generalized displacement vector and the generalized stress potential vector for the inhomogeneity can be written as

$$\mathbf{u}^{in} = x_1 \boldsymbol{\varepsilon}_1^{in} + x_2 \boldsymbol{\varepsilon}_2^{in}, \quad \boldsymbol{\phi}^{in} = x_1 \mathbf{t}_2^{in} - x_2 \mathbf{t}_1^{in}, \quad (2.4.27)$$

where

$$\boldsymbol{\varepsilon}_1^{in} = [\boldsymbol{\varepsilon}_{11}^{in}, \boldsymbol{\Theta}, 2\boldsymbol{\varepsilon}_{13}^{in}, -\boldsymbol{E}_1^{in}]^T, \quad \boldsymbol{\varepsilon}_2^{in} = [2\boldsymbol{\varepsilon}_{12}^{in} - \boldsymbol{\Theta}, \boldsymbol{\varepsilon}_{22}^{in}, 2\boldsymbol{\varepsilon}_{23}^{in}, -\boldsymbol{E}_2^{in}]^T, \quad (2.4.28)$$

$$\mathbf{t}_1^{in} = [\boldsymbol{\sigma}_{11}^{in}, \boldsymbol{\sigma}_{12}^{in}, \boldsymbol{\sigma}_{13}^{in}, \boldsymbol{D}_1^{in}]^T, \quad \mathbf{t}_2^{in} = [\boldsymbol{\sigma}_{21}^{in}, \boldsymbol{\sigma}_{22}^{in}, \boldsymbol{\sigma}_{23}^{in}, \boldsymbol{D}_2^{in}]^T. \quad (2.4.29)$$

$\boldsymbol{\Theta}$ in Eq. (2.4.28) is the rigid body rotation of the inhomogeneity.

Substitution of Eq. (2.4.1) into Eq. (2.4.27) yields

$$\mathbf{u}^{in} = a\boldsymbol{\varepsilon}_1^{in} \cos \psi + b\boldsymbol{\varepsilon}_2^{in} \sin \psi, \quad \boldsymbol{\phi}^{in} = a\mathbf{t}_2^{in} \cos \psi - b\mathbf{t}_1^{in} \sin \psi. \quad (2.4.30)$$

By comparing Eq. (2.4.30) to Eq. (2.4.16) and Eq. (2.4.17), it is obvious that

$$f(Z_\alpha) = \zeta_\alpha^{-1}. \quad (2.4.31)$$

$\mathbf{u}$ and $\boldsymbol{\phi}$ on the boundary Γ are then expressed as

$$\begin{aligned} \mathbf{u}|_\Gamma &= \cos \psi (\mathbf{h}_1 + a\boldsymbol{\varepsilon}_1^\infty) - \sin \psi (\mathbf{H}\mathbf{g}_1 + \mathbf{S}\mathbf{h}_1 - b\boldsymbol{\varepsilon}_2^\infty), \\ \boldsymbol{\phi}|_\Gamma &= \cos \psi (\mathbf{g}_1 + a\mathbf{t}_2^\infty) - \sin \psi (\mathbf{S}^T \mathbf{g}_1 - \mathbf{L}\mathbf{h}_1 + b\mathbf{t}_1^\infty). \end{aligned} \quad (2.4.32)$$

The boundary condition on Γ can be expressed as

$$\mathbf{u}^{in}|_\Gamma = \mathbf{u}|_\Gamma, \quad \boldsymbol{\phi}^{in}|_\Gamma = \boldsymbol{\phi}|_\Gamma. \quad (2.4.33)$$

The boundary condition of stresses and electric displacements are expressed by the generalized stress potential $\boldsymbol{\phi}$ in Eq. (2.4.33). Substituting Eq. (2.4.30) into Eq. (2.4.32) we have

$$\begin{bmatrix} \mathbf{h}_1 \\ \mathbf{g}_1 \end{bmatrix} = a \begin{bmatrix} \boldsymbol{\varepsilon}_1^{in} \\ \mathbf{t}_2^{in} \end{bmatrix} - a \begin{bmatrix} \boldsymbol{\varepsilon}_1^\infty \\ \mathbf{t}_2^\infty \end{bmatrix}, \quad (2.4.34)$$

$$\begin{bmatrix} \mathbf{S} & \mathbf{H} \\ -\mathbf{L} & \mathbf{S}^T \end{bmatrix} \begin{bmatrix} \mathbf{h}_1 \\ \mathbf{g}_1 \end{bmatrix} = b \begin{bmatrix} \boldsymbol{\varepsilon}_2^\infty \\ -\mathbf{t}_1^\infty \end{bmatrix} - b \begin{bmatrix} \boldsymbol{\varepsilon}_2^{in} \\ \mathbf{t}_1^{in} \end{bmatrix}. \quad (2.4.35)$$

Eqs. (2.4.34) and (2.4.35) are not sufficient for solving the six unknown vectors $\mathbf{h}_1, \mathbf{g}_1, \boldsymbol{\varepsilon}_1^{in}, \boldsymbol{\varepsilon}_2^{in}, \mathbf{t}_1^{in}, \mathbf{t}_2^{in}$. Additional relations are obtained from Eq. (2.3.30) as

$$\mathbf{N}^* \begin{bmatrix} \boldsymbol{\varepsilon}_1^{in} \\ \mathbf{t}_2^{in} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\varepsilon}_2^{in} \\ -\mathbf{t}_1^{in} \end{bmatrix}, \quad \mathbf{N} \begin{bmatrix} \boldsymbol{\varepsilon}_1^\infty \\ \mathbf{t}_2^\infty \end{bmatrix} = \begin{bmatrix} \boldsymbol{\varepsilon}_2^\infty \\ -\mathbf{t}_1^\infty \end{bmatrix}, \quad (2.4.36)$$

where $\mathbf{N}^*$ is obtained from Eq. (2.3.18) with material constants of the inhomogeneity C_{ijkl}^* . From the above equations, the solution for the elastic and electric fields inside the inhomogeneity is given by

$$\begin{bmatrix} \boldsymbol{\varepsilon}_1^{in} \\ \mathbf{t}_2^{in} \end{bmatrix} = \mathbf{D}^{-1} \left\{ \frac{b}{a} \mathbf{N} + \begin{bmatrix} \mathbf{S} & \mathbf{H} \\ -\mathbf{L} & \mathbf{S}^T \end{bmatrix} \right\} \begin{bmatrix} \boldsymbol{\varepsilon}_1^\infty \\ \mathbf{t}_2^\infty \end{bmatrix}, \quad (2.4.37)$$

where

$$\mathbf{D} = \frac{b}{a} \mathbf{N}^* + \begin{bmatrix} \mathbf{S} & \mathbf{H} \\ -\mathbf{L} & \mathbf{S}^T \end{bmatrix}. \quad (2.4.38)$$

The solution for $\mathbf{h}_1, \mathbf{g}_1$ is given by

$$\begin{bmatrix} \mathbf{h}_1 \\ \mathbf{g}_1 \end{bmatrix} = \mathbf{D}^{-1} (\mathbf{N} - \mathbf{N}^*) \begin{bmatrix} \boldsymbol{\varepsilon}_1^\infty \\ \mathbf{t}_2^\infty \end{bmatrix}. \quad (2.4.39)$$

The rigid body rotation is easily obtained from Eq. (2.4.34) as

$$\Theta = \frac{(\mathbf{h}_1)_2}{a}, \quad (2.4.40)$$

where $(\mathbf{h}_1)_2$ denotes the second component of the vector $\mathbf{h}_1$. The generalized stress vector $\mathbf{t}_m$ and the generalized hoop stress vector $\mathbf{t}_n$ are deduced from Eqs. (2.4.19)–(2.4.22) as

$$\mathbf{t}_m = \frac{-1}{\rho(\psi)} [a \sin \psi \mathbf{t}_2^\infty + b \cos \psi \mathbf{t}_1^\infty] + \frac{-1}{\rho(\psi)} [\mathbf{g}_1 \sin \psi + (\mathbf{S}^T \mathbf{g}_1 - \mathbf{L} \mathbf{h}_1) \cos \psi], \quad (2.4.41)$$

$$\begin{aligned} \mathbf{t}_n = & \frac{-1}{\rho(\psi)} [b \cos \psi \mathbf{t}_2^\infty - a \sin \psi \mathbf{t}_1^\infty] - \frac{\mathbf{N}_3(\omega)}{\rho(\psi)} [\mathbf{h}_1 \sin \psi + (\mathbf{H} \mathbf{g}_1 + \mathbf{S} \mathbf{h}_1) \cos \psi] \\ & - \frac{\mathbf{N}_1^T(\omega)}{\rho(\psi)} [\mathbf{g}_1 \sin \psi + (\mathbf{S}^T \mathbf{g}_1 - \mathbf{L} \mathbf{h}_1) \cos \psi]. \end{aligned} \quad (2.4.42)$$

2.4.1.3 Holes

If we replace the inhomogeneity with a hole occupied by vacuum, when the

infinite matrix is subjected to uniform generalized strain fields ε_{kl}^∞ (or stress fields σ_{kl}^∞) at infinity, the interior fields are not considered, and the boundary conditions become

$$\phi|_r = 0. \quad (2.4.43)$$

In Eq. (2.4.43), the fourth equation $\phi_4|_r = 0$ is equivalent to Eq. (2.2.10). Since the hole is also a kind of inhomogeneity, the expression for the core function $f(Z_\alpha)$ is still given by Eq. (2.4.31). Thus $\mathbf{u}$ and ϕ on the boundary Γ are still given by Eq. (2.4.32). For Eq. (2.4.43) we have

$$\begin{aligned} \mathbf{g}_1 &= -a\mathbf{t}_2^\infty, \\ \mathbf{h}_1 &= \mathbf{L}^{-1}(\mathbf{S}^T \mathbf{g}_1 + b\mathbf{t}_1^\infty). \end{aligned} \quad (2.4.44)$$

The generalized stress vector $\mathbf{t}_m$ and the generalized hoop stress vector $\mathbf{t}_n$ can be calculated from Eqs. (2.4.41) and (2.4.42) using Eq. (2.4.44).

If the material of the matrix reduces to transversely isotropic (point group $6mm$), Sosa (1991) gave the solution of the electroelastic fields by using the Lekhnitskii formalism. The core function of the Lekhnitskii formalism $f_k(Z_k)$ takes the form

$$f_k(Z_k) = (B_k + iB_k^*)Z_k + (A_{k1}I_1 + A_{k2}I_2 + A_{k3}I_3) \frac{Z_k - \sqrt{Z_k^2 - (a^2 + \mu_k^2 b^2)}}{a + i\mu_k b} \quad (\text{not summed on } k). \quad (2.4.45)$$

where B_k and B_k^* ($k=1,2,3$) are real numbers determined by

$$\begin{aligned} \sigma_{11}^\infty &= 2 \operatorname{Re}\{\mu_k^2 (B_k + iB_k^*)\}, \\ \sigma_{22}^\infty &= 2(B_1 + B_2 + B_3), \\ \sigma_{12}^\infty &= -2 \operatorname{Re}\{\mu_k (B_k + iB_k^*)\}. \end{aligned} \quad (2.4.46)$$

μ_k ($k=1,2,3$) are solved from the following equation with positive imaginary parts

$$\begin{aligned} &a_{11}\delta_{11}\mu^6 + (a_{11}\delta_{22} + 2a_{12}\delta_{11} + a_{33}\delta_{11} + b_{21}^2 + b_{13}^2 + 2b_{21}b_{13})\mu^4 + \\ &(a_{22}\delta_{11} + 2a_{12}\delta_{22} + a_{33}\delta_{22} + 2b_{21}b_{22} + 2b_{13}b_{22})\mu^2 + (a_{22}\delta_{22} + b_{22}^2) = 0, \end{aligned} \quad (2.4.47)$$

where a_{ij} , b_{ij} and δ_{ij} are the material constants obtained from the constitutive equations

$$\begin{bmatrix} \varepsilon_{11} \\ \varepsilon_{22} \\ 2\varepsilon_{12} \\ -E_1 \\ -E_2 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & 0 & 0 & b_{21} \\ a_{12} & a_{22} & 0 & 0 & b_{22} \\ 0 & 0 & a_{33} & b_{13} & 0 \\ 0 & 0 & b_{13} & \delta_{11} & 0 \\ b_{21} & b_{22} & 0 & 0 & \delta_{22} \end{bmatrix} \begin{bmatrix} \sigma_{11} \\ \sigma_{22} \\ \sigma_{12} \\ D_1 \\ D_2 \end{bmatrix}. \quad (2.4.48)$$

Λ_{ij} are components of the matrix

$$\Lambda = \frac{1}{\Delta} \begin{bmatrix} \mu_2 \lambda_3 - \mu_3 \lambda_2 & \lambda_2 - \lambda_3 & \mu_3 - \mu_2 \\ \mu_3 \lambda_1 - \mu_1 \lambda_3 & \lambda_3 - \lambda_1 & \mu_1 - \mu_3 \\ \mu_1 \lambda_2 - \mu_2 \lambda_1 & \lambda_1 - \lambda_2 & \mu_2 - \mu_1 \end{bmatrix}, \quad (2.4.49)$$

$$\Delta = (\lambda_2 - \lambda_3)\mu_1 + (\lambda_3 - \lambda_1)\mu_2 + (\lambda_1 - \lambda_2)\mu_3. \quad (2.4.50)$$

λ_k are related to μ_k by

$$\lambda_k(\mu_k) = -\frac{(b_{21} + b_{13})\mu_k^2 + b_{22}}{\delta_{11}\mu_k^2 + \delta_{22}}. \quad (2.4.51)$$

The coefficients l_i ($i = 1, 2, 3$) are given by

$$l_1 = -\frac{a\sigma_{22}^\infty}{2} + i\frac{b\sigma_{12}^\infty}{2}, \quad l_2 = \frac{a\sigma_{12}^\infty}{2} - i\frac{b\sigma_{11}^\infty}{2}, \quad l_3 = \frac{aD_2^\infty}{2} - i\frac{bD_1^\infty}{2}. \quad (2.4.52)$$

The components of strains and electric fields are related to the potential function as

$$\begin{aligned} \varepsilon_{11} &= 2 \operatorname{Re} \sum_{k=1}^3 (a_{11}\mu_k^2 + a_{12} - b_{21}\lambda_k) f'_k, \\ \varepsilon_{22} &= 2 \operatorname{Re} \sum_{k=1}^3 (a_{12}\mu_k^2 + a_{22} - b_{22}\lambda_k) f'_k, \\ 2\varepsilon_{12} &= 2 \operatorname{Re} \sum_{k=1}^3 (-a_{33}\mu_k + b_{13}\lambda_k\mu_k) f'_k, \\ E_1 &= 2 \operatorname{Re} \sum_{k=1}^3 (b_{13} + \delta_{11}\lambda_k) \mu_k \phi'_k, \\ E_2 &= -2 \operatorname{Re} \sum_{k=1}^3 (b_{21}\mu_k^2 + b_{22} + \delta_{22}\lambda_k) f'_k. \end{aligned} \quad (2.4.53)$$

The components of stresses and electric displacements are related to the

potential function as

$$\begin{aligned}\sigma_{11} &= 2 \operatorname{Re} \sum_{k=1}^3 \mu_k^2 f'_k, \quad \sigma_{22} = 2 \operatorname{Re} \sum_{k=1}^3 f'_k, \quad \sigma_{12} = -2 \operatorname{Re} \sum_{k=1}^3 \mu_k f'_k, \\ D_1 &= 2 \operatorname{Re} \sum_{k=1}^3 \lambda_k \mu_k f'_k, \quad D_2 = -2 \operatorname{Re} \sum_{k=1}^3 \lambda_k f'_k.\end{aligned}\quad (2.4.54)$$

2.4.2 Cracks

2.4.2.1 A Griffith Crack in a Piezoelectric Matrix

The configuration of a crack is defined as a special hole where $b \rightarrow 0$ (without loss of generality, the crack is assumed to be located along the x axis). Thus the full-field solution for an infinite matrix containing a crack with generalized constant stresses or strains applied at infinity is easily obtained as

$$\begin{aligned}\mathbf{u} &= \mathbf{u}^\infty + 2 \operatorname{Re} \left\{ \mathbf{A} \langle f(Z_\alpha) \rangle \mathbf{A}^\top \right\} \mathbf{g}_1 + 2 \operatorname{Re} \left\{ \mathbf{A} \langle f(Z_\alpha) \rangle \mathbf{B}^\top \right\} \mathbf{h}_1, \\ \mathbf{j} &= \mathbf{j}^\infty + 2 \operatorname{Re} \left\{ \mathbf{B} \langle f(Z_\alpha) \rangle \mathbf{A}^\top \right\} \mathbf{g}_1 + 2 \operatorname{Re} \left\{ \mathbf{B} \langle f(Z_\alpha) \rangle \mathbf{B}^\top \right\} \mathbf{h}_1,\end{aligned}\quad (2.4.55)$$

where

$$\mathbf{g}_1 = -a \mathbf{t}_2^\infty, \quad \mathbf{h}_1 = \mathbf{L}^{-1} (\mathbf{S}^\top \mathbf{g}_1 + b \mathbf{t}_1^\infty), \quad (2.4.56)$$

and

$$f(Z_\alpha) = \frac{a}{Z_\alpha + \sqrt{Z_\alpha^2 - a^2}}. \quad (2.4.57)$$

For transversely isotropic materials, the solution was obtained by Sosa (1992) as

$$f_k = (B_k + i B_k^*) Z_k + \frac{1}{2} (-A_{k1} \sigma_{22}^\infty + A_{k2} \sigma_{12}^\infty + A_{k3} D_2^\infty) (Z_k - \sqrt{Z_k^2 - a^2}) \quad (\text{not summed on } k). \quad (2.4.58)$$

Eq. (2.4.58) holds for arbitrary loads applied at infinity. Thus we learn that for transversely isotropic materials, the disturbance fields are determined by three components only: σ_{22}^∞ , σ_{12}^∞ and D_2^∞ .

• Crack tip fields

Following Sosa's work (Sosa, 1992), the crack tip fields for transversely isotropic materials are solved here. For points near the right crack tip of a Griffith crack of

length $2l$ shown in Fig. 2.4, the polar coordinates can be established as

$$x_1 = l + r \cos \theta, \quad x_2 = r \sin \theta, \quad r > 0, \quad -\pi \leq \theta \leq \pi. \quad (2.4.59)$$

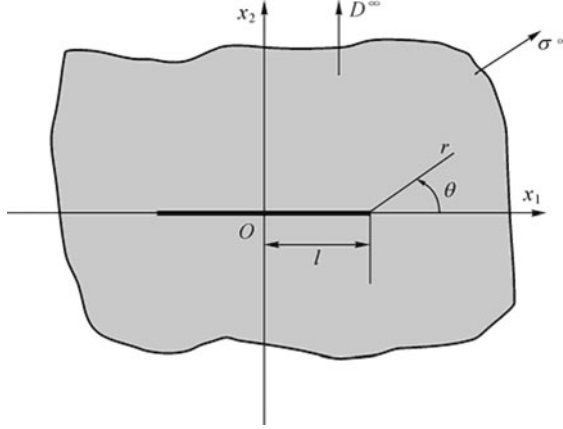


Fig. 2.4 A Griffith crack in piezoelectric solid subjected to uniform electromechanical loads at infinity

Thus, near the crack tip, we have $r \ll l$, which makes the following approximation available:

$$Z_k \approx l, \quad \sqrt{Z_k^2 - l^2} \approx \sqrt{2lr} \sqrt{\cos \theta + \mu_k \sin \theta}. \quad (2.4.60)$$

The derivative of Eq. (2.4.60) gives

$$f'_k = \frac{\sqrt{l}}{\sqrt{8r}} (A_{k1} \sigma_{22}^\infty - A_{k2} \sigma_{12}^\infty - A_{k3} D_2^\infty) \frac{1}{\sqrt{\cos \theta + \mu_k \sin \theta}} \quad (\text{not summed on } k). \quad (2.4.61)$$

Introduce

$$C_k = A_{k1} K_{\text{I}} - A_{k2} K_{\text{II}} - A_{k3} K_{\text{e}}, \quad (2.4.62)$$

where

$$K_{\text{I}} = \sqrt{l} \sigma_{22}^\infty, \quad K_{\text{II}} = \sqrt{l} \sigma_{12}^\infty, \quad K_{\text{e}} = \sqrt{l} D_2^\infty. \quad (2.4.63)$$

From Eq. (2.4.54) and Eq. (2.4.61), the generalized stresses near the crack tip can be expressed in the form

$$\begin{aligned}
 \sigma_{11} &= \frac{1}{\sqrt{2r}} \operatorname{Re} \sum_{k=1}^3 \frac{C_k \mu_k^2}{\sqrt{\cos \theta + \mu_k \sin \theta}}, \\
 \sigma_{22} &= \frac{1}{\sqrt{2r}} \operatorname{Re} \sum_{k=1}^3 \frac{C_k}{\sqrt{\cos \theta + \mu_k \sin \theta}}, \\
 \sigma_{12} &= -\frac{1}{\sqrt{2r}} \operatorname{Re} \sum_{k=1}^3 \frac{C_k \mu_k}{\sqrt{\cos \theta + \mu_k \sin \theta}}, \\
 D_1 &= \frac{1}{\sqrt{2r}} \operatorname{Re} \sum_{k=1}^3 \frac{C_k \lambda_k \mu_k}{\sqrt{\cos \theta + \mu_k \sin \theta}}, \\
 D_2 &= -\frac{1}{\sqrt{2r}} \operatorname{Re} \sum_{k=1}^3 \frac{C_k \lambda_k}{\sqrt{\cos \theta + \mu_k \sin \theta}}.
 \end{aligned} \tag{2.4.64}$$

A similar expression for the generalized strains can be obtained by using Eq. (2.4.53). The most important feature of Eq. (2.4.64) is that the characteristic singularity at the crack tip is of the order $r^{-1/2}$, which is a general conclusion in anisotropic fracture mechanics. K_I , K_{II} and K_e in Eq. (2.4.63) describe the impact of the applied load and the geometry of the crack on the crack tip fields, and are referred to as the “mechanical intensity factors” for mode I and mode II and the “electric intensity factor”, respectively.

• *J-integral for cracks in piezoelectric materials*

The solution for the electroelastic fields at the crack tip given by Eq. (2.4.64) possesses characteristic singularity of the order $r^{-1/2}$, which means that close to the crack tip the stresses and electric displacements will approach infinity. This is obviously against the physical reality, since infinite stresses will definitely cause crack growth. An alternative method to avoid this problem was constructed for elastic material by Rice (1968), named the *J*-integral. The *J*-integral for piezoelectric materials was first constructed by Suo et al. (1992).

For two bounded materials with a crack area A left on the interface, the crack faces being free of external charge and traction, the increase in the electric Gibbs free energy G_2 , according to Eq. (2.1.4) and its following paragraph, can be given as

$$dG_2 = F d\Delta - Q dV - \mathcal{G} dA, \tag{2.4.65}$$

where the system is loaded by a displacement Δ and voltage V , and F and Q denote force and charge, respectively. Comparing Eq. (2.4.65) with Eq. (2.1.4) we have

$$F\Delta = \int_V \sigma_{ij} \varepsilon_{ij} dV = \int_s n_i \sigma_{ij} u_j ds, \quad QV = \int_V D_i E_i dV = - \int_s n_i D_i \phi ds, \quad (2.4.66)$$

and

$$\mathcal{G} = - \left(\frac{\partial G_2}{\partial A} \right) \bigg|_{\Delta, V}. \quad (2.4.67)$$

In the deduction process of Eq. (2.4.66) the principle of virtual work is used. $\mathcal{G}$ can be understood as the decrease in energy G_2 associated with the creation of a unit crack area, if the displacement Δ and voltage V are kept still. For a bimaterial with the interface along the x -axis, the integral

$$J = \int (G_2 n_1 - n_i \sigma_{ip} u_{p,1} - n_i D_i \phi_{,1}) ds \quad (2.4.68)$$

vanishes over contours not enclosing any part of A . When the crack faces are free of external charge and traction,

$$\mathcal{G} = J, \quad (2.4.69)$$

where the integral path begins at one point on the lower surface of the crack face and ends at another point on the upper surface of the crack face. From Eq. (2.4.68) we see that if the integral path is chosen far enough from the crack tip, the singularity problem of the stress fields at the crack tip can be avoided. The J -integral discussed here is defined for cracks at interfaces between two bounded materials A and B. If A and B are the same material, the J -integral for a Griffith crack in a homogeneous material is obtained.

The J -integral plays the role of the energy driving force for the singular crack tip to move. It is path-independent and can be evaluated along the most convenient integral path.

2.4.2.2 A Griffith Crack at the Interface of a Piezoelectric Bimaterial

In this subsection we discuss a Griffith crack at the interface of a piezoelectric bimaterial; the result presented here was derived by Suo et al. (1992).

Consider a Griffith crack of length $2l$ lying along the x_1 axis, as illustrated in Fig. 2.4. We assume that the upper half space $x_2 > 0$ is occupied by material 1, and the lower half space $x_2 < 0$ is occupied by material 2. Terms with superscripts (1) and (2), respectively, refer to material 1 and material 2. Using the Stroh formalism, the solution form of the generalized displacement vector and the generalized stress potential function vector is given by

$$\mathbf{u} = 2 \operatorname{Re} \left\{ \sum_{\alpha=1}^4 \mathbf{a}_{\alpha} f_{\alpha}(Z_{\alpha}) \right\}, \quad \phi = 2 \operatorname{Re} \left\{ \sum_{\alpha=1}^4 \mathbf{b}_{\alpha} f_{\alpha}(Z_{\alpha}) \right\}, \quad (2.4.70)$$

or

$$\mathbf{u} = \mathbf{A}\mathbf{f}(Z_{\alpha}) + \bar{\mathbf{A}}\bar{\mathbf{f}}(\bar{Z}_{\alpha}), \quad \phi = \mathbf{B}\mathbf{f}(Z_{\alpha}) + \bar{\mathbf{B}}\bar{\mathbf{f}}(\bar{Z}_{\alpha}), \quad (2.4.71)$$

where

$$\begin{aligned} \mathbf{A} &= [\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \mathbf{a}_4], \\ \mathbf{B} &= [\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3, \mathbf{b}_4], \\ \mathbf{f}(Z_{\alpha}) &= [f_1(Z_1), f_2(Z_2), f_3(Z_3), f_4(Z_4)]^T. \end{aligned} \quad (2.4.72)$$

The boundary conditions on the interface are given by

$$\mathbf{u}^{(1)}(x) = \mathbf{u}^{(2)}(x), \quad |x| > l, \quad (2.4.73)$$

$$\mathbf{t}^{(1)}(x) = \mathbf{t}^{(2)}(x), \quad |x| > l, \quad (2.4.74)$$

where

$$\mathbf{t}^{(1)}(x) = [\sigma_{21}, \sigma_{22}, \sigma_{23}, D_2]^T. \quad (2.4.75)$$

The boundary conditions on the crack are given by

$$\mathbf{t}^{(1)}(x) = \mathbf{t}^{(2)}(x) = 0, \quad |x| < l. \quad (2.4.76)$$

Continuity of $\mathbf{t}$ is required for both the interface segment and the crack segment, which gives

$$\mathbf{B}^{(1)}\mathbf{f}'^{(1)}(x) + \bar{\mathbf{B}}^{(1)}\bar{\mathbf{f}}'^{(1)}(x) = \mathbf{B}^{(2)}\mathbf{f}'^{(2)}(x) + \bar{\mathbf{B}}^{(2)}\bar{\mathbf{f}}'^{(2)}(x), \quad -\infty < x < \infty. \quad (2.4.77)$$

Eq. (2.4.77) can be rearranged as

$$\mathbf{B}^{(1)}\mathbf{f}'^{(1)}(x) - \bar{\mathbf{B}}^{(2)}\bar{\mathbf{f}}'^{(2)}(x) = \mathbf{B}^{(2)}\mathbf{f}'^{(2)}(x) - \bar{\mathbf{B}}^{(1)}\bar{\mathbf{f}}'^{(1)}(x), \quad -\infty < x < \infty. \quad (2.4.78)$$

The left hand side of Eq. (2.4.78) is the boundary value of a function which is analytical in the upper half-space, while the right hand side is the boundary value of another function which is analytical in the lower half-space. Since both functions vanish at infinity, they must vanish for any $Z = x + py$ ($\operatorname{Im} p > 0$).

$$\mathbf{B}^{(1)}\mathbf{f}'^{(1)}(Z) - \bar{\mathbf{B}}^{(2)}\bar{\mathbf{f}}'^{(2)}(Z) = 0, \quad y > 0. \quad (2.4.79)$$

Continuity of the generalized displacement on the interface requires that

$$\mathbf{A}^{(1)}\mathbf{f}^{(1)}(x) + \bar{\mathbf{A}}^{(1)}\bar{\mathbf{f}}^{(1)}(x) = \mathbf{A}^{(2)}\mathbf{f}^{(2)}(x) + \bar{\mathbf{A}}^{(2)}\bar{\mathbf{f}}^{(2)}(x), \quad |x| > l. \quad (2.4.80)$$

Define

$$\mathbf{Y} = i\mathbf{A}\mathbf{B}^{-1}, \quad (2.4.81)$$

and

$$\mathbf{H} = \mathbf{Y}^{(1)} + \bar{\mathbf{Y}}^{(2)}. \quad (2.4.82)$$

$\mathbf{H}$ and $\mathbf{Y}$ possess the same properties. $\mathbf{H}$ is Hermitian, independent of the normalization for $\mathbf{A}$. $\mathbf{Y}$ is defined for one material while $\mathbf{H}$ is defined for a bimaterial. After manipulation we obtain from Eq. (2.4.80) that

$$\mathbf{H}\mathbf{B}^{(1)}\mathbf{f}'^{(1)}(x) = \bar{\mathbf{H}}\mathbf{B}^{(2)}\bar{\mathbf{f}}'^{(2)}(x), \quad |x| > l. \quad (2.4.83)$$

Continuity of displacement across the interface is satisfied when the following function

$$\mathbf{h}(Z_\alpha) = \begin{cases} \mathbf{B}^{(1)}\mathbf{f}'^{(1)}(Z_\alpha), & y > 0 \\ \mathbf{H}^{-1}\bar{\mathbf{H}}\mathbf{B}^{(1)}\mathbf{f}'^{(1)}(Z_\alpha), & y < 0 \end{cases} \quad (2.4.84)$$

is analytic in the whole plane except on the crack lines. The remaining task is to find an appropriate $\mathbf{h}(Z_\alpha)$.

• *Crack tip field*

Consider the asymptotic problem first. The boundary condition for the crack given by Eq. (2.4.76) can be recast as

$$\mathbf{h}^+(x) + \bar{\mathbf{H}}^{-1}\mathbf{H}\mathbf{h}^-(x) = 0, \quad |x| < l. \quad (2.4.85)$$

For isotropic piezoelectric bimaterial, the solution was obtained by Suo (1992) as

$$\mathbf{h}(Z) = \frac{e^{\pi\varepsilon}KZ^{i\varepsilon}\mathbf{w} + e^{-\pi\varepsilon}\bar{K}\bar{Z}^{-i\varepsilon}\bar{\mathbf{w}}}{2\sqrt{2\pi Z} \cosh \pi\varepsilon} + \frac{K_3Z^\kappa\mathbf{w}_3 + K_4Z^{-\kappa}\mathbf{w}_4}{2\sqrt{2\pi Z} \cosh \pi\kappa}, \quad (2.4.86)$$

where K (complex), K_3 and K_4 (real) are called intensity factors, and ε and κ are determined by

$$\begin{aligned}\varepsilon &= \frac{1}{\pi} \operatorname{arctanh}[(b^2 - c)^{1/2} - b]^{1/2}, \\ \kappa &= \frac{1}{\pi} \operatorname{arctanh}[(b^2 - c)^{1/2} + b]^{1/2}.\end{aligned}\quad (2.4.87)$$

And

$$b = \frac{1}{4} \operatorname{tr}[(\mathbf{D}^{-1} \mathbf{W})^2], \quad c = \det(\mathbf{D}^{-1} \mathbf{W}), \quad (2.4.88)$$

$$\mathbf{H} = \mathbf{D} + i \mathbf{W}. \quad (2.4.89)$$

In Eq. (2.4.89), the matrices $\mathbf{D}$ and $\mathbf{W}$ are both real while $\mathbf{H}$ is complex. In a solution given by Eq. (2.4.86), $\mathbf{w}$, $\mathbf{w}_3$ and $\mathbf{w}_4$ are determined by the applied loads. The solution of the asymptotic problem permits understanding of the behavior of the bimaterial near the crack tip. For example, the jump in the generalized displacement $\mathbf{d}$ at a distance r behind the tip is

$$\begin{aligned}\mathbf{d} &= (\mathbf{H} + \bar{\mathbf{H}}) \sqrt{\frac{r}{2\pi}} \left[\frac{K r^{i\varepsilon} \mathbf{w}}{(1 + 2i\varepsilon) \cosh(\pi\varepsilon)} + \frac{\bar{K} r^{-i\varepsilon} \bar{\mathbf{w}}}{(1 - 2i\varepsilon) \cosh(\pi\varepsilon)} \right. \\ &\quad \left. + \frac{K_3 r^{\kappa} \mathbf{w}_3}{(1 + 2\kappa) \cos(\pi\kappa)} + \frac{K_4 r^{-\kappa} \mathbf{w}_4}{(1 - 2\kappa) \cos(\pi\kappa)} \right].\end{aligned}\quad (2.4.90)$$

The energy release rate related to crack growth, which is given by the J -integral and discussed in Section 2.4.2.1, is given by

$$\mathcal{G} = \frac{\bar{\mathbf{w}}(\mathbf{H} + \bar{\mathbf{H}})\mathbf{w}}{4 \cosh^2(\pi\varepsilon)} |K|^2 + \frac{\mathbf{w}_3(\mathbf{H} + \bar{\mathbf{H}})\mathbf{w}_4}{4 \cosh^2(\pi\kappa)} K_3 K_4. \quad (2.4.91)$$

• Full-field solutions

Considering a set of applied traction-charge loads prescribed on the crack surface by a vector $\mathbf{t}_0(x)$, the boundary condition is then given by

$$\mathbf{h}^+(x) + \bar{\mathbf{H}}^{-1} \mathbf{H} \mathbf{h}^-(x) = -\mathbf{t}_0(x), \quad |x| < l. \quad (2.4.92)$$

If we expand $\mathbf{h}(x)$ and $\mathbf{t}_0(x)$ in the $\mathbf{w}$ -space (assuming that $\mathbf{w}$ is normalized), it gives

$$\begin{aligned}\mathbf{h}(x) &= h_1 \mathbf{w} + h_2 \bar{\mathbf{w}} + h_3 \mathbf{w}_3 + h_4 \mathbf{w}_4, \\ \mathbf{t}_0(x) &= t \mathbf{w} + \bar{t} \bar{\mathbf{w}} + t_3 \mathbf{w}_3 + t_4 \mathbf{w}_4.\end{aligned}\quad (2.4.93)$$

Eq. (2.4.92) is rewritten as

$$\begin{aligned}
h_1^+ + e^{-2\pi\epsilon} h_1^- &= -t, \\
h_2^+ + e^{+2\pi\epsilon} h_2^- &= -\bar{t}, \\
h_3^+ + e^{2\pi i\kappa} h_3^- &= -t_3, \\
h_4^+ + e^{-2\pi i\kappa} h_4^- &= -t_4.
\end{aligned} \tag{2.4.94}$$

The solution of Eq. (2.4.94) for isotropic piezoelectric material is given by

$$\begin{aligned}
\int h_1(Z) dZ &= \frac{t}{1 + e^{-2\pi\epsilon}} \left[\left(\frac{Z-a}{Z+a} \right)^{i\epsilon} (Z^2 - a^2)^{1/2} - Z \right], \\
\int h_2(Z) dZ &= \frac{\bar{t}}{1 + e^{2\pi\epsilon}} \left[\left(\frac{Z-a}{Z+a} \right)^{-i\epsilon} (Z^2 - a^2)^{1/2} - Z \right], \\
\int h_3(Z) dZ &= \frac{t_3}{1 + e^{2\pi i\kappa}} \left[\left(\frac{Z-a}{Z+a} \right)^{\kappa} (Z^2 - a^2)^{1/2} - Z \right], \\
\int h_4(Z) dZ &= \frac{t_4}{1 + e^{-2\pi i\kappa}} \left[\left(\frac{Z-a}{Z+a} \right)^{-\kappa} (Z^2 - a^2)^{1/2} - Z \right].
\end{aligned} \tag{2.4.95}$$

The intensity factors are

$$\begin{aligned}
K &= t(1 + 2i\epsilon) \cosh(\pi\epsilon)(\pi a)^{1/2} (2a)^{-i\epsilon}, \\
K_3 &= t_3(1 + 2\kappa) \cos(\pi\kappa)(\pi a)^{1/2} (2a)^{-\kappa}, \\
K_4 &= t_4(1 - 2\kappa) \cos(\pi\kappa)(\pi a)^{1/2} (2a)^{\kappa}.
\end{aligned} \tag{2.4.96}$$

2.4.2.3 Nonlocal Theory for Cracks

As seen in Section 2.4.2.1, the solution for the crack tip fields using classic field theory possesses characteristic singularity of the order $r^{-1/2}$. Although, by using the J -integral, it is possible to obtain the crack moving force by choosing an integral path far enough away from the tip, the actual values of the electroelastic fields at the crack tip are yet unreasonable. The nonlocal continuum field theory can be used to solve this problem. For elastic materials, Eringen et al. (Eringen et al., 1977; Eringen, 1977; 1978; 1979) had successfully used the method to discuss the crack tip fields on an elastic plate subject to uniform tension, shear and anti-plane shear. Zhou et al. (1999; 2002a; 2002b; 2003; 2004; 2006) used the method to deal with Griffith's cracks in piezoelectric materials. Although the nonlocal theory is powerful, due to its mathematical complexity, the solutions based on this method are limited and restricted to piezoelectric materials with isotropy. Zhou et al. (1999) illustrated the manipulation process of this method.

A basic standpoint of the nonlocal theory is that the behavior of one material point is influenced by the state of all points of the body. Consider the behavior of a Griffith crack in an isotropic piezoelectric solid, as plotted in Fig. 2.4, subjected to uniform anti-plane shear strains. The basic equations are given by

$$\frac{\partial \tau_{xz}}{\partial x} + \frac{\partial \tau_{yz}}{\partial y} = 0, \quad (2.4.97)$$

$$\frac{\partial D_x}{\partial x} + \frac{\partial D_y}{\partial y} = 0, \quad (2.4.98)$$

$$\tau_{kz}(\mathbf{x}) = \int_V [C'_{44}(|\mathbf{x}' - \mathbf{x}|) u_{3,k}(\mathbf{x}') + e'_{15}(|\mathbf{x}' - \mathbf{x}|) \phi_k(\mathbf{x}')] dV(\mathbf{x}') \quad (k = 1, 2), \quad (2.4.99)$$

$$D_k(\mathbf{x}) = \int_V [e'_{15}(|\mathbf{x}' - \mathbf{x}|) u_{3,k}(\mathbf{x}') + \alpha'_{11}(|\mathbf{x}' - \mathbf{x}|) \phi_k(\mathbf{x}')] dV(\mathbf{x}') \quad (k = 1, 2), \quad (2.4.100)$$

where τ_{kz} , u_3 and ϕ denote the shear stress, the displacement in thickness direction and the electric potential, respectively. In Eqs. (2.4.99) and (2.4.100), it is shown that the shear stress and electric displacement $\tau_{kz}(\mathbf{x})$ and $D_k(\mathbf{x})$ at any point $\mathbf{x}$ are related to the $u_{3,k}(\mathbf{x}')$ and $\phi_k(\mathbf{x}')$ at all points of the body. The three independent material constants $C'_{44}(|\mathbf{x}' - \mathbf{x}|)$, $e'_{15}(|\mathbf{x}' - \mathbf{x}|)$ and $\alpha'_{11}(|\mathbf{x}' - \mathbf{x}|)$ are assumed to be functions of the distance $(|\mathbf{x}' - \mathbf{x}|)$. The expressions of $C'_{44}(|\mathbf{x}' - \mathbf{x}|)$, $e'_{15}(|\mathbf{x}' - \mathbf{x}|)$ and $\alpha'_{11}(|\mathbf{x}' - \mathbf{x}|)$ are called dispersion relations. Zhou et al. (1999) quoted Eringen (1974; 1977)'s results obtained from lattice dynamics and wrote $C'_{44}(|\mathbf{x}' - \mathbf{x}|)$, $e'_{15}(|\mathbf{x}' - \mathbf{x}|)$ and $\alpha'_{11}(|\mathbf{x}' - \mathbf{x}|)$ as

$$(C'_{44}, e'_{15}, \alpha'_{11}) = (C_{44}, e_{15}, \alpha_{11}) \lambda(|\mathbf{x}' - \mathbf{x}|), \quad (2.4.101)$$

$$\lambda(|\mathbf{x}' - \mathbf{x}|) = \lambda_0 \exp[-(\beta/a)^2 (\mathbf{x}' - \mathbf{x})^T (\mathbf{x}' - \mathbf{x})], \quad (2.4.102)$$

where β is a constant, a is the lattice parameter, and $C_{44}, e_{15}, \alpha_{11}$ are the macroscopic shear modulus, piezoelectric coefficient and dielectric parameter, respectively. When $u_{3,k}(\mathbf{x}')$ and $\phi_k(\mathbf{x}')$ in Eqs. (2.4.99) and (2.4.100) are invariant to the position $\mathbf{x}'$, the normalization relation is obtained:

$$\int_V \lambda(|\mathbf{x}' - \mathbf{x}|) dV(\mathbf{x}') = 1. \quad (2.4.103)$$

For two-dimensional problems, substitution of Eq. (2.4.102) into Eq. (2.4.103) gives

$$\lambda_0 = \frac{1}{\pi}(\beta/a)^2. \quad (2.4.104)$$

Eqs. (2.4.99) and (2.4.100) can be rewritten as

$$\tau_{kz}(\mathbf{x}) = \int_V \lambda(|\mathbf{x}' - \mathbf{x}|) \sigma_{kz}(\mathbf{x}') dV(\mathbf{x}') \quad (k=1, 2), \quad (2.4.105)$$

$$D_k(\mathbf{x}) = \int_V \lambda(|\mathbf{x}' - \mathbf{x}|) D_k^c(\mathbf{x}') dV(\mathbf{x}') \quad (k=1, 2), \quad (2.4.106)$$

where

$$\begin{aligned} \sigma_{kz} &= C_{44} u_{3,k} + e_{15} \phi_{,k}, \\ D_k^c &= e_{15} u_{3,k} - \alpha_{11} \phi_{,k}. \end{aligned} \quad (2.4.107)$$

Eq. (2.4.107) is the classic constitutive equation. For a uniform shear strain $\tau_{yz} = \tau_0$ and electric displacement $D_y = D_0$ applied along the surface of the crack, and by considering only the out-of-plane displacement and in-plane elastic fields, the boundary conditions are given by

$$\tau_{yz}(x, 0) = \tau_0, \quad |\mathbf{x}| \leq l, \quad (2.4.108)$$

$$D_y(x, 0) = D_0, \quad |\mathbf{x}| \leq l, \quad (2.4.109)$$

$$u_3(x, 0) = \phi_0(x, 0) = 0, \quad |\mathbf{x}| \leq l, \quad (2.4.110)$$

$$u_3(x, y) = \phi_0(x, y) = 0, \quad |\mathbf{x}| \rightarrow \infty. \quad (2.4.111)$$

For a piezoelectric body containing a Griffith crack, substitution of Eqs. (2.4.105) and (2.4.106) into Eqs. (2.4.97) and (2.4.98) yields

$$\begin{aligned} & \iint_V \lambda(|\mathbf{x}' - \mathbf{x}|, |\mathbf{y}' - \mathbf{y}|) [C_{44} \nabla^2 u_3(\mathbf{x}', \mathbf{y}') + e_{15} \nabla^2 \phi(\mathbf{x}', \mathbf{y}')] d\mathbf{x}' d\mathbf{y}' \\ & - \int_{-l}^l \lambda(|\mathbf{x}' - \mathbf{x}|, 0) \llbracket \sigma_{yz}(\mathbf{x}', 0) \rrbracket d\mathbf{x}' = 0, \end{aligned} \quad (2.4.112)$$

$$\begin{aligned} & \iint_V \lambda(|\mathbf{x}' - \mathbf{x}|, |\mathbf{y}' - \mathbf{y}|) [e_{15} \nabla^2 u_3(\mathbf{x}', \mathbf{y}') - \alpha_{11} \nabla^2 \phi(\mathbf{x}', \mathbf{y}')] d\mathbf{x}' d\mathbf{y}' \\ & - \int_{-l}^l \lambda(|\mathbf{x}' - \mathbf{x}|, 0) \llbracket D_y^c(\mathbf{x}', 0) \rrbracket d\mathbf{x}' = 0, \end{aligned} \quad (2.4.113)$$

where $\llbracket \sigma_{yz}(\mathbf{x}', 0) \rrbracket$ denotes a jump across the crack line and ∇^2 denotes the two-dimensional Laplace operator. In the derivation of Eqs. (2.4.112) and (2.4.113), the Green-Gauss theorem was used. Under the anti-plane boundary conditions, the displacement and the electric potential satisfy the following relations:

$$u_3(x, -y) = -u_3(x, y), \quad \phi(x, -y) = -\phi(x, y). \quad (2.4.114)$$

According to Eq. (2.4.114), on the surface of the crack we have

$$\llbracket \sigma_{yz}(x', 0) \rrbracket = \llbracket D_y^e(x', 0) \rrbracket = 0. \quad (2.4.115)$$

Thus Eqs. (2.4.112) and (2.4.113) can be rewritten as

$$\iint_V \lambda(|x' - x|, |y' - y|) [C_{44} \nabla^2 u_3(x', y') + e_{15} \nabla^2 \phi(x', y')] dx' dy' = 0, \quad (2.4.116)$$

$$\iint_V \lambda(|x' - x|, |y' - y|) [e_{15} \nabla^2 u_3(x', y') - \alpha_{11} \nabla^2 \phi(x', y')] dx' dy' = 0. \quad (2.4.117)$$

The solution for Eqs. (2.4.116) and (2.4.117) is immediately obtained by using the Fourier transform technique as follows:

$$u_3(x, y) = \frac{2}{\pi} \int_0^\infty A(s) e^{-sy} \cos(sx) ds, \quad (2.4.118)$$

$$\phi(x, y) = \frac{2}{\pi} \int_0^\infty \left[\frac{e_{15}}{\alpha_{11}} A(s) + B(s) \right] e^{-sy} \cos(sx) ds, \quad (2.4.119)$$

where the form of $A(s)$ and $B(s)$ are to be determined by the boundary conditions. Substitution of Eqs. (2.4.116) and (2.4.117) into Eqs. (2.4.108)–(2.4.111) and further manipulations lead to the following results:

$$A(s) = \sum_{n=1}^{\infty} a_n B_n \frac{1}{s} J_{2n-1}(sl), \quad (2.4.120)$$

$$B(s) = \sum_{n=1}^{\infty} \left(b_n - \frac{e_{15}}{\alpha_{11}} a_n \right) B_n \frac{1}{s} J_{2n-1}(sl), \quad (2.4.121)$$

$$B_n = 2\sqrt{\pi}(-1)^{n-1} \frac{\Gamma(2n - \frac{1}{2})}{(2n - 2)!}, \quad (2.4.122)$$

where $\Gamma(n)$ and $J_n(x)$ are the Gamma and Bessel functions, respectively. The coefficients a_n and b_n are determined by

$$\sum_{n=1}^{\infty} a_n B_n \int_0^\infty \operatorname{erfc}\left(\frac{a}{2\beta}s\right) J_{2n-1}(sl) \cos(sx) ds = -\frac{\pi}{2\mu} \tau_0 (1 + \chi), \quad (2.4.123)$$

$$\sum_{n=1}^{\infty} (b_n - \frac{e_{15}}{\alpha_{11}} a_n) B_n \int_0^{\infty} \operatorname{erfc}(\frac{a}{2\beta} s) J_{2n-1}(sl) \cos(sx) ds = \frac{\pi D_0}{2\alpha_{11}}, \quad (2.4.124)$$

where

$$\chi = \frac{e_{15} D_0}{\alpha_{11} \tau_0}, \quad (2.4.125)$$

$$\operatorname{erfc}(s) = 1 - \Phi(s), \quad \Phi(s) = \frac{2}{\sqrt{\pi}} \int_0^s \exp(-t^2) dt, \quad \mu = C_{44} + \frac{e_{15}^2}{\alpha_{11}}. \quad (2.4.126)$$

Eqs. (2.4.123) and (2.4.124) were treated numerically in Zhou et al. (1999) by using the Schmidt method. For details of the calculation methods and the deduction process, readers should refer to Zhou et al. (1999).

The shear stresses and electric displacement are given by

$$\begin{aligned} \tau_{yz}(x, y) = & \frac{2}{\pi} \int_0^{\infty} [-\mu s A(s) - e_{15} s B(s)] ds \int_0^{\infty} dy' \int_{-\infty}^{\infty} [\lambda(|x' - x|, |y' - y|) \\ & + \lambda(|x' - x|, |y' + y|)] e^{-sy'} \cos(sx') dx', \end{aligned} \quad (2.4.127)$$

$$\begin{aligned} D_y(x, y) = & \frac{2}{\pi} \int_0^{\infty} \alpha_{11} s B(s) ds \int_0^{\infty} dy' \int_{-\infty}^{\infty} [\lambda(|x' - x|, |y' - y|) \\ & + \lambda(|x' - x|, |y' + y|)] e^{-sy'} \cos(sx') dx'. \end{aligned} \quad (2.4.128)$$

The shear stresses and electric displacement along the crack line are given by

$$\tau_{yz}(x, 0) = \frac{2\mu}{\pi(1 + \chi)} \sum_{n=1}^{\infty} a_n B_n \int_0^{\infty} \operatorname{erfc}(\frac{a}{2\beta} s) J_{2n-1}(sl) \cos(xs) ds, \quad (2.4.129)$$

$$D_y(x, 0) = \frac{D_0}{\tau_0} \tau_{yz}(x, 0). \quad (2.4.130)$$

For different values of $a/(2\beta l)$, the distributions of shear stresses along the x axis were plotted in Zhou et al. (1999).

The following conclusions can be made from the solutions obtained by using the nonlocal theory:

(1) No singularity exists at the crack tip for the electroelastic fields unless the lattice parameter $a \rightarrow 0$ or the crack length $l \rightarrow \infty$. In other words, for a decreasing lattice parameter or increasing crack length, the stress concentration enhances.

(2) The maximum stress does not occur at the crack tip, but at points slightly

beside it. This phenomenon had been thoroughly discussed by Eringen (1983). The distance between the crack tip and the maximum stress point is very small and it depends on the crack length and the lattice parameter.

(3) The significance of this result is that the fracture criteria are unified at both macroscopic and microscopic scales; viz., it may solve the problem of any scale cracks (by accommodating the value of $a/(2\beta l)$).

2.4.3 Dislocations and Line Charges

Let a line charge q per unit length coincident with a straight dislocation with Burgers vector $\mathbf{b}$ be applied to the x_3 axis at the point of origin of the (x_1, x_2) -plane. Without loss of generality, define a slip plane at $x_2 = 0$, $x_1 < 0$. For a polar coordinate

$$r > 0, \quad -\pi \leq \theta \leq \pi, \quad (2.4.131)$$

the boundary conditions at the origin are given by

$$\mathbf{u}(r, \pi) - \mathbf{u}(r, -\pi) = \mathbf{b}, \quad \phi(r, \pi) - \phi(r, -\pi) = \mathbf{f}, \quad (2.4.132)$$

where

$$b_4 = 0, \quad f_1 = f_2 = f_3 = 0, \quad f_4 = -q. \quad (2.4.133)$$

For $\mathbf{u}$ and ϕ , the two types of core functions can be constructed as follows:

$$\begin{aligned} \mathbf{u}^I &= 2 \operatorname{Re} \left\{ \mathbf{A} \langle \mathbf{In}(\zeta_\alpha) \rangle \mathbf{A}^T \right\} \mathbf{g}_0 + 2 \operatorname{Re} \left\{ \mathbf{A} \langle \mathbf{In}(\zeta_\alpha) \rangle \mathbf{B}^T \right\} \mathbf{h}_0, \\ \phi^I &= 2 \operatorname{Re} \left\{ \mathbf{B} \langle \mathbf{In}(\zeta_\alpha) \rangle \mathbf{A}^T \right\} \mathbf{g}_0 + 2 \operatorname{Re} \left\{ \mathbf{B} \langle \mathbf{In}(\zeta_\alpha) \rangle \mathbf{B}^T \right\} \mathbf{h}_0, \end{aligned} \quad (2.4.134)$$

and

$$\begin{aligned} \mathbf{u}^{II} &= 2 \sum_{k=1}^{\infty} \operatorname{Re} \left\{ \mathbf{A} \langle \zeta_\alpha^{-k} \rangle \mathbf{A}^T \right\} \mathbf{g}_k + 2 \sum_{k=1}^{\infty} \operatorname{Re} \left\{ \mathbf{A} \langle \zeta_\alpha^{-k} \rangle \mathbf{B}^T \right\} \mathbf{h}_k, \\ \phi^{II} &= 2 \sum_{k=1}^{\infty} \operatorname{Re} \left\{ \mathbf{B} \langle \zeta_\alpha^{-k} \rangle \mathbf{A}^T \right\} \mathbf{g}_k + 2 \sum_{k=1}^{\infty} \operatorname{Re} \left\{ \mathbf{B} \langle \zeta_\alpha^{-k} \rangle \mathbf{B}^T \right\} \mathbf{h}_k. \end{aligned} \quad (2.4.135)$$

$\mathbf{In}(\zeta_\alpha)$ is the only one which is multiple-valued of θ (it increases its value by $2\pi i$ when θ is increased by 2π) and the increased value is independent of α . This relation can be mathematically expressed as

$$\mathbf{In}(\zeta_\alpha + \pi) - \mathbf{In}(\zeta_\alpha - \pi) = 2\pi i. \quad (2.4.136)$$

Thus we have

$$\begin{bmatrix} \mathbf{u}^I(r, \pi) - \mathbf{u}^I(r, -\pi) \\ \boldsymbol{\phi}^I(r, \pi) - \boldsymbol{\phi}^I(r, -\pi) \end{bmatrix} = 2\pi \begin{bmatrix} \mathbf{S} & \mathbf{H} \\ -\mathbf{L} & \mathbf{S}^T \end{bmatrix} \begin{bmatrix} \mathbf{h}_0 \\ \mathbf{g}_0 \end{bmatrix}, \quad (2.4.137)$$

where $\mathbf{S}$, $\mathbf{H}$ and $\mathbf{L}$ are real matrices defined in Eq. (2.3.23). From Eq. (2.4.137) we have

$$\begin{bmatrix} \mathbf{h}_0 \\ \mathbf{g}_0 \end{bmatrix} = \frac{1}{2\pi} \begin{bmatrix} \mathbf{S} & \mathbf{H} \\ -\mathbf{L} & \mathbf{S}^T \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{b} \\ \mathbf{f} \end{bmatrix}. \quad (2.4.138)$$

The expressions for the potentials are finally given by Eq. (2.4.134). The results obtained in Eq. (2.4.138) can obviously be extended to include a coincident line force $\mathbf{f} = [f_1, f_2, f_3, 0]$ and a coincident electric dipole layer $b_4 \neq 0$ along the tip of the slip plane $x_2 = 0$, $x_1 < 0$.

2.5 Solution Methods for Three-Dimensional Problems

In this section, methods for the solution of the elastic and electric fields related to all three components x_1, x_2, x_3 of the coordinate for piezoelectric materials are constructed. The study of three-dimensional problems for elastic materials has a rather long history. A lot of fundamental work considering the effects of heterostructures such as inclusions, dislocations, cracks, etc., was done by Eshelby (1951; 1957; 1959). The most comprehensive and systematic summary of the basic three-dimensional micromechanical problems for anisotropic elastic bodies was given in Mura (1987). A clear expatiation on the physical concept of eigenstrain, inclusion and inhomogeneity was also presented in that book.

For piezoelectric solids, these concepts are explained here, as well as several analytical solution methods which are commonly used in treating three-dimensional problems.

2.5.1 Eigenstrains and Equivalent Inclusion Method

The word ‘eigenstrains’ refers to stress-free strains, such as strains caused by thermal expansion, phase transition, crystal lattice misfit, etc. In a homogeneous

piezoelectric body occupying R , when a certain distribution of eigenstrains ε_{ij}^* is prescribed in a subdomain Ω , Ω is called an inclusion of the piezoelectric body. Inside the inclusion, the total strain satisfies

$$\varepsilon_{ij} = \varepsilon_{ij}^e + \varepsilon_{ij}^*, \quad (2.5.1)$$

where ε_{ij}^e denote the elastic strains. The corresponding constitutive equations are given by

$$\begin{aligned} \sigma_{ij} &= C_{ijkl}^E (\varepsilon_{kl} - \varepsilon_{kl}^*) - e_{kij} E_k, \\ D_i &= e_{ikl} (\varepsilon_{kl} - \varepsilon_{kl}^*) + \alpha_{ik}^e E_k, \end{aligned} \quad (2.5.2)$$

or

$$\sigma_{iJ} = C_{iJKl} (\varepsilon_{KL} - \varepsilon_{KL}^*), \quad (2.5.3)$$

where C_{iJKl} , σ_{iJ} and ε_{KL} are defined in Eqs. (2.1.30) and (2.1.31), and

$$\varepsilon_{KL}^* = \begin{cases} \varepsilon_{kl}^*, & K = 1, 2, 3 \\ 0, & K = 4 \end{cases}. \quad (2.5.4)$$

Inside the inclusion Eq. (2.1.28) becomes

$$C_{iJKl} u_{K,li} = C_{iJKl} \varepsilon_{KL,i}^*. \quad (2.5.5)$$

Eq. (2.5.5) can be used to describe the equilibrium equation for the whole piezoelectric body, once we define the expression of ε_{KL}^* in R as

$$\varepsilon_{KL}^* = f_{KL}(\mathbf{x}) H(\mathbf{x}), \quad \mathbf{x} \in R, \quad (2.5.6)$$

where

$$H(\mathbf{x}) = \begin{cases} 1, & \mathbf{x} \in \Omega \\ 0, & \mathbf{x} \notin \Omega \end{cases}, \quad (2.5.7)$$

and $f_{KL}(\mathbf{x})$ denotes the distribution of the eigenstrains ε_{KL}^* in Ω .

Now we obtain the equilibrium equation for a piezoelectric body with inclusions. Yet the definition of inclusion is based on a homogeneous body, which indicates uniform material constants inside and outside the inclusion. For a piezoelectric body with material constants denoted by C_{iJKl} occupying R , if a

subdomain Ω of R is replaced by another kind of piezoelectric material whose material constants are denoted by C_{ijkl}^* , Ω is called an inhomogeneity. The equilibrium equation for the piezoelectric material in R is given by

$$C_{ijkl} u_{K,li} = \Delta C_{ijkl} u_{K,li}, \quad (2.5.8)$$

where

$$\Delta C_{ijkl} = (C_{ijkl} - C_{ijkl}^*) H(\mathbf{x}). \quad (2.5.9)$$

Letting

$$C_{ijkl} \varepsilon_{kl}^{**} = \Delta C_{ijkl} u_{K,li}, \quad (2.5.10)$$

Eq. (2.5.8) can be transformed to

$$C_{ijkl} u_{K,li} = C_{ijkl} \varepsilon_{kl,i}^{**}, \quad (2.5.11)$$

which is similar to Eq. (2.5.5). ε_{kl}^{**} is called the equivalent eigenstrain of the inhomogeneity. By using this equivalent inclusion method, it is possible to establish a relation between the solutions obtained for inclusions and for inhomogeneities and avoid redundant calculations for the two problems. In practical problems, the two problems may be coupled. For example, in the phase transition system, there will be spontaneous deformation associated with the phase transition, as well as a change in the elastic constants. As before, we denote such a subdomain as Ω : Ω is called an inhomogeneous inclusion. For a set of eigenstrains ε_{kl}^* prescribed in an inhomogeneity Ω , whose material constants are denoted by C_{ijkl}^* , the equilibrium equation is given by

$$C_{ijkl} u_{K,li} = \Delta C_{ijkl} u_{K,li} + C_{ijkl}^* \varepsilon_{kl,i}^*. \quad (2.5.12)$$

The equivalent inclusion method gives

$$C_{ijkl} u_{K,li} = C_{ijkl} (\varepsilon_{kl,i}^* + \varepsilon_{kl,i}^{**}), \quad (2.5.13)$$

where ε_{kl}^{**} is solved from

$$\Delta C_{ijkl} (u_{K,li} - \varepsilon_{kl,i}^*) = C_{ijkl} \varepsilon_{kl,i}^{**}. \quad (2.5.14)$$

Up to now, we can conclude that the solution of Eq. (2.5.5) is the fundamental problem for all kinds of heterostructures. In the following section, possible solution methods of Eq. (2.5.5) are introduced.

2.5.2 Method of Fourier Integrals

Due to the linear nature of Eq. (2.1.28), a superposition of solutions is allowed. Assume a distribution of eigenstrains ε_{KI}^* given by the Fourier integral form

$$\varepsilon_{KI}^*(\mathbf{x}) = \int_{-\infty}^{\infty} \bar{\varepsilon}_{KI}^*(\xi) \exp(i\xi \cdot \mathbf{x}) d\xi, \quad (2.5.15)$$

where

$$\bar{\varepsilon}_{KI}^*(\xi) = (-2\pi)^{-3} \int_{-\infty}^{\infty} \varepsilon_{KI}^*(\mathbf{x}) \exp(-i\xi \cdot \mathbf{x}) d\mathbf{x} \quad (2.5.16)$$

and

$$\int_{-\infty}^{\infty} d\xi = \int_{-\infty}^{\infty} d\xi_1 \int_{-\infty}^{\infty} d\xi_2 \int_{-\infty}^{\infty} d\xi_3. \quad (2.5.17)$$

The total eigenstrains $\varepsilon_{KI}^*(\mathbf{x})$ defined in Eq. (2.5.15) can be considered as a cumulated effect of a set of periodic eigenstrains $\bar{\varepsilon}_{KI}^*(\xi) \exp(i\xi \cdot \mathbf{x})$ with continuously varying wave vector $\xi = (\xi_1, \xi_2, \xi_3)$. Thus it is first necessary to discuss the displacement fields caused by the periodic eigenstrains

$$\varepsilon_{KI}^*(\mathbf{x}) = \bar{\varepsilon}_{KI}^*(\xi) \exp(i\xi \cdot \mathbf{x}). \quad (2.5.18)$$

The generalized displacement fields take the form

$$u_J(\mathbf{x}) = \bar{u}_J(\xi) \exp(i\xi \cdot \mathbf{x}). \quad (2.5.19)$$

Substitution of Eqs. (2.5.18) and (2.5.19) into Eq. (2.5.5) yields

$$C_{iJKI} \bar{u}_K \xi_i \xi_i = -i C_{iJKI} \bar{\varepsilon}_{KI}^* \xi_i. \quad (2.5.20)$$

Eq. (2.5.20) gives four equations ($J = 1, 2, 3, 4$) which determine $\bar{u}_K$ for given $\bar{\varepsilon}_{KI}^*(\xi)$. The solution for Eq. (2.5.20) is given by

$$\bar{u}_K(\mathbf{x}) = -i C_{iJKI} \bar{\varepsilon}_{KI}^*(\xi) \xi_i [D^{-1}]_{JK} \exp(i\xi \cdot \mathbf{x}), \quad (2.5.21)$$

where $[D^{-1}]_{JK}$ denotes the components of the inverse matrix of $\mathbf{D}$, whose components are defined by

$$D_{JK} = C_{iJKI} \xi_i \xi_i. \quad (2.5.22)$$

The corresponding strain and stress for Eq. (2.5.21) are

$$\varepsilon_{Kl}(\mathbf{x}) = \begin{cases} \frac{1}{2} C_{iJ Mn} \bar{\varepsilon}_{Mn}^*(\xi) \xi_i \{ [D^{-1}]_{JK} \xi_l + [D^{-1}]_{Jl} \xi_K \} \exp(i\xi \cdot \mathbf{x}), & K = 1, 2, 3 \\ C_{iJ Mn} \bar{\varepsilon}_{Mn}^*(\xi) \xi_i [D^{-1}]_{JK} \xi_l \exp(i\xi \cdot \mathbf{x}), & K = 4 \end{cases} \quad (2.5.23)$$

and

$$\sigma_{ij}(\mathbf{x}) = C_{iJKl} \{ C_{pQMn} \bar{\varepsilon}_{Mn}^*(\xi) \xi_p [D^{-1}]_{QK} \xi_l \exp(i\xi \cdot \mathbf{x}) - \varepsilon_{Kl}^*(\mathbf{x}) \}. \quad (2.5.24)$$

Now let us get back to the eigenstrain expression given by Eq. (2.5.15). The corresponding expressions for the electroelastic fields are

$$u_K(\mathbf{x}) = -i \int_{-\infty}^{\infty} C_{iJ Mn} \bar{\varepsilon}_{Mn}^*(\xi) \xi_i [D^{-1}]_{JK} \exp(i\xi \cdot \mathbf{x}) d\xi, \quad (2.5.25)$$

$$\varepsilon_{Kl}(\mathbf{x}) = \begin{cases} \frac{1}{2} \int_{-\infty}^{\infty} C_{iJ Mn} \bar{\varepsilon}_{Mn}^*(\xi) \xi_i \{ [D^{-1}]_{JK} \xi_l + [D^{-1}]_{Jl} \xi_K \} \exp(i\xi \cdot \mathbf{x}) d\xi, & K = 1, 2, 3 \\ \int_{-\infty}^{\infty} C_{iJ Mn} \bar{\varepsilon}_{Mn}^*(\xi) \xi_i [D^{-1}]_{JK} \xi_l \exp(i\xi \cdot \mathbf{x}) d\xi, & K = 4 \end{cases} \quad (2.5.26)$$

and

$$\sigma_{ij}(\mathbf{x}) = C_{iJKl} \{ \int_{-\infty}^{\infty} C_{pQMn} \bar{\varepsilon}_{Mn}^*(\xi) \xi_p [D^{-1}]_{QK} \xi_l \exp(i\xi \cdot \mathbf{x}) d\xi - \varepsilon_{Kl}^*(\mathbf{x}) \}. \quad (2.5.27)$$

The advantage of the method of Fourier integrals is obvious. Instead of solving the differential equations in Eq. (2.5.15), we can treat the algebraic equations in the Fourier space described by the wave vector ξ given in Eq. (2.5.20). As shown in the above discussion, the solution in the Fourier space is easily obtained. There are two challenges when using the method of Fourier integrals. One is how to obtain $\bar{\varepsilon}_{Mn}^*(\xi)$ for any given distribution of eigenstrains $\varepsilon_{Kl}^*(\mathbf{x})$. The other is how to perform the inverse Fourier transforms given in Eqs. (2.5.25)–(2.5.27). As a matter of fact, it is often not possible to obtain the explicit analytical expression for the inverse Fourier transforms.

2.5.3 Method of Green's Function

By considering Eq. (2.5.5), for a given distribution of $\varepsilon_{Kl}^*(\mathbf{x})$, the right hand side can be simplified to X_j , which is similar to the body force term f_j in Eq. (2.1.26). Thus, in Eq. (2.5.5), the generalized displacement field u_j is obtained as the

disturbance field caused by the existence of $\varepsilon_{kl}^*(\mathbf{x})$. And the effect of $\varepsilon_{kl}^*(\mathbf{x})$ on the equilibrium equation acts like an imaginary generalized body force X_J (X_4 denotes extraneous body charge density), where $X_J = C_{ijkl}\varepsilon_{kl,i}^*$ is distributed in the subdomain Ω . At this point, it is natural to think that if the generalized displacement field caused by a generalized unit point body force is known, the integration of this basic solution in the subdomain Ω with varying amplitude $X_J = C_{ijkl}\varepsilon_{kl,i}^*$ for different points yields the final disturbance field caused by the existence of $\varepsilon_{kl}^*(\mathbf{x})$. Of course, the validity of this solution is guaranteed by the linearity of the equations in Eq. (2.5.5). When a generalized unit point body force is applied to the point $\mathbf{x}'$ in the x_J direction (for $J = 4$, a unit extraneous body charge density is defined in $\mathbf{x}'$), it causes a generalized displacement at point $\mathbf{x}$ in the x_K direction (for $K = 4$, the electric potential is considered). Let us denote this generalized displacement component by $G_{KJ}(\mathbf{x} - \mathbf{x}')$. We have

$$C_{ijkl}G_{KM,li}(\mathbf{x} - \mathbf{x}') = -\delta_{JM}\delta(\mathbf{x} - \mathbf{x}'), \quad (2.5.28)$$

where $\delta(\mathbf{x} - \mathbf{x}')$ is the three-dimensional Dirac delta function. Here $G_{KJ}(\mathbf{x} - \mathbf{x}')$ is called the Green's function. The solution for Eq. (2.5.5) is then given by

$$u_K(\mathbf{x}) = - \int_{-\infty}^{\infty} G_{KJ}(\mathbf{x} - \mathbf{x}') C_{iJmn} \varepsilon_{mn,i}^*(\mathbf{x}') d\mathbf{x}'. \quad (2.5.29)$$

Integrating by parts, Eq. (2.5.29) is transformed as

$$\begin{aligned} u_K(\mathbf{x}) &= \int_{-\infty}^{\infty} C_{iJmn} \frac{\partial}{\partial x'_i} G_{KJ}(\mathbf{x} - \mathbf{x}') \varepsilon_{mn}^*(\mathbf{x}') d\mathbf{x}' \\ &\quad - \int_{-\infty}^{\infty} C_{iJmn} \frac{\partial}{\partial x'_i} [G_{KJ}(\mathbf{x} - \mathbf{x}') \varepsilon_{mn}^*(\mathbf{x}')] d\mathbf{x}'. \end{aligned} \quad (2.5.30)$$

By assuming that $\varepsilon_{mn}^*(\mathbf{x}')$ takes zero value at infinity, the second term in Eq. (2.5.30) vanishes. Since $\frac{\partial}{\partial x'_i} G_{KJ}(\mathbf{x} - \mathbf{x}') = -G_{KJ,i}(\mathbf{x} - \mathbf{x}')$, Eq. (2.5.30) can be further simplified as

$$u_K(\mathbf{x}) = - \int_{-\infty}^{\infty} C_{iJmn} G_{KJ,i}(\mathbf{x} - \mathbf{x}') \varepsilon_{mn}^*(\mathbf{x}') d\mathbf{x}'. \quad (2.5.31)$$

The corresponding expressions for the generalized strain and stress are

$$\varepsilon_{KI}(\mathbf{x}) = \begin{cases} -\frac{1}{2} \int_{-\infty}^{\infty} C_{iJMn} \varepsilon_{Mn}^*(\mathbf{x}') [G_{KJ,li}(\mathbf{x}-\mathbf{x}') + G_{Ki,lJ}(\mathbf{x}-\mathbf{x}')] d\mathbf{x}', & K = 1, 2, 3 \\ -\int_{-\infty}^{\infty} C_{iJMn} \varepsilon_{Mn}^*(\mathbf{x}') G_{KJ,li}(\mathbf{x}-\mathbf{x}') d\mathbf{x}', & K = 4 \end{cases} \quad (2.5.32)$$

and

$$\sigma_{iJ}(\mathbf{x}) = -C_{iJKl} \left\{ \int_{-\infty}^{\infty} C_{pQMn} \varepsilon_{Mn}^*(\mathbf{x}') G_{KQ,lp}(\mathbf{x}-\mathbf{x}') d\mathbf{x}' - \varepsilon_{KI}^*(\mathbf{x}) \right\}. \quad (2.5.33)$$

Eqs. (2.5.31)–(2.5.33) provide the solution of Eq. (2.5.5) by using the method of Green's function. The solution for Green's function is itself a very important and skillful problem. Green's functions for various specific conditions are solved in Section 2.5.4. Yet, as can be seen in Eqs. (2.5.31)–(2.5.33), while using the method of Green's function, the actual terms entering the expressions are the derivatives of Green's function rather than Green's function itself.

2.6 Basic Solution for Three-Dimensional Problems

The three methods introduced in Section 2.5 are not independent of each other. The equivalent inclusion method synthesizes all possible micromechanical problems for piezoelectric solids with heterostructure to the “equivalent inclusion” problem. It is more like a conceptual breakthrough than a specific mathematical tool. In Section 2.6, the usage of these methods is shown in the basic solutions.

In Section 2.6.1, the basic solution for electroelastic fields with an infinite piezoelectric matrix containing an ellipsoidal inclusion is presented. Wang (1992a) derived the analytical solutions for the coupled electric and elastic field of this problem. While the versatility of an ellipsoidal shape—which can be extended to fibers, spheres, discs, etc.—is evident, the solution has turned out to have as far-reaching implications in the development of this discipline in science as the preceding legendary counterpart solution in elasticity given by Eshelby (1957).

In Section 2.6.2, the basic solution for electroelastic fields with an infinite piezoelectric matrix containing a penny shaped crack is presented. This solution was obtained by Wang (1992b).

In Section 2.6.3, the elastic problem of an ellipsoidal inhomogeneity embedded in an infinite matrix is treated, where both phases possess eigenstrains. This problem is associated with a thermodynamic phenomenology treating ferroelectric composites introduced in Chapter 3.

In Section 2.6.4, the three-dimensional Green's function for piezoelectric materials is derived. As introduced in Section 2.5.3, the Green's function serves as the basic solution in solving micromechanical problems of heterostructures. Wang and Achenbach (1995) solved the three-dimensional time-harmonic Green's

function for anisotropic elastic materials using the mathematical technique of Radon transform (Helgason, 1980). Later, the analytical expression for the static Green's function was obtained by Wang (1997), using the same method. The three-dimensional Green's function in anisotropic piezoelectric solids was first obtained by Pan and Tonon (2000), by using the Radon transform. Li and Wang (2007) provided an alternative method based on the Fourier transform and complex variables. Here we deduce the solution for the three-dimensional static Green's function based on Pan and Tonon (2000).

As explained at the beginning of Chapter 2, the methods introduced here aim at solving piezoelectric materials with general anisotropy.

2.6.1 Ellipsoidal Inhomogeneous Inclusions

By extending Eshelby's approach (Eshelby, 1957) for elastic materials, Wang (1992a) gave the three-dimensional solution on the coupled elastic and electric fields of an anisotropic piezoelectric ellipsoidal inhomogeneity embedded in an infinite matrix. In Wang's work, the techniques of Green's function and Fourier integrals were combined. Exact analytical expressions were obtained for interior fields of the inhomogeneity. The exterior fields just outside the inhomogeneity were also expressed in explicit forms. Here, we extend Wang's solution to the case of an ellipsoidal piezoelectric inhomogeneous inclusion embedded in an infinite matrix.

The governing equation for inhomogeneous inclusions is obtained by using the equivalent inclusion method in Eq. (2.5.14) as

$$C_{ijkl}u_{K,i|l} = C_{ijkl}(\varepsilon_{Kl,i}^* + \varepsilon_{Kl,i}^{**}), \quad (2.6.1)$$

where ε_{Kl}^{**} is solved from

$$\Delta C_{ijkl}(u_{K,i|l} - \varepsilon_{Kl,i}^*) = C_{ijkl}\varepsilon_{Kl,i}^{**}, \quad (2.6.2)$$

and

$$\Delta C_{ijkl} = (C_{ijkl} - C_{ijkl}^*)H(\mathbf{x}), \quad (2.6.3)$$

$$H(\mathbf{x}) = \begin{cases} 1, & \mathbf{x} \in \Omega \\ 0, & \mathbf{x} \notin \Omega \end{cases}. \quad (2.6.4)$$

C_{ijkl} and C_{ijkl}^* denote components of the generalized stiffness tensor for the matrix and the inhomogeneous inclusion. Ω denotes the subdomain occupied by the ellipsoidal inhomogeneous inclusion. The Cartesian coordinates $\mathbf{x} = (x_1, x_2, x_3)$ are set at the directions of the three principal axes of the ellipsoid, with length

denoted by a , b and c .

The solution for Eq. (2.6.1) is obtained by using the method of Green's function as

$$u_K(\mathbf{x}) = u_K^\infty - \int_{-\infty}^{\infty} C_{iJm} G_{KJ,i}(\mathbf{x} - \mathbf{x}') [\varepsilon_{m}^*(\mathbf{x}') + \varepsilon_{m}^{**}(\mathbf{x}')] d\mathbf{x}', \quad (2.6.5)$$

where u_K^∞ denotes the generalized displacement fields caused by generalized constant strain fields ε_{KL}^∞ (or generalized constant stress fields σ_{iJ}^∞) applied at infinity. The corresponding expressions for the generalized strains and stresses are

$$\varepsilon_{KL}(\mathbf{x}) = \begin{cases} \varepsilon_{KL}^\infty - \frac{1}{2} \int_{-\infty}^{\infty} C_{iJm} [\varepsilon_m^*(\mathbf{x}') + \varepsilon_m^{**}(\mathbf{x}')] [G_{KJ,il}(\mathbf{x} - \mathbf{x}') + G_{iJ,Kl}(\mathbf{x} - \mathbf{x}')] d\mathbf{x}', & K = 1, 2, 3 \\ \varepsilon_{KL}^\infty - \int_{-\infty}^{\infty} C_{iJm} [\varepsilon_m^*(\mathbf{x}') + \varepsilon_m^{**}(\mathbf{x}')] G_{KJ,il}(\mathbf{x} - \mathbf{x}') d\mathbf{x}', & K = 4 \end{cases} \quad (2.6.6)$$

and

$$\sigma_{iJ}(\mathbf{x}) = \sigma_{iJ}^\infty - C_{iJKl} \left\{ \int_{-\infty}^{\infty} C_{pQM} [\varepsilon_M^*(\mathbf{x}') + \varepsilon_M^{**}(\mathbf{x}')] G_{KQ,lp}(\mathbf{x} - \mathbf{x}') d\mathbf{x}' - [\varepsilon_{KL,i}^*(\mathbf{x}') + \varepsilon_{KL,i}^{**}(\mathbf{x}')] \right\}. \quad (2.6.7)$$

The Green's function $G_{KJ}(\mathbf{x} - \mathbf{x}')$ can be solved from Eq. (2.5.28). Since the solution of Green's function is itself a basic inclusion problem, by using the method of Fourier integrals we have

$$G_{KJ}(\mathbf{x} - \mathbf{x}') = \frac{1}{8\pi^3} \int_{-\infty}^{\infty} \bar{G}_{KJ}(\boldsymbol{\xi}) \exp[i\boldsymbol{\xi} \cdot (\mathbf{x} - \mathbf{x}')] d\boldsymbol{\xi}, \quad (2.6.8)$$

where $\bar{G}_{KJ}(\boldsymbol{\xi})$ satisfies

$$C_{iJKl} \bar{G}_{KM}(\boldsymbol{\xi}) \xi_i \xi_j = -\delta_{JM}. \quad (2.6.9)$$

Eq. (2.6.9) gives

$$\bar{G}_{KJ}(\boldsymbol{\xi}) = [D^{-1}]_{KJ}, \quad (2.6.10)$$

where $[D^{-1}]_{KJ}$ is defined in Eq. (2.5.22). The first- and second-order derivatives of Green's function are given by

$$G_{KJ,l}(\mathbf{x}-\mathbf{x}') = \frac{i}{8\pi^3} \int_{-\infty}^{\infty} \bar{G}_{KJ}(\xi) \xi_l \exp[i\xi \cdot (\mathbf{x}-\mathbf{x}')] d\xi, \quad (2.6.11)$$

$$G_{KJ,li}(\mathbf{x}-\mathbf{x}') = -\frac{1}{8\pi^3} \int_{-\infty}^{\infty} \bar{G}_{KJ}(\xi) \xi_l \xi_i \exp[i\xi \cdot (\mathbf{x}-\mathbf{x}')] d\xi. \quad (2.6.12)$$

Substitution of Eq. (2.6.12) into Eq. (2.6.6) gives

$$\begin{aligned} \mathcal{E}_{Kl}(\mathbf{x}) = & \mathcal{E}_{Kl}^{\infty} + \\ & \begin{cases} \frac{1}{16\pi^3} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} C_{iJMn} [\mathcal{E}_{Mn}^*(\mathbf{x}') + \mathcal{E}_{Mn}^{**}(\mathbf{x}')] [\bar{G}_{KJ}(\xi) \xi_l \xi_i + \bar{G}_{lJ}(\xi) \xi_K \xi_i] \exp[i\xi \cdot (\mathbf{x}-\mathbf{x}')] d\xi d\mathbf{x}', & K=1, 2, 3 \\ \frac{1}{8\pi^3} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} C_{iJMn} [\mathcal{E}_{Mn}^*(\mathbf{x}') + \mathcal{E}_{Mn}^{**}(\mathbf{x}')] \bar{G}_{KJ}(\xi) \xi_l \xi_i \exp[i\xi \cdot (\mathbf{x}-\mathbf{x}')] d\xi d\mathbf{x}', & K=4 \end{cases} \end{aligned} \quad (2.6.13)$$

Due to their similar structure, the core integral of Eq. (2.6.13) is given by

$$I_{Kl} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} C_{iJMn} [\mathcal{E}_{Mn}^*(\mathbf{x}') + \mathcal{E}_{Mn}^{**}(\mathbf{x}')] \bar{G}_{KJ}(\xi) \xi_l \xi_i \exp[i\xi \cdot (\mathbf{x}-\mathbf{x}')] d\xi d\mathbf{x}'. \quad (2.6.14)$$

The integration with respect to the ξ -space (the Fourier space) is considered first. According to Eq. (2.5.17), we have

$$d\xi = d\xi_1 d\xi_2 d\xi_3 = \xi^2 d\xi dS(\mathbf{w}), \quad (2.6.15)$$

where

$$\xi = (\xi_1^2 + \xi_2^2 + \xi_3^2)^{1/2}, \quad \mathbf{w} = \xi / \xi. \quad (2.6.16)$$

And $dS(\mathbf{w})$ is a surface element on the unit sphere S^2 in the ξ -space. Eq. (2.6.14) is then rewritten as

$$\begin{aligned} I_{Kl} = & \int_{-\infty}^{\infty} d\mathbf{x}' \int_0^{\infty} \xi^2 d\xi \int_{S^2} C_{iJMn} [\mathcal{E}_{Mn}^*(\mathbf{x}') + \mathcal{E}_{Mn}^{**}(\mathbf{x}')] \bar{G}_{KJ}(\xi) \xi_l \xi_i \exp[i\xi \mathbf{w} \cdot (\mathbf{x}-\mathbf{x}')] dS(\mathbf{w}) \\ = & \frac{1}{2} \int_{\Omega} C_{iJMn} [\mathcal{E}_{Mn}^*(\mathbf{x}') + \mathcal{E}_{Mn}^{**}(\mathbf{x}')] d\mathbf{x}' \int_{S^2} \bar{G}_{KJ}(\xi) \xi_l \xi_i dS(\mathbf{w}) \int_0^{\infty} \xi^2 \exp[i\xi \mathbf{w} \cdot (\mathbf{x}-\mathbf{x}')] d\xi \\ = & -\pi \int_{\Omega} C_{iJMn} [\mathcal{E}_{Mn}^*(\mathbf{x}') + \mathcal{E}_{Mn}^{**}(\mathbf{x}')] d\mathbf{x}' \int_{S^2} \bar{G}_{KJ}(\mathbf{w}) w_l w_i \delta''[\mathbf{w} \cdot (\mathbf{x}-\mathbf{x}')] dS(\mathbf{w}). \end{aligned} \quad (2.6.17)$$

In deduction of Eq. (2.6.17), the following properties

$$\xi^2 \bar{G}_{KM}(\xi) = \bar{G}_{KM}(\mathbf{w}), \quad \int_{-\infty}^{\infty} \exp(i\xi \eta) d\xi = 2\pi \delta(\eta) \quad (2.6.18)$$

have been used.

The integration with respect to the subdomain Ω is then considered. The boundary of Ω is expressed by

$$\frac{(x'_1)^2}{a^2} + \frac{(x'_2)^2}{b^2} + \frac{(x'_3)^2}{c^2} = 1. \quad (2.6.19)$$

The following transformations of variables are useful in treating the integration:

$$\begin{aligned} x_1 / a &= y_1, \quad x_2 / b = y_2, \quad x_3 / c = y_3, \quad x'_1 / a = y'_1, \quad x'_2 / b = y'_2, \\ x'_3 / c &= y'_3, \quad aw_1 = \rho_1, \quad bw_2 = \rho_2, \quad cw_3 = \rho_3, \\ aw_1 / \rho &= \bar{w}_1, \quad bw_2 / \rho = \bar{w}_2, \quad cw_3 / \rho = \bar{w}_3, \quad \rho = (\rho_1^2 + \rho_2^2 + \rho_3^2)^{1/2}. \end{aligned} \quad (2.6.20)$$

The volume element dx' is then transformed as

$$dx' = dx'_1 dx'_2 dx'_3 = abc dy'_1 dy'_2 dy'_3 = abcr dr d\theta dz, \quad (2.6.21)$$

where

$$z = \bar{\mathbf{w}} \cdot \mathbf{y}' = \mathbf{w} \cdot \mathbf{x}' / \rho. \quad (2.6.22)$$

Substitution of Eq. (2.6.21) and Eq. (2.6.20) into Eq. (2.6.17) gives

$$\begin{aligned} I_{KI} &= -\pi abc \int_{-1}^1 dz \int_0^{2\pi} d\phi \int_0^R r dr \int_{S^2} C_{iMn} [\varepsilon_{Mn}^*(\mathbf{x}') \\ &\quad + \varepsilon_{Mn}^{**}(\mathbf{x}')] \bar{G}_{KJ}(\mathbf{w}) w_i w_j \delta''(\rho \bar{w}_p y_p - \rho z) dS(\mathbf{w}), \end{aligned} \quad (2.6.23)$$

where $R = \sqrt{1-z^2}$ must be a real number. The argument of the second order derivative of Dirac's delta function becomes zero when

$$\bar{w}_p y_p = z. \quad (2.6.24)$$

From Eq. (2.6.23) we see that z is restricted to $(-1, 1)$, while $\mathbf{y}$ can be arbitrarily defined. Thus, for Eq. (2.6.24) to hold, effective values of the unit vector $\mathbf{w}$ are restricted by the selection of $\mathbf{y}$.

2.6.1.1 Formulae for Interior Points

Now let us consider the case when $\mathbf{y}$ is located inside the unit sphere, or we can say when point $\mathbf{x}$ is inside the ellipsoidal inclusion. Integration of Eq. (2.6.23) by

parts with respect to z leads to

$$\begin{aligned}
 I_{KI} = & -\pi abc \int_0^{2\pi} d\phi \left\{ \int_0^R r dr \frac{\partial^2}{\partial z^2} C_{iJMn} [\varepsilon_{Mn}^*(\mathbf{x}') + \varepsilon_{Mn}^{**}(\mathbf{x}')] - z \left\{ \frac{\partial}{\partial z} C_{iJMn} [\varepsilon_{Mn}^*(\mathbf{x}') + \varepsilon_{Mn}^{**}(\mathbf{x}')] \right\} \right\}_{r=R} \\
 & - z \frac{\partial}{\partial z} \left\{ C_{iJMn} [\varepsilon_{Mn}^*(\mathbf{x}') + \varepsilon_{Mn}^{**}(\mathbf{x}')] \right\}_{r=R} - \left\{ C_{iJMn} [\varepsilon_{Mn}^*(\mathbf{x}') + \varepsilon_{Mn}^{**}(\mathbf{x}')] \right\}_{r=R} \right\}_{z=\bar{w}_p, y_p} \\
 & \times \int_{S^2} \bar{G}_{KJ}(\mathbf{w}) \rho^{-3} w_i w_j dS(\mathbf{w}).
 \end{aligned} \tag{2.6.25}$$

When the eigenstrains $\varepsilon_{Mn}^*(\mathbf{x}')$ and $\varepsilon_{Mn}^{**}(\mathbf{x}')$ are constants inside the ellipsoidal, Eq. (2.6.25) is transformed to

$$I_{KI} = 2\pi^2 C_{iJMn} (\varepsilon_{Mn}^* + \varepsilon_{Mn}^{**}) \int_{S^2} \bar{G}_{KJ}(\mathbf{w}) w_i w_j dS(\bar{\mathbf{w}}), \tag{2.6.26}$$

where the relation

$$dS(\bar{\mathbf{w}}) = abc \rho^{-3} dS(\mathbf{w}) \tag{2.6.27}$$

is used. Eq. (2.6.26) can be transformed to

$$I_{KI} = 2\pi^2 C_{iJMn} (\varepsilon_{Mn}^* + \varepsilon_{Mn}^{**}) \int_{-1}^1 d\bar{w}_3 \int_0^{2\pi} \bar{G}_{KJ}(\mathbf{w}) w_i w_j d\theta, \tag{2.6.28}$$

where the relation between w_i and $\bar{w}_i$ is given in Eq. (2.6.20), and

$$\bar{w}_1 = (1 - \bar{w}_3^2)^{1/2} \cos \theta, \quad \bar{w}_2 = (1 - \bar{w}_3^2)^{1/2} \sin \theta. \tag{2.6.29}$$

The integration in Eq. (2.6.28) is a constant which is independent of the location of $\mathbf{x}$. Thus, when the eigenstrains $\varepsilon_{Mn}^*(\mathbf{x}')$ and $\varepsilon_{Mn}^{**}(\mathbf{x}')$ are constants inside the ellipsoidal, the total strains are also constants. In this case, the lengthy expression in Eq. (2.6.13) can be simplified to a very concise one:

$$\varepsilon_{KI} = \varepsilon_{KI}^\infty + S_{IKMn} (\varepsilon_{Mn}^* + \varepsilon_{Mn}^{**}), \tag{2.6.30}$$

where

$$S_{IKMn} = \begin{cases} \frac{1}{8\pi} C_{iJMn} (N_{KJli} + N_{lJKi}), & K = 1, 2, 3 \\ \frac{1}{4\pi} C_{iJMn} N_{KJli}, & K = 4 \end{cases} \tag{2.6.31}$$

and

$$N_{KJli} = \int_{-1}^1 d\bar{w}_3 \int_0^{2\pi} \bar{G}_{KJ}(\mathbf{w}) w_i w_l d\theta. \quad (2.6.32)$$

The fourth-order tensor S_{IKMn} can be called the Eshelby tensor for anisotropic piezoelectric materials. Although the exact expressions of the integral in Eq. (2.6.32) are too lengthy to be presented, it is actually rather convenient for numerical calculation of Eshelby tensor for a general ellipsoidal shaped inclusion.

When the Eshelby tensor is known, the total strain field can be solved from Eq. (2.6.30). Now we obtain the corresponding solutions for inclusions, inhomogeneities and inhomogeneous inclusions.

For *inclusions* with prescribed eigenstrains ε_{Mn}^* , the generalized strains for interior points are given by

$$\varepsilon_{KI} = \varepsilon_{KI}^\infty + S_{IKMn} \varepsilon_{Mn}^*. \quad (2.6.33)$$

The associated expression for the generalized stresses is given by

$$\sigma_{iJ} = C_{ijKI} (\varepsilon_{KI}^\infty + S_{IKMn} \varepsilon_{Mn}^* - \varepsilon_{KI}^*). \quad (2.6.34)$$

For *inhomogeneities* governed by Eq. (2.5.10), the solutions for the total strains for interior points are provided in matrix notation as

$$\boldsymbol{\varepsilon} = (\mathbf{I} - \mathbf{S}\mathbf{C}^{-1}\Delta\mathbf{C})^{-1} \boldsymbol{\varepsilon}^\infty, \quad (2.6.35)$$

where

$$\mathbf{C} = \begin{bmatrix} C_{11} & C_{12} & C_{13} & C_{14} & C_{15} & C_{16} & e_{11} & e_{21} & e_{31} \\ C_{12} & C_{22} & C_{23} & C_{24} & C_{25} & C_{26} & e_{12} & e_{22} & e_{32} \\ C_{13} & C_{23} & C_{33} & C_{34} & C_{35} & C_{36} & e_{13} & e_{23} & e_{33} \\ C_{14} & C_{24} & C_{34} & C_{44} & C_{45} & C_{46} & e_{14} & e_{24} & e_{34} \\ C_{15} & C_{25} & C_{35} & C_{45} & C_{55} & C_{56} & e_{15} & e_{25} & e_{35} \\ C_{16} & C_{26} & C_{36} & C_{46} & C_{56} & C_{66} & e_{16} & e_{26} & e_{36} \\ e_{11} & e_{12} & e_{13} & e_{14} & e_{15} & e_{16} & \alpha_{11} & \alpha_{12} & \alpha_{13} \\ e_{21} & e_{22} & e_{23} & e_{24} & e_{25} & e_{26} & \alpha_{12} & \alpha_{22} & \alpha_{23} \\ e_{31} & e_{32} & e_{33} & e_{34} & e_{35} & e_{36} & \alpha_{13} & \alpha_{23} & \alpha_{33} \end{bmatrix}, \quad (2.6.36)$$

and

$$\mathbf{S} = \begin{bmatrix} S_{1111} & S_{1122} & S_{1133} & S_{1123} & S_{1113} & S_{1112} & S_{1141} & S_{1142} & S_{1143} \\ S_{2211} & S_{2222} & S_{2233} & S_{2223} & S_{2213} & S_{2212} & S_{2241} & S_{2242} & S_{2243} \\ S_{3311} & S_{3322} & S_{3333} & S_{3323} & S_{3313} & S_{3312} & S_{3341} & S_{3342} & S_{3343} \\ 2S_{2311} & 2S_{2322} & 2S_{2333} & 2S_{2323} & 2S_{2313} & 2S_{2312} & 2S_{2341} & 2S_{2342} & 2S_{2343} \\ 2S_{1311} & 2S_{1322} & 2S_{1333} & 2S_{1323} & 2S_{1313} & 2S_{1312} & 2S_{1341} & 2S_{1342} & 2S_{1343} \\ 2S_{1211} & 2S_{1222} & 2S_{1233} & 2S_{1223} & 2S_{1213} & 2S_{1212} & 2S_{1241} & 2S_{1242} & 2S_{1243} \\ S_{1411} & S_{1422} & S_{1433} & S_{1423} & S_{1413} & S_{1412} & S_{1441} & S_{1442} & S_{1443} \\ S_{2411} & S_{2422} & S_{2433} & S_{2423} & S_{2413} & S_{2412} & S_{2441} & S_{2442} & S_{2443} \\ S_{3411} & S_{3422} & S_{3433} & S_{3423} & S_{3413} & S_{3412} & S_{3441} & S_{3442} & S_{3443} \end{bmatrix}. \quad (2.6.37)$$

$\Delta \mathbf{C}$ has the same structure as Eq. (2.6.36), whose components are defined in Eq. (2.6.3), and $\mathbf{C}^{-1}$ denotes the inverse matrix of $\mathbf{C}$. $\mathbf{I}$ is the ninth-order unit matrix.

The associated expression for the generalized stresses written in matrix notation is given by

$$\boldsymbol{\sigma} = \mathbf{C}^* (\mathbf{I} - \mathbf{S} \mathbf{C}^{-1} \Delta \mathbf{C})^{-1} \boldsymbol{\varepsilon}^\infty, \quad (2.6.38)$$

where

$$\boldsymbol{\sigma} = [\sigma_{11}, \sigma_{22}, \sigma_{33}, \sigma_{23}, \sigma_{13}, \sigma_{12}, D_1, D_2, D_3]^T. \quad (2.6.39)$$

For *inhomogeneous inclusions* governed by Eq. (2.5.12), the solution for the total strains is provided in matrix notation as

$$\boldsymbol{\varepsilon} = (\mathbf{I} - \mathbf{S} \mathbf{C}^{-1} \Delta \mathbf{C})^{-1} [\boldsymbol{\varepsilon}^\infty + \mathbf{S} (\mathbf{I} - \mathbf{C}^{-1} \Delta \mathbf{C}) \boldsymbol{\varepsilon}^*]. \quad (2.6.40)$$

The associated expression for the generalized stresses written in matrix notation is given by

$$\boldsymbol{\sigma} = \mathbf{C}^* \{ (\mathbf{I} - \mathbf{S} \mathbf{C}^{-1} \Delta \mathbf{C})^{-1} [\boldsymbol{\varepsilon}^\infty + \mathbf{S} (\mathbf{I} - \mathbf{C}^{-1} \Delta \mathbf{C}) \boldsymbol{\varepsilon}^*] - \boldsymbol{\varepsilon}^* \}. \quad (2.6.41)$$

Eqs. (2.6.33)–(2.6.41) are obtained without specifying the expression of Eshelby tensor.

For transversely isotropic piezoelectric materials and for several specific shapes of inclusion, explicit expressions of the integral in Eq. (2.6.32) can be obtained.

• Cylindrical inclusions

When both the inclusion and the matrix are assumed to be transversely isotropic, non-zero components of the fourth-order tensor $N_{K,li}$ are

$$\begin{aligned}
N_{1111} &= N_{2222} = \frac{5C_{11} - 3C_{12}}{2C_{11}(C_{11} - C_{12})} \pi, \\
N_{1122} &= N_{2211} = \frac{7C_{11} - C_{12}}{2C_{11}(C_{11} - C_{12})} \pi, \\
N_{3311} &= N_{3322} = \frac{2\pi\alpha_{11}}{C_{44}\alpha_{11} + e_{15}^2}, \\
N_{1212} &= N_{2121} = N_{2112} = N_{1221} = -\frac{C_{11} + C_{12}}{2C_{11}(C_{11} - C_{12})} \pi, \\
N_{3411} &= N_{3422} = N_{4311} = N_{4322} = \frac{2\pi e_{15}}{C_{44}\alpha_{11} + e_{15}^2}, \\
N_{4411} &= N_{4422} = -\frac{2C_{44}\pi}{C_{44}\alpha_{11} + e_{15}^2}.
\end{aligned} \tag{2.6.42}$$

• *Spheroidal inclusions*

For spheroidal inclusions ($a = b$, $a/c = \beta$, β is called the aspect ratio), N_{Kjli} can be expressed as an explicit integral:

$$\begin{aligned}
N_{1111} &= N_{2222} = \frac{\pi}{2} \beta^2 \int_0^1 (1-w^2) M \{ \{ [\beta^2(1-w^2)C_{44} + C_{33}w^2] \\
&\quad \cdot [\beta^2(1-w^2)(C_{11} + 3C_{66}) + 4C_{44}w^2] - \beta^2w^2(1-w^2)(C_{13} + C_{44})^2 \} \\
&\quad \cdot [\beta^2(1-w^2)\alpha_{11} + \alpha_{33}w^2] + [\beta^2(1-w^2)(C_{11} + 3C_{66}) + 4C_{44}w^2] \\
&\quad \cdot [\beta^2(1-w^2)e_{15} + e_{33}w^2]^2 - \beta^2w^2(1-w^2)(e_{31} + e_{15}) \\
&\quad \cdot [\beta^2(1-w^2)(2e_{15}C_{13} + e_{15}C_{44} - e_{31}C_{44}) \\
&\quad + w^2(2e_{15}C_{13} + e_{15}C_{44} - e_{31}C_{44} - e_{15}C_{33})] \} dw,
\end{aligned} \tag{2.6.43a}$$

$$\begin{aligned}
N_{3333} &= 4\pi \int_0^1 w^2 M \{ [\beta^2(1-w^2)C_{44} + C_{33}w^2][\beta^2(1-w^2)C_{11} + C_{33}w^2] \\
&\quad \cdot [\beta^2(1-w^2)\alpha_{11} + \alpha_{33}w^2] + \beta^2w^2(1-w^2)(e_{31} + e_{15})^2 \\
&\quad \cdot [\beta^2(1-w^2)C_{66} + C_{44}w^2] \} dw,
\end{aligned} \tag{2.6.43b}$$

$$\begin{aligned}
N_{1122} &= N_{2211} = \frac{\pi}{2} \beta^2 \int_0^1 (1-w^2) M \{ \{ [\beta^2(1-w^2)(3C_{11} + C_{66}) + 4C_{44}w^2] \\
&\quad \cdot [\beta^2(1-w^2)C_{44} + C_{33}w^2] - 3\beta^2w^2(1-w^2)(C_{13} + C_{44})^2 \} \\
&\quad \cdot [\beta^2(1-w^2)\alpha_{11} + \alpha_{33}w^2] - 3\beta^2w^2(1-w^2)(e_{31} + e_{15}) \\
&\quad \cdot [\beta^2(1-w^2)(2e_{15}C_{13} + e_{15}C_{44} - e_{31}C_{44}) \\
&\quad + w^2(2e_{33}C_{13} + 2e_{33}C_{44} - e_{31}C_{33} - e_{15}C_{33})] \\
&\quad + [\beta^2(1-w^2)(3C_{11} + C_{66}) \\
&\quad + 4C_{44}w^2][\beta^2(1-w^2)e_{15} + e_{33}w^2]^2 \} dw,
\end{aligned} \tag{2.6.43c}$$

$$\begin{aligned}
N_{1133} = N_{2233} = 2\pi \int_0^1 w^2 M \{ & [\beta^2(1-w^2)(C_{11} + C_{66}) + 2C_{44}w^2] \\
& \cdot [\beta^2(1-w^2)C_{44} + C_{33}w^2] - \beta^2w^2(1-w^2)(C_{13} + C_{44})^2 \} \\
& \cdot [\beta^2(1-w^2)\alpha_{11} + \alpha_{33}w^2] - \beta^2w^2(1-w^2)(e_{31} + e_{15}) \\
& \cdot [\beta^2(1-w^2)(2e_{15}C_{13} + e_{15}C_{44} - e_{31}C_{44}) \\
& + w^2(2e_{33}C_{13} + 2e_{33}C_{44} - e_{31}C_{33} - e_{15}C_{33})] \\
& + [\beta^2(1-w^2)(C_{11} + C_{66}) \\
& + 2C_{44}w^2][\beta^2(1-w^2)e_{15} + e_{33}w^2]^2 \} dw,
\end{aligned} \tag{2.6.43d}$$

$$\begin{aligned}
N_{3311} = N_{3322} = 2\pi\beta^2 \int_0^1 (1-w^2) M \{ & [\beta^2(1-w^2)C_{66} + C_{44}w^2] \\
& \cdot [\beta^2(1-w^2)C_{11} + C_{44}w^2][\beta^2(1-w^2)\alpha_{11} + \alpha_{33}w^2] \\
& + \beta^2w^2(1-w^2)(e_{31} + e_{15})^2[\beta^2(1-w^2)C_{66} + C_{44}w^2] \} dw,
\end{aligned} \tag{2.6.43e}$$

$$\begin{aligned}
N_{1212} = \frac{\pi}{2} \beta^4 \int_0^1 (1-w^2) M \{ & [\beta^2(1-w^2)(2e_{15}C_{13} + e_{15}C_{44} - e_{31}C_{44}) \\
& + w^2(2e_{33}C_{13} + 2e_{33}C_{44} - e_{31}C_{33} - e_{15}C_{33})]w^2(e_{15} + e_{31}) \\
& + \{w^2(C_{13} + C_{44})^2 - (C_{11} - C_{66})[\beta^2(1-w^2)C_{44} + C_{33}w^2]\} \\
& \cdot [\beta^2(1-w^2)\alpha_{11} + \alpha_{33}w^2] \\
& - (C_{11} - C_{66})[\beta^2(1-w^2)e_{15} + e_{33}w^2]^2 \} dw,
\end{aligned} \tag{2.6.43f}$$

$$\begin{aligned}
N_{1313} = N_{2323} = -2\pi\beta^2 \int_0^1 w(1-w^2) M [& \beta^2(1-w^2)C_{66} + C_{44}w^2] \\
& \cdot \{ (C_{13} + C_{44})[\beta^2(1-w^2)\alpha_{11} + \alpha_{33}w^2] + (e_{31} + e_{15}) \\
& \cdot [\beta^2(1-w^2)e_{15} + e_{33}w^2] \} dw,
\end{aligned} \tag{2.6.43g}$$

$$\begin{aligned}
N_{1413} = 2\pi\beta^2 \int_0^1 w(1-w^2) M [& \beta^2(1-w^2)C_{66} + C_{44}w^2] \\
& \cdot [\beta^2(1-w^2)(e_{31}C_{44} - e_{15}C_{13}) \\
& + w^2(e_{31}C_{33} + e_{15}C_{33} - e_{31}C_{33} - e_{33}C_{44})] dw,
\end{aligned} \tag{2.6.43h}$$

$$\begin{aligned}
N_{3411} = 2\pi\beta^2 \int_0^1 w(1-w^2) M \{ & [\beta^2(1-w^2)C_{66} + C_{44}w^2] \\
& \cdot [\beta^2(1-w^2)C_{11} + w^2C_{44}][\beta^2(1-w^2)e_{15} + e_{33}w^2] \\
& - \beta^2w^2(1-w^2)(e_{31} + e_{15})(C_{13} + C_{44}) \\
& \cdot [\beta^2(1-w^2)C_{66} + C_{44}w^2] \} dw,
\end{aligned} \tag{2.6.43i}$$

$$\begin{aligned}
N_{3433} = 4\pi \int_0^1 w^2 M \{ & [\beta^2(1-w^2)C_{66} + C_{44}w^2] \\
& \cdot [\beta^2(1-w^2)C_{11} + w^2C_{44}][\beta^2(1-w^2)e_{15} + e_{33}w^2] \\
& - \beta^2w^2(1-w^2)(e_{31} + e_{15})(C_{13} + C_{44}) \\
& \cdot [\beta^2(1-w^2)C_{66} + C_{44}w^2] \} dw,
\end{aligned} \tag{2.6.43j}$$

$$\begin{aligned}
N_{4411} = 2\pi\beta^2 \int_0^1 w(1-w^2) M \{ & \beta^2 w^2 (1-w^2) \\
& \cdot [\beta^2 (1-w^2)C_{66} + C_{44}w^2](C_{13} + C_{44})^2 \\
& - [\beta^2 (1-w^2)C_{66} + w^2C_{44}][\beta^2 (1-w^2)C_{11} + w^2C_{44}] \\
& \cdot [\beta^2 (1-w^2)C_{44} + w^2C_{33}] \} dw,
\end{aligned} \tag{2.6.43k}$$

$$\begin{aligned}
N_{4433} = 4\pi \int_0^1 w^2 M \{ & \beta^2 w^2 (1-w^2)[\beta^2 (1-w^2)C_{66} + C_{44}w^2](C_{13} + C_{44})^2 \\
& - [\beta^2 (1-w^2)C_{66} + w^2C_{44}][\beta^2 (1-w^2)C_{11} + w^2C_{44}] \\
& \cdot [\beta^2 (1-w^2)C_{44} + w^2C_{33}] \} dw,
\end{aligned} \tag{2.6.43l}$$

$$\begin{aligned}
M^{-1} = & [\beta^2 (1-w^2)C_{66} + C_{44}w^2] \{ [\beta^2 (1-w^2)\alpha_{11} + \alpha_{33}w^2] \\
& \cdot [\beta^2 (1-w^2)C_{11} + C_{44}w^2][\beta^2 (1-w^2)C_{44} + C_{33}w^2] \\
& - \beta^2 w^2 (1-w^2)(C_{13} + C_{44})^2 + [\beta^2 (1-w^2)C_{11} + C_{44}w^2] \\
& \cdot [\beta^2 (1-w^2)e_{15} + e_{33}w^2]^2 - 2\beta^2 w^2 (1-w^2)(C_{13} + C_{44})(e_{31} + e_{15}) \\
& \cdot [\beta^2 (1-w^2)e_{15} + e_{33}w^2] + 2\beta^2 w^2 (1-w^2)(e_{31} + e_{15})^2 \\
& \cdot [\beta^2 (1-w^2)C_{44} + C_{33}w^2] \}.
\end{aligned} \tag{2.6.43m}$$

Eq. (2.6.35) and Eq. (2.6.38) provide the general relation of the internal strains and stresses of the inhomogeneity to the applied strains at the boundary. For spheroidal inhomogeneity with transverse isotropy, Eq. (2.6.35) can be rewritten as (Wang and Du, 1993; Wang, 1994)

$$\boldsymbol{\varepsilon} = \mathbf{B}\boldsymbol{\varepsilon}^\infty, \tag{2.6.44}$$

where

$$\mathbf{B} = \begin{bmatrix} B_{11}^1 & B_{12}^1 & B_{13}^1 & 0 & 0 & 0 & 0 & 0 & B_{13}^2 \\ B_{12}^1 & B_{11}^1 & B_{13}^1 & 0 & 0 & 0 & 0 & 0 & B_{13}^2 \\ B_{31}^1 & B_{31}^1 & B_{33}^1 & 0 & 0 & 0 & 0 & 0 & B_{33}^2 \\ 0 & 0 & 0 & B_{44}^1 & 0 & 0 & 0 & 2B_{42}^2 & 0 \\ 0 & 0 & 0 & 0 & B_{44}^1 & 0 & 2B_{42}^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & B_{66}^1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1/2B_{15}^3 & 0 & B_{11}^4 & 0 & 0 \\ 0 & 0 & 0 & 1/2B_{15}^3 & 0 & 0 & 0 & B_{11}^4 & 0 \\ B_{31}^3 & B_{31}^3 & B_{33}^3 & 0 & 0 & 0 & 0 & 0 & B_{33}^4 \end{bmatrix}, \tag{2.6.45}$$

and

$$\begin{bmatrix} B_{11}^I & B_{12}^I & B_{13}^I \\ B_{12}^I & B_{11}^I & B_{13}^I \\ B_{31}^I & B_{31}^I & B_{33}^I \end{bmatrix} = \begin{bmatrix} R_{11} & R_{12} & R_{13} \\ R_{12} & R_{11} & R_{13} \\ R_{31} & R_{31} & R_{33} \end{bmatrix}^{-1}. \quad (2.6.46)$$

The components in Eq. (2.6.46) are given by

$$\begin{aligned} R_{11} = & 1 + \frac{1}{4\pi} (G_{11}^I \Delta C_{11} + G_{12}^I \Delta C_{12} + G_{13}^I \Delta C_{13} + G_{31}^2 \Delta e_{31}) \\ & - \frac{\Xi}{16\pi^2} (G_{11}^I \Delta e_{31} + G_{12}^I \Delta e_{31} + G_{13}^I \Delta e_{33} - G_{31}^2 \Delta \alpha_{33}) \\ & \times (G_{31}^2 \Delta C_{11} + G_{31}^2 \Delta C_{12} + G_{33}^2 \Delta C_{33} + G_{33}^3 \Delta e_{31}), \end{aligned} \quad (2.6.47)$$

$$\begin{aligned} R_{33} = & 1 + \frac{1}{4\pi} (2G_{13}^I \Delta C_{13} + G_{33}^I \Delta C_{33} + G_{33}^2 \Delta e_{33}) \\ & - \frac{\Xi}{16\pi^2} (2G_{13}^I \Delta C_{31} + G_{33}^I \Delta e_{33} - G_{33}^2 \Delta \alpha_{33}) \\ & \times (G_{31}^2 \Delta C_{13} + G_{31}^2 \Delta C_{13} + G_{33}^2 \Delta C_{33} + G_{33}^3 \Delta e_{33}), \end{aligned} \quad (2.6.48)$$

$$\begin{aligned} R_{12} = & \frac{1}{4\pi} (G_{11}^I \Delta C_{12} + G_{12}^I \Delta C_{11} + G_{13}^I \Delta C_{13} + G_{31}^2 \Delta e_{31}) \\ & - \frac{\Xi}{16\pi^2} (G_{11}^I \Delta e_{31} + G_{12}^I \Delta e_{31} + G_{13}^I \Delta e_{33} - G_{31}^2 \Delta \alpha_{33}) \\ & \times (G_{31}^2 \Delta C_{11} + G_{31}^2 \Delta C_{12} + G_{33}^2 \Delta C_{33} + G_{33}^3 \Delta e_{31}), \end{aligned} \quad (2.6.49)$$

$$\begin{aligned} R_{13} = & \frac{1}{4\pi} (G_{11}^I \Delta C_{13} + G_{12}^I \Delta C_{13} + G_{13}^I \Delta C_{33} + G_{31}^2 \Delta e_{33}) \\ & - \frac{\Xi}{16\pi^2} (G_{11}^I \Delta e_{31} + G_{12}^I \Delta e_{31} + G_{13}^I \Delta e_{33} - G_{31}^2 \Delta \alpha_{33}) \\ & \times (G_{31}^2 \Delta C_{13} + G_{31}^2 \Delta C_{13} + G_{33}^2 \Delta C_{33} + G_{33}^3 \Delta e_{33}), \end{aligned} \quad (2.6.50)$$

$$\begin{aligned} R_{31} = & \frac{1}{4\pi} (G_{13}^I \Delta C_{11} + G_{13}^I \Delta C_{12} + G_{33}^I \Delta C_{13} + G_{33}^2 \Delta e_{31}) \\ & - \frac{\Xi}{16\pi^2} (G_{13}^I \Delta e_{31} + G_{13}^I \Delta e_{31} + G_{33}^I \Delta e_{33} - G_{33}^2 \Delta \alpha_{33}) \\ & \times (G_{31}^2 \Delta C_{11} + G_{31}^2 \Delta C_{12} + G_{33}^2 \Delta C_{33} + G_{33}^3 \Delta e_{31}), \end{aligned} \quad (2.6.51)$$

where

$$\Xi = [1 + \frac{1}{4\pi} (2G_{31}^2 \Delta e_{31} + G_{33}^2 \Delta e_{33} - G_{33}^3 \Delta \alpha_{32})]^{-1}, \quad (2.6.52)$$

and ΔC_{ij} , Δe_{ij} and $\Delta \alpha_{ij}$ are obtained from the material matrix

$$\Delta \mathbf{C} = \mathbf{C}^* - \mathbf{C} = \begin{bmatrix} \Delta C_{11} & \Delta C_{12} & \Delta C_{13} & 0 & 0 & 0 & 0 & 0 & \Delta e_{31} \\ \Delta C_{12} & \Delta C_{11} & \Delta C_{13} & 0 & 0 & 0 & 0 & 0 & \Delta e_{31} \\ \Delta C_{13} & \Delta C_{13} & \Delta C_{33} & 0 & 0 & 0 & 0 & 0 & \Delta e_{33} \\ 0 & 0 & 0 & \Delta C_{44} & 0 & 0 & 0 & \Delta e_{15} & 0 \\ 0 & 0 & 0 & 0 & \Delta C_{44} & 0 & \Delta e_{15} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \Delta C_{66} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \Delta e_{15} & 0 & \Delta \alpha_{11} & 0 & 0 \\ 0 & 0 & 0 & \Delta e_{15} & 0 & 0 & 0 & \Delta \alpha_{11} & 0 \\ \Delta e_{31} & \Delta e_{31} & \Delta e_{33} & 0 & 0 & 0 & 0 & 0 & \Delta \alpha_{33} \end{bmatrix}. \quad (2.6.53)$$

The other components in Eq. (2.6.45) are given by

$$B_{44}^1 = [1 + \frac{1}{\pi} G_{44}^1 \Delta C_{44} + \frac{1}{2\pi} G_{24}^2 \Delta e_{24} - \frac{1}{2\pi^2} (G_{44}^1 \Delta e_{24} - \frac{1}{2} G_{24}^2 \Delta \alpha_{11}) \\ \times (1 + \frac{1}{2\pi} G_{24}^2 \Delta e_{24} - \frac{1}{4\pi} G_{11}^3 \Delta \alpha_{11})^{-1} (G_{24}^2 \Delta C_{44} + \frac{1}{2} G_{11}^3 \Delta e_{24})]^{-1}, \quad (2.6.54)$$

$$B_{66}^1 = (1 + \frac{1}{\pi} G_{66}^1 \Delta C_{66})^{-1}, \quad (2.6.55)$$

$$B_{13}^2 = \frac{\Xi}{4\pi} (B_{11}^1 + B_{12}^1) (G_{11}^1 \Delta e_{31} + G_{12}^1 \Delta e_{31} + G_{13}^1 \Delta e_{33} - G_{31}^2 \Delta \alpha_{33}) \\ + \frac{\Xi}{4\pi} B_{13}^1 (2G_{13}^1 \Delta e_{31} + G_{33}^1 \Delta e_{33} - G_{33}^2 \Delta \alpha_{33}), \quad (2.6.56)$$

$$B_{33}^2 = \frac{\Xi}{2\pi} B_{13}^1 (G_{11}^1 \Delta e_{31} + G_{12}^1 \Delta e_{31} + G_{13}^1 \Delta e_{33} - G_{31}^2 \Delta \alpha_{33}) \\ + \frac{\Xi}{4\pi} B_{33}^1 (2G_{13}^1 \Delta e_{31} + G_{33}^1 \Delta e_{33} - G_{33}^2 \Delta \alpha_{33}), \quad (2.6.57)$$

$$B_{42}^2 = \frac{1}{2\pi} B_{44}^1 (1 + \frac{1}{2\pi} G_{24}^2 \Delta e_{24} - \frac{1}{4\pi} G_{11}^3 \Delta \alpha_{11})^{-1} (G_{44}^1 \Delta e_{24} - \frac{1}{2} G_{24}^2 \Delta \alpha_{11}), \quad (2.6.58)$$

$$B_{31}^3 = \frac{\Xi}{4\pi} (B_{11}^1 + B_{12}^1) (G_{31}^2 \Delta C_{11} + G_{31}^2 \Delta C_{12} + G_{33}^1 \Delta C_{13} + G_{33}^3 \Delta e_{33}) \\ + \frac{\Xi}{4\pi} B_{13}^1 (2G_{13}^2 \Delta C_{13} + G_{33}^2 \Delta C_{33} + G_{33}^3 \Delta e_{33}), \quad (2.6.59)$$

$$B_{33}^3 = \frac{\Xi}{2\pi} B_{13}^1 (G_{31}^2 \Delta C_{11} + G_{31}^2 \Delta C_{12} + G_{33}^1 \Delta C_{13} + G_{33}^3 \Delta e_{33}) \\ + \frac{\Xi}{4\pi} B_{33}^1 (2G_{31}^2 \Delta C_{13} + G_{33}^2 \Delta C_{33} + G_{33}^3 \Delta e_{33}), \quad (2.6.60)$$

$$B_{15}^3 = \frac{1}{\pi} B_{44}^1 \left(1 + \frac{1}{2\pi} G_{24}^2 \Delta e_{24} - \frac{1}{4\pi} G_{11}^3 \Delta \alpha_{11} \right)^{-1} (G_{24}^1 \Delta C_{44} + \frac{1}{2} G_{11}^3 \Delta e_{24}), \quad (2.6.61)$$

$$B_{11}^4 = \left(1 + \frac{1}{2\pi} G_{24}^2 \Delta e_{24} - \frac{1}{4\pi} G_{11}^3 \Delta \alpha_{11} \right)^{-1} \left[1 + \frac{1}{\pi} B_{42}^2 (G_{24}^2 \Delta C_{44} + \frac{1}{2} G_{11}^3 \Delta e_{24}) \right], \quad (2.6.62)$$

$$\begin{aligned} B_{33}^4 = & \Xi + \frac{\Xi}{2\pi} B_{13}^2 (G_{31}^2 \Delta C_{11} + G_{31}^2 \Delta C_{12} + G_{33}^1 \Delta C_{13} + G_{33}^3 \Delta e_{31}) \\ & + \frac{\Xi}{4\pi} B_{33}^2 (2G_{31}^2 \Delta C_{13} + G_{33}^2 \Delta C_{33}). \end{aligned} \quad (2.6.63)$$

The components of G_{ij}^k can be found in Eq. (2.6.43) as

$$G_{11}^1 = N_{1111}, \quad (2.6.64a)$$

$$G_{33}^1 = N_{3333}, \quad (2.6.64b)$$

$$G_{12}^1 = N_{1212}, \quad (2.6.64c)$$

$$G_{13}^1 = N_{1313}, \quad (2.6.64d)$$

$$G_{44}^1 = \frac{1}{4} (2N_{1313} + N_{1133} + N_{3311}), \quad (2.6.64e)$$

$$G_{66}^1 = \frac{1}{2} (N_{1212} + N_{1122}), \quad (2.6.64f)$$

$$G_{13}^2 = N_{1413}, \quad (2.6.64g)$$

$$G_{31}^2 = N_{3411}, \quad (2.6.64h)$$

$$G_{24}^2 = \frac{1}{2} (N_{1413} + N_{3411}), \quad (2.6.64i)$$

$$G_{33}^2 = N_{3433}, \quad (2.6.64j)$$

$$G_{11}^3 = N_{4411}, \quad (2.6.64k)$$

$$G_{33}^3 = N_{4433}. \quad (2.6.64l)$$

2.6.1.2 Jump of the Strain and Electric Field on the Boundary

For the discussion of Eqs. (2.6.23) and (2.6.24) it is shown that when we consider exterior points of the inclusion (i.e. $\mathbf{x}$ is located outside Ω), the integration in

Eq. (2.6.23) on the unit sphere surface S^2 is only partially permitted. The expressions of the elastic fields for exterior points were derived by Mura (1987) for anisotropic elastic materials. Despite the mathematical skills presented, the deduction of the elastic fields for exterior points is an extremely complicated task out of proportion to its usefulness. Instead, we consider here the jumps of the strain field and the electric field on the inclusion boundary. Denote physical quantities for points just outside the inclusion with superscript *Ex* (e.g. $u_J^{Ex}, \sigma_{ij}^{Ex}, \varepsilon_{kl}^{Ex}$), and physical quantities for interior points with superscript *In* (e.g. $u_J^{In}, \sigma_{ij}^{In}, \varepsilon_{kl}^{In}$). The continuity conditions are given by

$$u_J^{Ex} = u_J^{In}, \quad (2.6.65)$$

$$\sigma_{ij}^{Ex} n_i = \sigma_{ij}^{In} n_i. \quad (2.6.66)$$

From Eq. (2.6.65), we know

$$[\varepsilon_{kl}] = \varepsilon_{kl}^{Ex} - \varepsilon_{kl}^{In} = \lambda_K n_l, \quad (2.6.67)$$

where λ_K are constants to be determined, and n_i are components of the outward normal of the inclusion boundary.

For **inclusions** with prescribed eigenstrains ε_{mn}^* , Eq. (2.6.66) can be expanded as

$$C_{ijkl} \varepsilon_{kl}^{Ex} n_i = C_{ijkl} (\varepsilon_{kl}^{In} - \varepsilon_{kl}^*) n_i, \quad (2.6.68)$$

which gives

$$C_{ijkl} \lambda_K n_i n_i = -C_{ijkl} \varepsilon_{kl}^* n_i. \quad (2.6.69)$$

By comparing Eq. (2.6.69) to Eq. (2.6.9), the solution for λ_K can be written as

$$\lambda_Q = -G_{QJ}(\mathbf{n}) C_{ijkl} \varepsilon_{kl}^* n_i. \quad (2.6.70)$$

Substitution of Eq. (2.6.70) into Eq. (2.6.67) gives

$$[\varepsilon_{kl}] = -G_{KJ}(\mathbf{n}) C_{ijkl} \varepsilon_{mn}^* n_i n_j. \quad (2.6.71)$$

For **inhomogeneities** governed by Eq. (2.5.8), Eq. (2.6.66) can be expanded as

$$C_{ijkl} \varepsilon_{kl}^{Ex} n_i = C_{ijkl} (\varepsilon_{kl}^{In} - \varepsilon_{kl}^{**}) n_i, \quad (2.6.72)$$

or

$$C_{ijkl} \varepsilon_{kl}^{Ex} n_i = C_{ijkl}^* \varepsilon_{kl}^{In} n_i. \quad (2.6.73)$$

After manipulation we have

$$[[\varepsilon_{kl}]] = -G_{kl}(\mathbf{n}) \Delta C_{ijkl} \varepsilon_{kl}^{In} n_i n_j. \quad (2.6.74)$$

For *inhomogeneous inclusions* governed by Eq. (2.5.12), Eq. (2.6.66) can be expanded as

$$C_{ijkl} \varepsilon_{kl}^{Ex} n_i = C_{ijkl} (\varepsilon_{kl}^{In} - \varepsilon_{kl}^* - \varepsilon_{kl}^{**}) n_i, \quad (2.6.75)$$

or

$$C_{ijkl} \varepsilon_{kl}^{Ex} n_i = C_{ijkl}^* (\varepsilon_{kl}^{In} - \varepsilon_{kl}^*) n_i. \quad (2.6.76)$$

After manipulation we have

$$[[\varepsilon_{kl}]] = -G_{kl}(\mathbf{n}) (\Delta C_{ijkl} \varepsilon_{kl}^{In} + C_{ijkl}^* \varepsilon_{kl}^*) n_i n_j. \quad (2.6.77)$$

From Eq. (2.6.71) we see that the jump of the generalized strain tensor across the boundary of inclusion can be directly obtained, while for inhomogeneities and inhomogeneous inclusions, the corresponding expressions in Eqs. (2.6.74) and (2.6.77) depend on ε_{kl}^{In} , the generalized strain tensor for interior points.

2.6.2 Flat Elliptical Cracks

Consider a flat elliptical crack ($x_1^2/a_1^2 + x_2^2/a_2^2 = 1$) in an anisotropic, piezoelectric medium, occupying the subdomain Ω in an infinite matrix. The crack is degenerated from the ellipsoidal cavity ($x_1^2/a_1^2 + x_2^2/a_2^2 + x_3^2/a_3^2 = 1$) by taking $a_3 \rightarrow 0$. It is assumed that the crack is filled with vacuum, and for the boundary condition of stresses and electric displacements Eqs. (2.2.1) and (2.2.3) are specified as

$$\sigma_{3J}^{In} = 0 \quad (J = 1, 2, 3, 4). \quad (2.6.78)$$

The solution for the generalized stress fields for inhomogeneity is given by Eq. (2.6.38) as

$$\boldsymbol{\sigma} = \mathbf{C}^* (\mathbf{I} - \mathbf{S} \mathbf{C}^{-1} \Delta \mathbf{C})^{-1} \boldsymbol{\varepsilon}^\infty. \quad (2.6.79)$$

Yet Eq. (2.6.79) is not suitable for analyzing cracks. The solution of the Eshelby tensor $\mathbf{S}$ relies on the equivalent inclusion method introduced in Section 2.5.2. As stated by Mura (1987), the application of the equivalent inclusion method to voids and cracks causes the so called “impotent eigenstrains”. (When the inhomogeneity is a void, the solution for the equivalent eigenstrains ε_{ij}^{**} is not unique. That is to say, there are a lot of choices of equivalent eigenstrains which do not generate any stress field within the material.)

Therefore, the problem for flat elliptical cracks needs re-deduction from the start. For the piezoelectric matrix, the solution was obtained by Wang (1992b).

The constitutive equation for the system of the piezoelectric matrix containing a crack is given by

$$\sigma_{ij} = C_{ijkl}^0(\mathbf{x}) \varepsilon_{kl}, \quad (2.6.80)$$

where

$$C_{ijkl}^0(\mathbf{x}) = C_{ijkl} - C_{ijkl} H(\mathbf{x}), \quad (2.6.81)$$

$$H(\mathbf{x}) = \begin{cases} 1, & \mathbf{x} \in \Omega \\ 0, & \mathbf{x} \notin \Omega \end{cases}. \quad (2.6.82)$$

The equilibrium equation is

$$C_{ijkl} u_{K,li} = -C_{ijkl} H(\mathbf{x}) u_{K,li}. \quad (2.6.83)$$

Using the Green's function method, the solution for the generalized displacement fields in Eq. (2.6.83) is given by

$$u_K(\mathbf{x}) = u_K^\infty + \int_{-\infty}^{\infty} C_{iJm} H(\mathbf{x}) u_{M,n} G_{KJ,i}(\mathbf{x} - \mathbf{x}') d\mathbf{x}', \quad (2.6.84)$$

where u_K^∞ denotes the generalized displacement fields caused by generalized constant strain fields ε_{kl}^∞ (or generalized constant stress fields σ_{ij}^∞) applied at infinity. The corresponding expressions for the generalized strains are

$$\varepsilon_{KL}(\mathbf{x}) = \begin{cases} \varepsilon_{KL}^{\infty} + \frac{1}{2} \int_{-\infty}^{\infty} C_{iJMn} H(\mathbf{x}) u_{M,n} [G_{KJ,li}(\mathbf{x} - \mathbf{x}') + G_{LJ,Ki}(\mathbf{x} - \mathbf{x}')] d\mathbf{x}', & K = 1, 2, 3 \\ \varepsilon_{KL}^{\infty} + \int_{-\infty}^{\infty} C_{iJMn} H(\mathbf{x}) u_{M,n} G_{KJ,li}(\mathbf{x} - \mathbf{x}') d\mathbf{x}', & K = 4 \end{cases} \quad (2.6.85)$$

Substitution of the Green's function and its derivatives from Eqs. (2.6.8)–(2.6.12) into Eq. (2.6.85) gives

$$\varepsilon_{KL}(\mathbf{x}) = \varepsilon_{KL}^{\infty} + \begin{cases} \frac{1}{16\pi^3} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} C_{iJMn} H(\mathbf{x}) u_{M,n} \{ \bar{G}_{KJ}(\xi) \xi_l \xi_i + \bar{G}_{LJ}(\xi) \xi_K \xi_i \} \exp[i\xi \cdot (\mathbf{x} - \mathbf{x}')] d\xi d\mathbf{x}', & K = 1, 2, 3 \\ \frac{1}{8\pi^3} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} C_{iJMn} H(\mathbf{x}) u_{M,n} \bar{G}_{KJ}(\xi) \xi_l \xi_i \exp[i\xi \cdot (\mathbf{x} - \mathbf{x}')] d\xi d\mathbf{x}', & K = 4 \end{cases} \quad (2.6.86)$$

By using the constitutive relation Eq. (2.6.80), the generalized stresses can be derived as

$$\begin{aligned} \sigma_{iJ} &= C_{iJKl}^0(\mathbf{x}) \varepsilon_{KL} \\ &= C_{iJKl} \varepsilon_{KL} - C_{iJKl} H(\mathbf{x}) \varepsilon_{KL} \\ &= C_{iJKl} \varepsilon_{KL}^{\infty} + C_{iJKl} \frac{1}{8\pi^3} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} C_{pQMn} H(\mathbf{x}) \varepsilon_{Mn} \bar{G}_{KQ}(\xi) \xi_l \xi_p \exp[i\xi \cdot (\mathbf{x} - \mathbf{x}')] d\xi d\mathbf{x}' \\ &\quad - C_{iJKl} H(\mathbf{x}) \varepsilon_{KL}. \end{aligned} \quad (2.6.87)$$

2.6.2.1 The Crack Opening Displacement and Interaction Energy

For uniform fields applied at infinity, it is proved that inside the crack the generalized strains are uniform. Denote the generalized strains and stresses for interior points as ε_{KL}^{In} and σ_{KL}^{In} . Eq. (2.6.87) is transformed as

$$\begin{aligned} \sigma_{ij}^{In} &= \sigma_{ij}^{\infty} - C_{iJKl} H(\mathbf{x}) \varepsilon_{KL}^{In} \\ &\quad + C_{iJKl} \frac{1}{8\pi^3} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} C_{pQMn} H(\mathbf{x}) \varepsilon_{Mn}^{In} \bar{G}_{KQ}(\xi) \xi_l \xi_p \exp[i\xi \cdot (\mathbf{x} - \mathbf{x}')] d\xi d\mathbf{x}'. \end{aligned} \quad (2.6.88)$$

Referring to the deduction procedure from Eq. (2.6.14) to Eq. (2.6.25), Eq. (2.6.87) can be transformed as

$$\sigma_{ij}^{In} = \sigma_{ij}^{\infty} - C_{ijkl} H(\mathbf{x}) \varepsilon_{kl}^{In} + \frac{a_1 a_2 a_3}{4\pi} C_{ijkl} C_{pqmn} \varepsilon_{mn}^{In} \int_{S^2} \bar{G}_{KQ}(\mathbf{w}) \rho^{-3} w_l w_p dS(\mathbf{w}). \quad (2.6.89)$$

Notice that for $a_3 \rightarrow 0$, certain components of ε_{Mn}^{In} approach infinity. To satisfy the boundary condition Eq. (2.6.78), $a_3 \varepsilon_{Mn}^{In}$ must remain finite. Since $\bar{G}_{KQ}(\mathbf{w}) w_l w_p$ are homogeneous functions of the same kind with degree zero, we will deal with only one integral in Eq. (2.6.89). It holds that

$$\begin{aligned} I_{iJQp} &= \frac{a_1 a_2 a_3}{4\pi} \int_{S^2} C_{ijkl} \bar{G}_{KQ}(\mathbf{w}) \rho^{-3} w_l w_p dS(\mathbf{w}) \\ &= \int_{S^2} \frac{C_{ijkl} \bar{G}_{KQ}(\mathbf{w}) w_l w_p dS(\mathbf{w})}{a_1^2 a_2^2 a_3^2 [w_1^2 / (a_2^2 a_3^2) + w_2^2 / (a_1^2 a_3^2) + w_3^2 / (a_1^2 a_2^2)]^{3/2}}. \end{aligned} \quad (2.6.90)$$

Introducing the following transformation:

$$\begin{aligned} w_1 &= (1 - w_3^2)^{1/2} \cos \theta, \\ w_2 &= (1 - w_3^2)^{1/2} \sin \theta, \\ t &= w_3 / (1 - w_3^2)^{1/2}, \\ dS(\mathbf{w}) &= dw_3 d\theta = (1 - w_3^2)^{3/2} dt d\theta, \end{aligned} \quad (2.6.91)$$

Eq. (2.6.90) can be transformed as

$$I_{iJQp} = \int_0^{2\pi} \frac{d\theta}{a_1^2 a_2^2 a_3^2} \int_{-\infty}^{\infty} \frac{B_{iJQp}(\cos \theta, \sin \theta, t) dt}{[\cos^2 \theta / (a_2^2 a_3^2) + \sin^2 \theta / (a_1^2 a_3^2) + t^2 / (a_1^2 a_2^2)]^{3/2}}, \quad (2.6.92)$$

where

$$B_{iJQp}(\mathbf{w}) = C_{ijkl} \bar{G}_{KQ}(\mathbf{w}) w_l w_p. \quad (2.6.93)$$

By integrating Eq. (2.6.92) by parts with respect to t , using

$$\begin{aligned} \frac{\partial}{\partial t} \{t / [\cos^2 \theta / (a_2^2 a_3^2) + \sin^2 \theta / (a_1^2 a_3^2) + t^2 / (a_1^2 a_2^2)]^{1/2}\} \\ = [\cos^2 \theta / (a_2^2 a_3^2) + \sin^2 \theta / (a_1^2 a_3^2)] / [\cos^2 \theta / (a_2^2 a_3^2) + \sin^2 \theta / (a_1^2 a_3^2) + t^2 / (a_1^2 a_2^2)]^{3/2} \end{aligned} \quad (2.6.94)$$

and

$$B_{iJQp} \Big|_{t \rightarrow \infty} = B_{iJQp} \Big|_{t \rightarrow -\infty} = B_{iJQp}(0, 0, 1), \quad (2.6.95)$$

Eq. (2.6.93) becomes

$$\begin{aligned} I_{iJQp} &= 2B_{iJQp}(0, 0, 1) \int_0^{2\pi} \frac{d\theta}{a_1 a_2 (\cos^2 \theta / a_2^2 + \sin^2 \theta / a_1^2)} - a_3 \int_0^{2\pi} \frac{d\theta}{a_1^2 a_2^2} \\ &\times \int_{-\infty}^{\infty} (\cos^2 \theta / a_2^2 + \sin^2 \theta / a_1^2)^{-1} [\cos^2 \theta / a_2^2 + \sin^2 \theta / a_1^2 + a_3^2 t^2 / (a_1^2 a_2^2)]^{-1/2} \\ &\times t \frac{\partial B_{iJQp}(\cos \theta, \sin \theta, t)}{\partial t} dt. \end{aligned} \quad (2.6.96)$$

Noticing that

$$\int_0^{2\pi} \frac{d\theta}{\cos^2 \theta / a_2^2 + \sin^2 \theta / a_1^2} = 2\pi a_1 a_2, \quad (2.6.97)$$

and that $a_3 \ll a_1, a_2$, we finally have

$$I_{iJQp} = 4\pi B_{iJQp}^2(0, 0, 1) - a_3 \eta_{iJQp}, \quad (2.6.98)$$

where

$$\eta_{iJQp} = \int_0^{2\pi} \frac{d\theta}{a_1^2 a_2^2} \int_{-\infty}^{\infty} (\cos^2 \theta / a_2^2 + \sin^2 \theta / a_1^2)^{-3/2} t \frac{\partial B_{iJQp}(\cos \theta, \sin \theta, t)}{\partial t} dt. \quad (2.6.99)$$

Substitution of Eq. (2.6.99) into Eq. (2.6.89) gives

$$\sigma_{iJ}^{In} = \sigma_{iJ}^{\infty} - C_{iJKl} H(\mathbf{x}) \varepsilon_{Kl}^{In} + \frac{1}{4\pi} C_{pQMn} \varepsilon_{Mn}^{In} [4\pi B_{iJQp}(0, 0, 1) - a_3 \eta_{iJQp}]. \quad (2.6.100)$$

According to the definition Eq. (2.6.93) and the property Eq. (2.6.9), one obtains

$$B_{3JQp}(0, 0, 1) = -\delta_{JQ} \delta_{p3}. \quad (2.6.101)$$

Substitution of Eq. (2.6.101) into Eq. (2.6.78) gives

$$\sigma_{3J}^{In} = \sigma_{3J}^{\infty} - \frac{a_3}{4\pi} C_{pQMn} \varepsilon_{Mn}^{In} \eta_{3JQp} = 0. \quad (2.6.102)$$

Since only $a_3 \varepsilon_{M3}^{In}$ remains finite for $a_3 \rightarrow 0$, Eq. (2.6.102) provides a set of

linear algebraic equations for determining $a_3 \varepsilon_{M3}^{In}$. Now we try to deal with the tensor η_{ijQp} .

$$\begin{aligned} \eta_{ijQp} &= \int_0^{2\pi} \frac{d\theta}{a_1^2 a_2^2} \int_{-\infty}^{\infty} (\cos^2 \theta / a_2^2 + \sin^2 \theta / a_1^2)^{-3/2} t \frac{\partial B_{ijQp}(\cos \theta, \sin \theta, t)}{\partial t} dt \\ &= C_{ijkl} L_{pQkl}, \end{aligned} \quad (2.6.103)$$

where

$$L_{pQkl} = \int_0^{2\pi} \frac{d\theta}{a_1^2 a_2^2} \int_{-\infty}^{\infty} (\cos^2 \theta / a_2^2 + \sin^2 \theta / a_1^2)^{-3/2} t \frac{\partial (\bar{G}_{kQ}(\mathbf{w}) w_l w_p)}{\partial t} dt, \quad (2.6.104)$$

and w_p can be obtained from Eq. (2.6.91). Since $\bar{G}_{kQ}(\mathbf{w}) w_l w_p$ are all polynomial functions of t with maximum degree 8 divided by a polynomial function of t with maximum degree 8, Eq. (2.6.104) can be expressed as

$$L_{pQkl} = - \int_0^{2\pi} \frac{d\theta}{a_1^2 a_2^2 (\cos^2 \theta / a_2^2 + \sin^2 \theta / a_1^2)^{3/2}} \int_{-\infty}^{\infty} M_{ijQp}(\cos \theta, \sin \theta, t) dt, \quad (2.6.105)$$

where M_{ijQp} are largest proper fractional expressions contained in $\bar{G}_{kQ}(\mathbf{w}) w_l w_p$. Eq. (2.6.105) certainly leads to

$$L_{pQkl} = - \int_0^{2\pi} \frac{d\theta}{a_1^2 a_2^2 (\cos^2 \theta / a_2^2 + \sin^2 \theta / a_1^2)^{3/2}} 2\pi i \sum_{\text{Im } z > 0} \text{Rez}[M_{ijQp}(\cos \theta, \sin \theta, t)], \quad (2.6.106)$$

where $\text{Rez}[M_{ijQp}(\cos \theta, \sin \theta, t)]$ denotes the residues of $M_{ijQp}(\cos \theta, \sin \theta, t)$. Since $M_{ijQp}(\cos \theta, \sin \theta, t)$ can be expressed as a polynomial of t , the residues can be easily calculated.

Using Eq. (2.6.103), Eq. (2.6.102) can be rewritten as

$$\sigma_{3J}^{In} = \sigma_{3J}^{\infty} - \frac{a_3}{4\pi} C_{pQMn} \varepsilon_{Mn}^{In} C_{ijkl} L_{pQkl} = 0. \quad (2.6.107)$$

Eq. (2.6.107) provides the basic equations for solving $a_3 \varepsilon_{M3}^{In}$.

Since the crack is regarded as an ellipsoidal void by letting $a_3 \rightarrow 0$, the generalized crack opening displacement is given by

$$U_K = 2 \int_0^h \varepsilon_{K3}^{In} dx_3 = 2a_3 \varepsilon_{K3}^{In} (1 - x_1^2 / a_1^2 - x_2^2 / a_2^2)^{1/2}, \quad (2.6.108)$$

where h is the half thickness of the ellipsoidal void in the x_3 -direction. It is necessary to explain here that U_4 of the generalized crack opening displacement is called the discontinuity in electric potential across the crack.

If the potential energy of the piezoelectric body associated with the presence of the crack is defined by W , one obtains

$$W = W^m + W^e. \quad (2.6.109)$$

W^m denotes the mechanical work, or in more detail, the work done by the applied stress acting (gradually) through the displacement of each crack face, which is given by

$$W^m = \frac{1}{2} \iint_A \sigma_{3i}^0 U_i dx_1 dx_2 = \frac{2}{3} \pi a_1 a_2 a_3 \sigma_{3i}^0 \varepsilon_{3i}^{In}, \quad (2.6.110)$$

where A is the area of the crack. W^e denotes the electrical work, or in more detail, the work done by the electric displacement through the discontinuity in electric potential across the crack, which is given by

$$W^e = -\frac{1}{2} \iint_A \sigma_{34}^0 U_4 dx_1 dx_2 = -\frac{2}{3} \pi a_1 a_2 a_3 \sigma_{34}^0 \varepsilon_{34}^{In} = \frac{2}{3} \pi a_1 a_2 a_3 D_3^0 E_3^{In}. \quad (2.6.111)$$

Eq. (2.6.109) is rewritten as

$$W = \frac{2}{3} \pi a_1 a_2 a_3 (\sigma_{3i}^0 \varepsilon_{3i}^{In} + D_3^0 E_3^{In}). \quad (2.6.112)$$

2.6.2.2 The Stress Intensity Factors

The deduction of the generalized stresses in Eq. (2.6.100) is restricted to interior points of the crack. For discussion of the stress intensity factors, we have to obtain knowledge about the trend of the generalized stresses when approaching the crack from the matrix. For uniform fields applied at infinity, the generalized stresses for exterior points of the crack are obtained from Eq. (2.6.88) as

$$\begin{aligned}
\sigma_{ij} &= C_{ijkl} \varepsilon_{kl}^{\infty} + C_{ijkl} \frac{1}{8\pi^3} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} C_{pQMn} H(\mathbf{x}) \varepsilon_{Mn} \bar{G}_{KQ}(\xi) \xi_l \xi_p \exp[i\xi \cdot (\mathbf{x} - \mathbf{x}')] d\xi d\mathbf{x}' \\
&= \sigma_{ij}^{\infty} + \frac{1}{8\pi^3} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} C_{pQMn} H(\mathbf{x}) \varepsilon_{Mn} B_{iQp}(\xi) \exp[i\xi \cdot (\mathbf{x} - \mathbf{x}')] d\xi d\mathbf{x}'.
\end{aligned} \tag{2.6.113}$$

Since the integration with respect to $\mathbf{x}'$ is performed in the crack, using Eq. (2.6.108), Eq. (2.6.113) can be transformed as

$$\begin{aligned}
\sigma_{ij} &= \sigma_{ij}^{\infty} + \frac{1}{8\pi^3} \int_{-\infty}^{\infty} B_{iQp}(\xi) d\xi \iint_A C_{pQM3} U_M \exp\{i[\xi_1(x_1 - x_1') \\
&\quad + \xi_2(x_2 - x_2') + \xi_3 x_3]\} dx_1 dx_2,
\end{aligned} \tag{2.6.114}$$

where A is the area of the crack surface.

According to Eq. (2.6.93), Eq. (2.6.9) and Eq. (2.5.22), $B_{iQp}(\xi)$ can be rewritten as

$$B_{iQp}(\xi) = \frac{N_{iQp}}{\det(\mathbf{D})}, \tag{2.6.115}$$

where $\mathbf{D}$ has components defined in Eq. (2.5.22), and

$$N_{iQp} = C_{ijkl} \hat{G}_{KQ}(\xi) \xi_l \xi_p. \tag{2.6.116}$$

$\hat{G}_{KQ}(\xi)$ are cofactors of $\mathbf{D}$.

The integral with the form of Eq. (2.6.114) was firstly treated by Willis (1967; 1968), for an anisotropic elastic matrix. For a piezoelectric matrix, the solution was obtained by Wang (1992b) as

$$\begin{aligned}
\sigma_{ij} &= \sigma_{ij}^{\infty} + \frac{a_3}{\pi} C_{pQM3} \varepsilon_{M3}^{\infty} \int_0^{2\pi} d\varphi \sum_{N=1}^4 R_{iQp}[\xi_1 / a_1, \xi_2 / a_2, \xi_3^N(\xi_1 / a_1, \xi_2 / a_2)] \\
&\quad \times \left(\frac{g}{g^2 - 1} - \frac{1}{2} \log \frac{g+1}{g-1} \right),
\end{aligned} \tag{2.6.117}$$

where $\xi_3^N = \xi_3^N(\xi_1, \xi_2)$ ($N=1,2,3,4$) are roots of the equation $\det(\mathbf{D})=0$ with positive imaginary parts for $x_3 > 0$, and

$$R_{iQp}(\xi_1, \xi_2, \xi_3^N) = N_{iQp}(\xi^N) \left\{ \left[\frac{\partial \det(\mathbf{D})}{\partial \xi_3} \right] \Big|_{\xi=\xi^N} \right\}^{-1}, \tag{2.6.118}$$

$$\varsigma_1 = \cos \varphi, \varsigma_2 = \sin \varphi, y_1 = x_1 / a_1, y_2 = x_2 / a_2, \quad (2.6.119)$$

$$g = -x_3 \xi_3^N (\varsigma_1 / a_1, \varsigma_2 / a_2) - (\varsigma_1 y_1 + \varsigma_2 y_2). \quad (2.6.120)$$

It should be noticed that $\xi_3^N(\xi_1, \xi_2)$ is homogeneous of degree 1.

The stress intensity factor K_{pq}^σ and the electric field intensity factor K_p^D are defined by

$$K_{pq}^\sigma = \lim_{r \rightarrow 0} (2\pi r)^{1/2} \sigma_{pq}, \quad (2.6.121)$$

$$K_p^D = \lim_{r \rightarrow 0} (2\pi r)^{1/2} \sigma_{p4} = \lim_{r \rightarrow 0} (2\pi r)^{1/2} D_p, \quad (2.6.122)$$

where r is the normal distance from the edge of the crack. For a flat elliptical crack,

$$r = (y-1) / (x_1^2 / a_1^4 + x_2^2 / a_2^4)^{1/2}. \quad (2.6.123)$$

When $\mathbf{x}$ approaches the crack edge, $\log \frac{g+1}{g-1} \ll \frac{g}{g^2-1}$. Accordingly, we consider the integral

$$V(\mathbf{x}) = \frac{1}{\pi} \int_0^{2\pi} \frac{f(\boldsymbol{\varsigma})}{g^2-1} d\varphi, \quad (2.6.124)$$

where

$$g \approx -(\varsigma_1 y_1 + \varsigma_2 y_2). \quad (2.6.125)$$

We have

$$V(\mathbf{x}) = \frac{1}{\pi} \int_0^{2\pi} \frac{f(\boldsymbol{\varsigma}) - f(\mathbf{y}/y)}{g^2-1} d\varphi + \frac{f(\mathbf{y}/y)}{\pi} \int_0^{2\pi} \frac{d\varphi}{g^2-1}. \quad (2.6.126)$$

The first term in Eq. (2.6.126) remains bounded as $y \rightarrow 1$, so that

$$V(\mathbf{x}) = \frac{2f(\mathbf{y}/y)}{(y^2-1)^{1/2}} i + o(1). \quad (2.6.127)$$

Thus Eq. (2.6.121) and Eq. (2.6.122) become

$$K_{ij}^{\sigma} = -2\sqrt{\pi i} a_3 (x_1^2 / a_1^4 + x_2^2 / a_2^4)^{-1/4} \times \left\{ C_{pQM3} \varepsilon_{M3}^{In} \sum_{N=1}^4 R_{ijQp} [\zeta_1 / a_1, \zeta_2 / a_2, \xi_3^N (\zeta_1 / a_1, \zeta_2 / a_2)] \right\}, \quad (2.6.128)$$

$$K_i^D = -2\sqrt{\pi i} a_3 (x_1^2 / a_1^4 + x_2^2 / a_2^4)^{-1/4} \times \left\{ C_{pQM3} \varepsilon_{M3}^{In} \sum_{N=1}^4 R_{i4Qp} [\zeta_1 / a_1, \zeta_2 / a_2, \xi_3^N (\zeta_1 / a_1, \zeta_2 / a_2)] \right\}. \quad (2.6.129)$$

When $a_3 \varepsilon_{M3}^{In}$ is solved from Eq. (2.6.108), the stress intensity factor K_{pq}^{σ} and the electric field intensity factor K_p^D near the crack edge can be calculated from Eq. (2.6.128) and Eq. (2.6.129). From the results of K_{pq}^{σ} and K_p^D , it is shown that for a flat elliptical crack, the singularity factor of the stresses and electric displacement at the crack edge coincides with the classical one, which is $1/\sqrt{r}$.

2.6.3 Ellipsoidal Inhomogeneity Embedded in an Infinite Matrix when both Phases Undergo Eigenstrains

Consider a more complicated situation, where an ellipsoidal inhomogeneity, occupying Ω , is embedded in an infinite matrix, and both the inhomogeneity and the matrix possess preassigned eigenstrains, denoted respectively by ε_{ij}^I and ε_{ij}^M . Here no electromechanical property is assumed for the materials. For such a material system, the total strains at infinity are restricted by ε_{ij}^0 . We have

$$\varepsilon_{ij} = e_{ij} + \varepsilon_{ij}^M + \varepsilon_{ij}^R H(\mathbf{x}), \quad (2.6.130)$$

where

$$H(\mathbf{x}) = \begin{cases} 1, & \mathbf{x} \in \Omega \\ 0, & \mathbf{x} \notin \Omega \end{cases}, \quad (2.6.131)$$

and

$$\varepsilon_{ij}^R = \varepsilon_{ij}^I - \varepsilon_{ij}^M. \quad (2.6.132)$$

ε_{ij} , e_{ij} , and ε_{ij}^R are the total strains, elastic strains and the relative eigenstrains of the inhomogeneity with respect to the matrix, respectively. On the other hand, the total strains can be rewritten as

$$\varepsilon_{ij} = \varepsilon_{ij}^0 + \tilde{\varepsilon}_{ij}, \quad (2.6.133)$$

where ε_{ij}^0 is the restricted strain applied at infinity, and $\tilde{\varepsilon}_{ij}$ is the disturbance strain caused by existing eigenstrains and the inhomogeneity. According to Eq. (2.6.130), we have

$$e_{ij} = \varepsilon_{ij}^\infty + \tilde{\varepsilon}_{ij} - \varepsilon_{ij}^R H(\mathbf{x}), \quad (2.6.134)$$

where

$$\varepsilon_{ij}^\infty = \varepsilon_{ij}^0 - \varepsilon_{ij}^M. \quad (2.6.135)$$

ε_{ij}^∞ denotes the elastic strain applied at infinity. The constitutive equations are given by

$$\sigma_{ij} = \begin{cases} C_{ijkl}(\varepsilon_{kl}^\infty + \tilde{\varepsilon}_{kl}), & \mathbf{x} \notin \Omega \\ C_{ijkl}^*(\varepsilon_{kl}^\infty + \tilde{\varepsilon}_{kl} - \varepsilon_{kl}^R), & \mathbf{x} \in \Omega \end{cases}. \quad (2.6.136)$$

Using the equivalent inclusion method

$$C_{ijkl}^*(\varepsilon_{kl}^\infty + \tilde{\varepsilon}_{kl} - \varepsilon_{kl}^R) = C_{ijkl}(\varepsilon_{kl}^\infty + \tilde{\varepsilon}_{kl} - \varepsilon_{kl}^R - \varepsilon_{kl}^{**}), \quad (2.6.137)$$

the equilibrium equation can be written as

$$C_{ijkl}(\varepsilon_{kl}^\infty + \tilde{\varepsilon}_{kl})_{,i} = H(\mathbf{x})C_{ijkl}(\varepsilon_{kl}^R + \varepsilon_{kl}^{**})_{,i}. \quad (2.6.138)$$

If $\varepsilon_{kl,i}^M = 0$, Eq. (2.6.138) is transformed as

$$C_{ijkl}\tilde{\varepsilon}_{kl,i} = H(\mathbf{x})C_{ijkl}(\varepsilon_{kl}^R + \varepsilon_{kl}^{**})_{,i}. \quad (2.6.139)$$

The solution for Eq. (2.6.139) for interior points of the inhomogeneity is given by

$$\tilde{\varepsilon}_{ij} = S_{ijkl}(\varepsilon_{kl}^R + \varepsilon_{kl}^{**}), \quad \mathbf{x} \in \Omega, \quad (2.6.140)$$

where S_{ijkl} is the Eshelby tensor. ε_{kl}^{**} can be obtained in Eq. (2.6.137) as

$$\varepsilon^{**} = \mathbf{C}^{-1} \Delta \mathbf{C}(\varepsilon^\infty + \tilde{\varepsilon} - \varepsilon^R), \quad (2.6.141)$$

where the matrix notation is defined in Eqs. (2.6.36) and (2.6.37). Substitution of Eq. (2.6.141) into Eq. (2.6.140) leads to

$$\boldsymbol{\varepsilon}^{**} = \mathbf{A}\boldsymbol{\varepsilon}^R + \mathbf{B}\boldsymbol{\varepsilon}^\infty, \quad (2.6.142)$$

where

$$\begin{aligned} \mathbf{A} &= (\mathbf{C} - \mathbf{S}\Delta\mathbf{C})^{-1} \mathbf{C}^* - \mathbf{I}, \\ \mathbf{B} &= (\mathbf{C} - \mathbf{S}\Delta\mathbf{C})^{-1} \Delta\mathbf{C}. \end{aligned} \quad (2.6.143)$$

Substitution of Eq. (2.6.142) into Eq. (2.6.140) gives

$$\tilde{\boldsymbol{\varepsilon}} = \mathbf{S}(\mathbf{A} + \mathbf{I})\boldsymbol{\varepsilon}^R + \mathbf{S}\mathbf{B}\boldsymbol{\varepsilon}^\infty. \quad (2.6.144)$$

2.6.4 Green's Function

The three-dimensional static Green's function is determined by the equation

$$C_{ijkl}G_{KM,li}(\mathbf{x} - \mathbf{x}') = -\delta_{JM}\delta(\mathbf{x} - \mathbf{x}'), \quad (2.6.145)$$

where $\delta(\mathbf{x} - \mathbf{x}')$ is the three-dimensional Dirac delta function. In Eq. (2.6.145), $G_{KJ}(\mathbf{x} - \mathbf{x}')$ denotes the generalized displacement component in point $\mathbf{x}$ in the x_K direction (for $K = 4$, the electric potential is considered) caused by application of a generalized unit point body force to the point $\mathbf{x}'$ in the x_J direction (for $J = 4$, a unit extraneous body charge density is defined in $\mathbf{x}'$). Without loss of generality, let $\mathbf{x}'$ be the origin of the Cartesian coordinate. Eq. (2.6.145) transforms to

$$C_{ijkl}G_{KM,li}(\mathbf{x}) = -\delta_{JM}\delta(\mathbf{x}). \quad (2.6.146)$$

Now we use the Radon transform introduced in Appendix A to solve the problem. Consider the Radon transform of the three-dimensional Dirac delta, which satisfies

$$\hat{\delta}(s, \mathbf{n}) = R[\delta(\mathbf{x})] = \int \delta(\mathbf{x})\delta(s - \mathbf{n} \cdot \mathbf{x})d\mathbf{x} = \delta(s - \mathbf{n} \cdot \mathbf{x})\big|_{\mathbf{x}=0} = \delta(s). \quad (2.6.147)$$

Eq. (2.6.147) indicates that the Radon transform of the three-dimensional Dirac delta gives a one-dimensional Dirac delta. Let both sides of Eq. (2.6.146) take the Radon transform simultaneously. We have

$$C_{ijkl}R[G_{KM,li}(\mathbf{x})] = -\delta_{JM}\delta(s). \quad (2.6.148)$$

According to the properties of Radon transform introduced in Appendix A, the left side of Eq. (2.6.148) can be transformed as

$$C_{ijkl}R[G_{KM,li}(\mathbf{x})] = C_{ijkl}n_l n_i \frac{\partial^2 \hat{G}_{KM}(s, \mathbf{n})}{\partial s^2}. \quad (2.6.149)$$

Substitution of Eq. (2.6.149) into Eq. (2.6.146) gives

$$\frac{\partial^2 \hat{G}_{KM}(s, \mathbf{n})}{\partial s^2} = -\delta_{JM}\delta(s)\Gamma_{JK}^{-1}(\mathbf{n}) = -\delta(s)\Gamma_{KM}^{-1}(\mathbf{n}), \quad (2.6.150)$$

where

$$\Gamma_{JK}(\mathbf{n}) = C_{ijkl}n_l n_i, \quad (2.6.151)$$

and $\Gamma_{JK}^{-1}(\mathbf{n})$ denote components of the inverse matrix of Γ . In deduction of Eq. (2.6.150), the symmetric property of Γ and its inverse matrix is used. The inverse Radon transform of the Green's function is given by

$$G_{KM}(\mathbf{x}) = -\frac{1}{8\pi^2} \int_S \frac{\partial^2 \hat{G}_{KM}(s, \mathbf{n})}{\partial s^2} \Big|_{s=\mathbf{n} \cdot \mathbf{x}} dS(\mathbf{n}). \quad (2.6.152)$$

Substitution of Eq. (2.6.150) into Eq. (2.6.152) yields

$$G_{KM}(\mathbf{x}) = \frac{1}{8\pi^2} \int_S [\delta(s)\Gamma_{KM}^{-1}(\mathbf{n})] \Big|_{s=\mathbf{n} \cdot \mathbf{x}} dS(\mathbf{n}) = \frac{1}{8\pi^2} \int_S \delta(\mathbf{n} \cdot \mathbf{x}) \Gamma_{KM}^{-1}(\mathbf{n}) dS(\mathbf{n}). \quad (2.6.153)$$

Eq. (2.6.153) can also be expressed as

$$G_{KM}(\mathbf{x}) = \frac{1}{8\pi^2} \int_S \frac{A_{KM}(\mathbf{n})}{D(\mathbf{n})} \delta(\mathbf{n} \cdot \mathbf{x}) dS(\mathbf{n}), \quad (2.6.154)$$

where $A_{KM}(\mathbf{n})$ denote components of the adjoint matrix of $\Gamma(\mathbf{n})$ and $D(\mathbf{n}) = \det[\Gamma(\mathbf{n})]$ is the determinant of $\Gamma(\mathbf{n})$.

Now we try to deal with the integral over the surface $S(\mathbf{n})$. To expand the terms

in Eq. (2.6.154) of $\mathbf{n}$, a new set of coordinates is constructed and denoted by $(\mathbf{e}, \mathbf{p}, \mathbf{q})$.

As illustrated in Fig. 2.5,

$$\mathbf{e} = \frac{\mathbf{x}}{r}, \quad (2.6.155)$$

where $r=|\mathbf{x}|$. $\mathbf{p}$ is a unit vector perpendicular to $\mathbf{e}$, and

$$\mathbf{q} = \mathbf{e} \times \mathbf{p}. \quad (2.6.156)$$

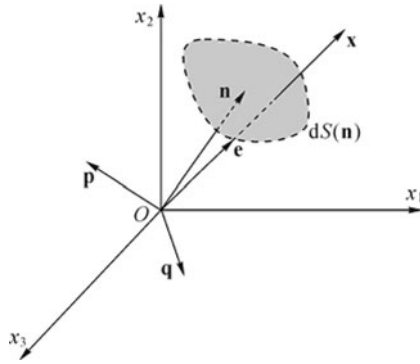


Fig. 2.5 Geometry of the new coordinate $(\mathbf{e}, \mathbf{p}, \mathbf{q})$ and its relation to $\mathbf{x}$ and $\mathbf{n}$

Mapping $\mathbf{n}$ to the new coordinate $(\mathbf{e}, \mathbf{p}, \mathbf{q})$, we have

$$\mathbf{n} = \xi \mathbf{p} + \zeta \mathbf{q} + \eta \mathbf{e}. \quad (2.6.157)$$

According to Eq. (2.6.155), we have

$$\mathbf{x} \cdot \mathbf{n} = \xi \mathbf{x} \cdot \mathbf{p} + \zeta \mathbf{x} \cdot \mathbf{q} + \eta \mathbf{x} \cdot \mathbf{e} = r\eta. \quad (2.6.158)$$

Therefore Eq. (2.6.154) can be expressed as

$$G_{KM}(\mathbf{x}) = \frac{1}{8\pi^2} \int_S \frac{A_{KM}(\xi \mathbf{p} + \zeta \mathbf{q} + \eta \mathbf{e})}{D(\xi \mathbf{p} + \zeta \mathbf{q} + \eta \mathbf{e})} \delta(r\eta) dS(\xi, \zeta, \eta). \quad (2.6.159)$$

Since the surface $S(\xi, \zeta, \eta)$ can be any closed one that contains the origin $\xi = \zeta = \eta = 0$, we assume $S(\xi, \zeta, \eta)$ to be the surface of a rectangular parallelepiped bounded by six rectangles S_1 and S_2 ($\xi = \pm 1$), S_3 and S_4 ($\zeta = \pm L$),

S_5 and S_6 ($\eta = \pm L$).

Eq. (2.6.159) can then be expressed as

$$G_{KM}(\mathbf{x}) = \frac{1}{8\pi^2} \sum_{i=1}^6 \int_{S_i} \frac{A_{KM}(\xi \mathbf{p} + \zeta \mathbf{q} + \eta \mathbf{e})}{D(\xi \mathbf{p} + \zeta \mathbf{q} + \eta \mathbf{e})} \delta(r\eta) dS_i(\xi, \zeta, \eta). \quad (2.6.160)$$

From Eq. (2.6.151), it is clear that $D(\xi \mathbf{p} + \zeta \mathbf{q} + \eta \mathbf{e})$ and $A_{KM}(\xi \mathbf{p} + \zeta \mathbf{q} + \eta \mathbf{e})$ are homogeneous polynomials of the argument $(\xi \mathbf{p} + \zeta \mathbf{q} + \eta \mathbf{e})$ with order 8 and 6, respectively. If we let $L \rightarrow \infty$, on surfaces S_3, S_4, S_5 and S_6 , $\frac{A_{KM}(\xi \mathbf{p} + \zeta \mathbf{q} + \eta \mathbf{e})}{D(\xi \mathbf{p} + \zeta \mathbf{q} + \eta \mathbf{e})}$ approaches zero with the speed $1/L^2$. Thus, for $L \rightarrow \infty$, Eq. (2.6.160) gives

$$G_{KM}(\mathbf{x}) = \frac{1}{4\pi^2} \int_{S_1} \frac{A_{KM}(\xi \mathbf{p} + \zeta \mathbf{q} + \eta \mathbf{e})}{D(\xi \mathbf{p} + \zeta \mathbf{q} + \eta \mathbf{e})} \delta(r\eta) dS_1(\xi, \zeta, \eta). \quad (2.6.161)$$

In the deduction of Eq. (2.6.161) the condition that in $\frac{A_{KM}(\xi \mathbf{p} + \zeta \mathbf{q} + \eta \mathbf{e})}{D(\xi \mathbf{p} + \zeta \mathbf{q} + \eta \mathbf{e})}$ only even power terms of ξ exist, is used. Substituting the equation for surface S_1 into Eq. (2.6.161), we have

$$G_{KM}(\mathbf{x}) = \frac{1}{4\pi^2} \int_0^\infty \int_0^\infty \frac{A_{KM}(\mathbf{p} + \zeta \mathbf{q} + \eta \mathbf{e})}{D(\mathbf{p} + \zeta \mathbf{q} + \eta \mathbf{e})} \delta(r\eta) d\zeta d\eta. \quad (2.6.162)$$

The integration with respect to η in Eq. (2.6.162) can easily be treated, which gives

$$G_{KM}(\mathbf{x}) = \frac{1}{4\pi^2} \int_0^\infty \frac{A_{KM}(\mathbf{p} + \zeta \mathbf{q})}{D(\mathbf{p} + \zeta \mathbf{q})} d\zeta. \quad (2.6.163)$$

To evaluate the integral Eq. (2.6.163) by residue calculus, we need to know the poles located at the roots of $D(\mathbf{p} + \zeta \mathbf{q})$. Since the inverse of $\mathbf{\Gamma}(\mathbf{n})$ exists (Dunn, 1994), $D(\mathbf{p} + \zeta \mathbf{q}) = 0$ does not have real roots. Thus, four pairs of conjugate complex roots exist for the equation $D(\mathbf{p} + \zeta \mathbf{q}) = 0$, which gives

$$D(\mathbf{p} + \zeta \mathbf{q}) = a_8 \sum_{m=1}^8 (\zeta - \zeta_m)(\zeta - \zeta_m^*), \quad (2.6.164)$$

where a_8 is the coefficient of ζ^8 in the expanded expression of $D(\mathbf{p} + \zeta \mathbf{q})$, and

$$\text{Im}(\zeta_m) > 0 \quad (m = 1, 2, 3, 4), \quad (2.6.165)$$

ζ_m^* is the conjugate of ζ_m . Thus the Green's function is obtained as

$$G_{KM}(\mathbf{x}) = -\frac{1}{2\pi r} \text{Im} \left(\sum_{i=1}^4 \frac{A_{KM}(\mathbf{p} + \zeta_i \mathbf{q})}{a_8 (\zeta_i - \zeta_i^*) \prod_{j=1, j \neq i}^4 (\zeta_i - \zeta_j)(\zeta_i - \zeta_j^*)} \right). \quad (2.6.166)$$

When the roots for the equation $D(\mathbf{p} + \zeta \mathbf{q}) = 0$ are obtained, Eq. (2.6.166) provides the explicit analytical expression of the Green's function.

2.7 Remarks

In Chapter 2, two-dimensional and three-dimensional problems for a piezoelectric body with heterostructures are treated independently. Intrinsically, the three-dimensional problem describes a more general situation, from which the two-dimensional problem can be derived. Yet, due to its mathematic elegance and technical usefulness, the Stroh formalism (or the alternative Lekhnitskii formalism) has always been the first choice for analyzing two-dimensional problems. Through using two-dimensional methods, it is possible to obtain the exact analytical solution for the electroelastic fields for exterior points of the inclusion in a material system with general anisotropy, which is unavailable if we use the solution methods developed for three-dimensional problems.

As introduced at the beginning of this chapter, the basic solutions provide knowledge of the local distribution of the electroelastic fields, and thus reflect the local properties of the materials. They serve as the foundation for analyzing more complicated problems such as interactions between inclusions, cracks and dislocations, etc., which are also widely studied yet not included in detail in this book.

For practical composites, due to the complexity of the microstructures, it is often impossible to obtain the exact solution for the electroelastic fields. For this reason, various micromechanics models are developed, which provides us with knowledge of the overall properties of the composites without actually solving the distribution of electroelastic fields inside them. These micromechanics models are introduced in Chapter 3 for piezoelectric and ferroelectric materials. The basic

solutions introduced in Chapter 2 are vital for these micromechanics models, since they allow the models to incorporate information of the local fields, which greatly enhances the validity of the results.

References

- [1] Barnett, D.M., Lothe, J.: Dislocations and line charges in anisotropic piezoelectric insulators. *Phys. Status Solidi B* **67**, 105-111 (1975)
- [2] Chen, T.Y.: Green's functions and the non-uniform transformation problem in a piezoelectric medium. *Mech. Res. Commun.* **20**, 271-278 (1993)
- [3] Chen, T.Y., Lin, F.Z.: Numerical evaluation of derivatives of the anisotropic piezoelectric Green's functions. *Mech. Res. Commun.* **20**, 501-506 (1993)
- [4] Chen, W.Q.: On the general solution for piezothermoelasticity for transverse isotropy with application. *J. App. Mech.* **67**, 705-711 (2000)
- [5] Chen, W.Q., Lim, C.W., Ding, H.J.: Point temperature solution for a penny-shaped crack in an infinite transversely isotropic thermo-piezo-elastic medium. *Eng. Anal. Bound. Elem.* **29**, 524-532 (2005)
- [6] Chung, M.Y.: Two-dimensional problems of piezoelectric solids with an elliptic inclusion. Ph.D. Dissertation, University of Illinois at Chicago (1995)
- [7] Chung, M.Y., Ting, T.C.T.: Line force, charge and dislocation in angularly inhomogeneous anisotropic piezoelectric wedges and spaces. *Philos. Maga. A* **71**, 1335-1343 (1995)
- [8] Chung, M.Y., Ting, T.C.T.: Piezoelectric solid with an elliptic inclusion or hole. *Int. J. Solids Struct.* **33**, 3343-3361 (1996)
- [9] Courant, R., Hilbert, D.: *Methods of Mathematical Physics*. John Wiley, New York (1953)
- [10] Deeg, W.F.J.: The analysis of dislocation, crack, and inclusion problems in piezoelectric solids. Ph.D. Thesis, Stanford University (1980)
- [11] Ding, H.J., Chen, W.: *Three Dimensional Problems of Piezoelectricity*. Nova Science Publishers, Huntington, New York (2001)
- [12] Ding, H.J., Liang, J., Chen, B.: Fundamental solutions for transversely isotropic piezoelectric media. *Sci. China Ser. A* **39**, 766-775 (1996)
- [13] Ding, H.J., Wang, G.Q., Chen, W.Q.: Fundamental solutions for plane problem of piezoelectric materials. *Sci. China Ser. E* **40**, 331-336 (1997a)
- [14] Ding, H.J., Wang, G.Q., Chen, W.Q.: Green's functions for a two-phase infinite piezoelectric plane. *Proc. R. Soc. Lond. A* **453**, 2241-2257 (1997b)
- [15] Ding, H.J., Chen, B., Liang, J.: On the Green's functions for two-phase transversely isotropic piezoelectric media. *Int. J. Solids Struct.* **34**, 3041-3057 (1997c)
- [16] Du, S.Y., Liang, J., Han, J.C., Wang, B.: General coupled solution of anisotropic piezoelectric materials with an elliptic inclusion. *Acta Mech. Sin.*

- 10**, 273-281 (1994)
- [17] Dunn, M.L.: Electroelastic Green's functions for transversely isotropic piezoelectric media and their application to the solution of inclusion and inhomogeneity problems. *Int. J. Eng. Sci.* **32**, 119-131 (1994)
 - [18] Dunn, M.L., Wienecke, H.A.: Green's functions for transversely isotropic piezoelectric solids. *Int. J. Solids Struct.* **33**, 4571-4581 (1996)
 - [19] Dunn, M.L., Wienecke, H.A.: Half-space Green's functions for transversely isotropic piezoelectric solids. *J. App. Mech.* **66**, 675-679 (1999)
 - [20] Eringen, A.C.: Non-local elasticity and waves. In: Thoft-Christensen, P. (ed.) *Continuum Mechanics Aspects of Geodynamics and Rock Fracture Mechanics*, pp. 81-105. Kluwer Academic Publishers, Dordrecht, Holland (1974)
 - [21] Eringen, A.C.: Continuum mechanics at the atomic scale. *Cryst. Lattice Defects* **7**, 109-130 (1977)
 - [22] Eringen, A.C.: Line crack subject to shear. *Int. J. Fract.* **14**, 367-379 (1978)
 - [23] Eringen, A.C.: Line crack subject to antiplane shear. *Eng. Fract. Mech.* **12**, 211-219 (1979)
 - [24] Eringen, A.C.: Interaction of a dislocation with a crack. *J. Appl. Phys.* **54**, 6811-6817 (1983)
 - [25] Eringen, A.C., Speziale, C.G., Kim, B.S.: Crack tip problem in non-local elasticity. *J. Mech. Phys. Solids* **25**, 339-355 (1977)
 - [26] Eshelby, J.D.: The force on an elastic singularity. *Phil. Trans. R. Soc. A* **244**, 87-112 (1951)
 - [27] Eshelby, J.D.: The determination of the elastic field of an ellipsoidal inclusion, and related problems. *Proc. R. Soc. Lond. A* **241**, 376-396 (1957)
 - [28] Eshelby, J.D.: The elastic field outside an ellipsoidal inclusion. *Proc. R. Soc. Lond. A* **252**, 561-569 (1959)
 - [29] Fang, D.N., Liu, J.X.: *Fracture Mechanics of Piezoelectric and Ferroelectric Solids*. Springer, Berlin (2009)
 - [30] Fu, R., Zhang, T.Y.: Effects of an electric field on the fracture toughness of poled lead zirconate titanate ceramics. *J. Am. Ceram. Soc.* **83**, 1215-1218 (2000)
 - [31] Gao, H.J., Zhang, T.Y., Tong, P.: Local and global energy release rates for an electrically yielded crack in a piezoelectric ceramic. *J. Mech. Phys. Solids* **45**, 491-510 (1997)
 - [32] Helgason, S.: *The Radon Transform*. Birkhauser, Boston (1980)
 - [33] Heyer, V., Schneider, G.A., Balke, H., Drescher, J., Bahr, H.A.: A fracture criterion for conducting cracks in homogeneously poled piezoelectric PZT-PIC151 ceramics. *Acta Mater.* **46**, 6615-6622 (1998)
 - [34] Hou, P.F., Leung, A.Y.T.: Three-dimensional Green's functions for two-phase transversely isotropic piezothermoelastic media. *J. Intel. Mater. Syst. Struct.* **20**, 11-21 (2009)
 - [35] Hou, P.F., Zhou, X.H., He, Y.J.: Green's functions for a semi-infinite

- transversely isotropic piezothermoelastic material. *Smart Mater. Struct.* **16**, 1915-1923 (2007)
- [36] IEEE Standards Board and American National Standards Institute: IEEE Standard on Piezoelectricity, an American National Standard. ANSI/IEEE Std 176-1987 (1987)
 - [37] Kuna, M.: Fracture mechanics of piezoelectric materials—Where are we right now? *Eng. Fract. Mech.* **77**, 309-326 (2010)
 - [38] Landau, L.D., Lifshitz, E.M.: *Electrodynamics of Continuous Media*. Pergamon Press, Oxford (1960)
 - [39] Lekhnitskii, S.G.: *Anisotropic Plates*. Gordon & Breach, New York (1968)
 - [40] Lekhnitskii, S.G.: *Theory of Elasticity of an Anisotropic Body*. MIR Publishers, Moscow (1981)
 - [41] Li, X.Y., Wang, M.Z.: Three-dimensional Green's functions for infinite anisotropic piezoelectric media. *Int. J. Solids Struct.* **44**, 1680-1684 (2007)
 - [42] Liang, J., Han, J.C., Wang, B., Du, S.Y.: Electroelastic modelling of anisotropic piezoelectric materials with an elliptic inclusion. *Int. J. Solids Struct.* **32**, 2989-3000 (1995)
 - [43] Liang, K.M., Liu, F., Miu, G.Q.: *Methods of Mathematical Physics*. Higher Education Press, Beijing, China (1998) (In Chinese)
 - [44] Lines, M.E., Glass, A.M.: *Principles and Applications of Ferroelectric and Related Materials*. Oxford University Press, New York (1977)
 - [45] Liu, J.X., Wang, B., Du, S.Y.: Two-dimensional electroelastic fundamental solutions for general anisotropic piezoelectric media. *Appl. Math. Mech.* **18**, 949-956 (1997a)
 - [46] Liu, J.X., Wang, B., Du, S.Y.: Electro-elastic fundamental solutions of anisotropic piezoelectric materials with an elliptical hole. *Acta Mech. Sin. (English Series)* **13**, 54-62 (1997b)
 - [47] Liu, J.X., Wang, B., Du, S.Y.: Electro-elastic Green's functions for a piezoelectric half-space and their application. *Appl. Math. Mech.* **18**, 1037-1043 (1997c)
 - [48] Lothe, J., Barnett, D.M.: Integral formalism for surface waves in piezoelectric crystals. Existence considerations. *J. Appl. Phys.* **47**, 1799-1807 (1976)
 - [49] Lu, P., Lee, H.P., Lu, C.: Exact solutions for simply supported functionally graded piezoelectric laminates by Stroh-like formalism. *Compos. Struct.* **72**, 352-363 (2006)
 - [50] McMeeking, R.M.: Electrostrictive stresses near crack-like flaws. *J. Appl. Math. Phys.* **40**, 615-627 (1989)
 - [51] Mulier, S., Debois, V., Cohen, S.S., Williams, D.R., Gao, C.F., Fan, W.X.: Green's functions for the plane problem in a half-infinite piezoelectric medium. *Mech. Res. Commun.* **25**, 69-74 (1998)
 - [52] Mura, T.: *Micromechanics of Defects in Solids*, 2nd ed. Martinus Nijhoff, Dordrecht (1987)

- [53] Pak, Y.E.: Crack extension force in a piezoelectric material. *J. Appl. Mech.* **57**, 647-653 (1990a)
- [54] Pak, Y.E.: Force on a piezoelectric screw dislocation. *J. Appl. Mech.* **57**, 863-869 (1990b)
- [55] Pak, Y.E.: Linear electro-elastic fracture mechanics of piezoelectric materials. *Int. J. Fract.* **54**, 79-100 (1992)
- [56] Pan, E., Tonon, F.: Three dimensional Green's functions in anisotropic piezoelectric solids. *Int. J. Solids Struct.* **37**, 943-958 (2000)
- [57] Pan, E., Yuan, F.G.: Three-dimensional Green's functions in anisotropic piezoelectric bimaterials. *Int. J. Eng. Sci.* **38**, 1939-1960 (2000)
- [58] Park, S.B., Sun, C.T.: Fracture criteria for piezoelectric ceramics. *J. Am. Ceram. Soc.* **78**, 1475-1480 (1995)
- [59] Qin, Q.H.: Thermoelastoelectric Green's function for a piezoelectric plate containing an elliptic hole. *Mech. Mater.* **30**, 21-29 (1998)
- [60] Qin, Q.H.: Green's function for thermopiezoelectric materials with holes of various shapes. *Arch. Appl. Mech.* **69**, 406-418 (1999a)
- [61] Qin, Q.H.: Thermoelastoelectric Green's function for thermal load inside or on the boundary of an elliptic inclusion. *Mech. Mater.* **31**, 611-626 (1999b)
- [62] Qin, Q.H., Yu, S.W.: An arbitrarily-oriented plane crack terminating at the interface between dissimilar piezoelectric materials. *Int. J. Solids Struct.* **34**, 581-590 (1997)
- [63] Qin, Q.H., Mai, Y.W., Yu, S.W.: Some problems in plane thermopiezoelectric materials with holes. *Int. J. Solids Struct.* **36**, 427-439 (1999)
- [64] Rice, J.R.: A path independent integral and the approximate analysis of strain concentration by notches and cracks. *J. Appl. Mech.* **35**, 379-386 (1968)
- [65] Smith, R.C.: *Smart Material Systems: Model Development*. SIAM, Philadelphia (2005)
- [66] Sosa, H.: Plane problems in piezoelectric media with defects. *Int. J. Solids Struct.* **28**, 491-505 (1991)
- [67] Sosa, H.: On the fracture mechanics of piezoelectric solids. *Int. J. Solids Struct.* **29**, 2613-2622 (1992)
- [68] Stroh, A.N.: Dislocations and cracks in anisotropic elasticity. *Phil. Mag.* **3**, 625-646 (1958)
- [69] Stroh, A.N.: Steady-state problems in anisotropic elasticity. *J. Math. Phys.* **41**, 77-103 (1962)
- [70] Sun, K., Zhang, X.F.: *Piezoelectricity*. National Defence Industrial Press, Beijing, China (1984) (In Chinese)
- [71] Suo, Z., Kuo, C.M., Barnett, D.M., Willis, J.R.: Fracture mechanics for piezoelectric ceramics. *J. Mech. Phys. Solids* **40**, 739-765 (1992)
- [72] Tiersten, H.F.: *Linear Piezoelectric Plate Vibrations*. Plenum Press, New York (1969)
- [73] Ting, T.C.T.: Some identities and the structure of N_i in the Stroh formalism of anisotropic elasticity. *Q. Appl. Math.* **46**, 109-120 (1988)

- [74] Ting, T.C.T.: Barnett-Lothe tensors and their associated tensors for monoclinic materials with the symmetry plane at $x_3 = 0$. *J. Elast.* **27**, 143-165 (1992)
- [75] Ting, T.C.T.: On anisotropic elastic materials that possess three identical Stroh eigenvalues as do isotropic materials. *Q. Appl. Math.* **52**, 363-375 (1994)
- [76] Ting, T.C.T.: *Anisotropic Elasticity: Theory and Applications*. Oxford University Press, USA (1996)
- [77] Ting, T.C.T.: A modified Lekhnitskii formalism a la Stroh for anisotropic elasticity and classifications of the 6×6 matrix N . *Proc. R. Soc. Lond. A* **455**, 69-89 (1999)
- [78] Ting, T.C.T., Yan, G.P.: The anisotropic elastic solid with an elliptic hole or rigid inclusion. *Int. J. Solids Struct.* **27**, 1879-1894 (1991)
- [79] Wang, B.: Three-dimensional analysis of an ellipsoidal inclusion in a piezoelectric material. *Int. J. Solids Struct.* **29**, 293-308 (1992a)
- [80] Wang, B.: Three-dimensional analysis of a flat elliptical crack in a piezoelectric material. *Int. J. Eng. Sci.* **30**, 781-791 (1992b)
- [81] Wang, B.: Effective behaviors of piezoelectric composites. *Appl. Mech. Rev.* **47**, S112-S121 (1994)
- [82] Wang, B., Du, S.Y.: Some explicit expressions for a spheroidal inclusion in a transversely isotropic, piezoelectric matrix. *Int. J. Appl. Electromagn. Mater.* **3**, 289-295 (1993)
- [83] Wang, C.Y.: Elastic fields produced by a point source in solids of general anisotropy. *J. Eng. Math.* **32**, 41-52 (1997)
- [84] Wang, C.Y., Achenbach, J.D.: Three-dimensional time-harmonic elastodynamic Green's functions for anisotropic solids. *Proc. R. Soc. Lond. A* **449**, 441-458 (1995)
- [85] Wang, H.Y., Singh, R.N.: Crack propagation in piezoelectric ceramics: Effect of applied electric fields. *J. Appl. Phys.* **81**, 7471-7479 (1997)
- [86] Willis, J.R.: Boussinesq problems for an anisotropic half-space. *J. Mech. Phys. Solids* **15**, 331-339 (1967)
- [87] Willis, J.R.: The stress field around an elliptical crack in an anisotropic elastic medium. *Int. J. Eng. Sci.* **6**, 253-263 (1968)
- [88] Xu, X.L., Rajapakse, R.K.N.D.: Analytical solution for an arbitrarily oriented void/crack and fracture of piezoceramics. *Acta Mater.* **47**, 1735-1747 (1999)
- [89] Xu, X.L., Rajapakse, R.K.N.D.: A theoretical study of branched cracks in piezoelectrics. *Acta Mater.* **48**, 1865-1882 (2000)
- [90] Zhou, Z.G., Wang, B.: Investigation of anti-plane shear behavior of two collinear impermeable cracks in the piezoelectric materials by using the non-local theory. *Int. J. Solids Struct.* **39**, 1731-1742 (2002)
- [91] Zhou, Z.G., Wang, B.: Nonlocal theory solution of two collinear cracks in the functionally graded materials. *Int. J. Solids Struct.* **43**, 887-898 (2006)
- [92] Zhou, Z.G., Han, J.C., Du, S.Y.: Investigation of a Griffith crack subject to

- anti-plane shear by using the non-local theory. *Int. J. Solids Struct.* **36**, 3891-3901 (1999)
- [93] Zhou, Z.G., Wang, B., Du, S.Y.: Analysis of the dynamic behavior of two parallel symmetric cracks using the non-local theory. *Int. J. Eng. Sci.* **40**, 1023-1035 (2002a)
- [94] Zhou, Z.G., Wang, B., Du, S.Y.: Investigation of antiplane shear behavior of two collinear permeable cracks in a piezoelectric material by using the nonlocal theory. *J. Appl. Mech.* **69**, 388-390 (2002b)
- [95] Zhou, Z.G., Sun, J.L., Wang, B.: Investigation of the behavior of a crack in a piezoelectric material subjected to a uniform tension loading by use of the non-local theory. *Int. J. Eng. Sci.* **42**, 2041-2063 (2004)

Micromechanics Models of Piezoelectric and Ferroelectric Composites

Due to the complicated heterostructures in actual materials and composites, it is extremely hard to obtain the exact solution for the electroelastic fields in them. Therefore, various kinds of micromechanics models for piezoelectric and ferroelectric composites are introduced, aimed at estimating the macroscopic mechanical and electrical properties of bulk piezoelectric and ferroelectric composites based on their microstructures. It should be emphasized that since the exact solution for the electroelastic fields is generally not available, the results obtained by these models are based on approximate solutions. In general, these models are based on the assumption that the distribution of heterostructures (inclusions, or inhomogeneities) is statistically uniform, so that the macroscopic composite can be concerned as a repetition of some “representative volume”, which is small relative to the whole bulk composite but large relative to the microscale constituents. And the representative volume is carefully chosen so that it correctly reflects the overall characterization of the composite. The solution for the electroelastic fields of the representative volume is related to the basic solutions given in Chapter 2.

Chapter 3 is arranged as follows: in Sections 3.1 and 3.2, the background information and some definitions are introduced; in Section 3.3, various kinds of micromechanics models for predicting the effective material properties of piezoelectric materials with heterostructures are introduced, together with the results obtained by using these models; in Section 3.4, the thermodynamic approach is used to incorporate the nonlinear electric properties and linear mechanical properties of ferroelectric materials; in Section 3.5, phase diagram analysis for ferroelectric composites is presented. Some remarks on Chapter 3 are provided in Section 3.6.

3.1 Background

Micromechanics models were developed independent of the research on piezoelectric and ferroelectric composites (e.g. Mori and Tanaka, 1973; Weng, 1984; Benveniste, 1987; Voigt, 1889; Reuss, 1929), and their application to elastic composites was summarized in (Du and Wang, 1998).

Early applications of micromechanics models to the investigation of piezoelectric composites included Newnham's connectivity theory (Newnham et al., 1978), Banno's extension (Banno, 1983) of Newnham's theory, and Chan and Unsworth (1989), Smith and Auld (1991), etc. These approaches share one similar assumption, namely that the stress or strain components are constants in the composites, which derives from Voigt (1889) and Reuss (1929). Then Hashin/Shtrikman-type bounds for describing the overall behavior have been developed (Bisegna and Luciano, 1996; 1997). These models provide the upper and lower bounds for the effective constants of piezoelectric composites. More rigorous attempts included Grekov et al. (1987a; 1987b) and Zhou (1991).

Wang (1992a) was the first to incorporate the rigorous solution for the coupled electroelastic fields in an ellipsoidal shaped piezoelectric inhomogeneity in the analysis of the effective behavior of composites. Yet the effective behavior was solved based on the dilute models, which is valid when the volume fraction of the inclusions is small. A most comprehensive extension of various kinds of micromechanics models to the study of piezoelectric composites was done by Dunn and Taya (1993c). These micromechanics models include the dilute, self-consistent, Mori-Tanaka and differential schemes. These models, before application to piezoelectric composites, were already widely used in ordinary composites by a lot of scholars (Mura, 1987; Weng, 1990; 1992; Qin and Weng, 1990; Benveniste et al., 1991; Taya and Arsenault, 1989). Other relevant works include Dunn and Taya (1993a; 1993b), Dunn (1995), Marco and Pieter (1998), Jiang et al. (2001), Jiang and Cheung (2001), Harald et al. (2005), Paolo and Raimondo (1996), Chen (1993; 1994), Huang and Kuo (1996) and Wu (2000).

The essence of these micromechanics models is the use of volume average physical quantities instead of their exact distributions. Unsatisfied with the mean field assumption, a lot of efforts are made by researchers regarding the exact distribution of the electroelastic fields by using the finite element method (FEM) (Sabina et al., 2001; Teply and Dvorak, 1988; Poizat and Sester, 1999; Smith et al., 1998; Böhm, 1993; Brockenbrough and Suresh, 1990; Cleveringa et al., 1997; Gunawarkena et al., 1993). By using the FEM approach, the unit cell of the piezoelectric composite is meshed and calculated. The results are then used to determine the complete tensors associated with the overall elastic, dielectric and piezoelectric behavior. Besides the FEM approach, the asymptotic homogenization approach (Suquet, 1989; Bakhvalov and Panasenko, 1987) is an alternative option for solving the effective properties without using the mean field assumption. The main restriction of these methods is that they can be used only to analyze piezoelectric composites with periodically distributed inclusions.

In the linear theories for piezoelectric materials, the electric properties for the materials are assumed to be linear. Yet, generally speaking, many piezoelectric materials are intrinsically ferroelectric, which implies a nonlinear relation between the polarization (or electric displacement) and the applied electric field. Chen et al. (1997), Nan and Clarke (1997), Li and Weng (1999) are among the scholars who first tried to use the micromechanics models to analyze the nonlinear behavior of the ferroelectric materials. Chen et al. (1997) tried to extend Dunn and Taya's method to treat the nonlinear switching behavior of ferroelectric materials caused by microstructure changes. Li and Weng (1999) sought to use the micromechanics models to study the physics of the domain switch and the coupled electromechanical properties of ferroelectric materials based on thermodynamics approaches. Further research concerning ferroelectric materials includes Lu et al. (1999), Huber and Fleck (2004), Li and Weng (2001; 2004), Huber (2005), Su and Weng (2006), Dommelen et al. (2004), Fulton and Gao (2001), Chen et al. (2002), Erhart and Cao (1999) and Wong et al. (2001). Much focus was put on the domain switching problems of ferroelectric materials. Yet a comprehensive understanding of the behavior of ferroelectric composites, which incorporates the effect of microstructures and the nonlinear electric property, is still lacking. In Section 3.3, based on the mean field theories, the energy formulation of ferroelectric composites is constructed, which incorporates the exact solution for the electroelastic fields in microstructures. In Section 3.4, phase diagram analysis is presented using energy formulation.

3.2 Some Definitions

We start our discussion with the following definitions of the volume averages of the elastic and electric quantities.

Notations for volume average physical quantities:

For space R occupied by bulk piezoelectric composite, denote $\langle f \rangle_R$ as the volume average of the physical quantity f over R , and we have

$$\langle \sigma_{ij} \rangle_R = \frac{1}{V_R} \int_R \sigma_{ij} dR, \quad (3.2.1)$$

$$\langle \varepsilon_{kl} \rangle_R = \frac{1}{V_R} \int_R \varepsilon_{kl} dR, \quad (3.2.2)$$

where V_R is the volume of the domain R and

$$\sigma_{i4} = D_i, \quad (3.2.3)$$

$$\varepsilon_{4l} = -E_l. \quad (3.2.4)$$

• *Volume average of the generalized stresses and strains (stress boundaries)*

The volume average of the generalized stresses $\langle \sigma_{ij} \rangle$ and strains $\langle \varepsilon_{ij} \rangle$ possesses some general properties, which can be extended from the pure elastic case. By denoting the material constants for the matrix by C_{ijkl} , and assuming that on the boundary of R , i.e. ∂R , the following boundary conditions are prescribed:

$$\sigma_{ij} = \sigma_{ij}^0, \quad (3.2.5)$$

where σ_{ij}^0 are given on the surface as constants, we have for Eq. (3.2.1)

$$\langle \sigma_{ij} \rangle_R = \frac{1}{V_R} \int_R \sigma_{ij} dR = \frac{1}{V_R} \int_R \sigma_{pj} \delta_{ip} dR = \frac{1}{V_R} \int_R \sigma_{pj} x_{i,p} dR, \quad (3.2.6)$$

where the relation $\delta_{ik} = x_{i,k}$ is used. Integration by part of the right hand side of Eq. (3.2.6) yields

$$\langle \sigma_{ij} \rangle_R = \frac{1}{V_R} \left(\int_{\partial R} \sigma_{pj} x_i n_p dS - \int_R \sigma_{pj,p} x_i dR \right). \quad (3.2.7)$$

For the second term in Eq. (3.2.7), noticing that for static problems $\sigma_{ij,i} = 0$ and $D_{i,i} = 0$, substitution of Eq. (3.2.5) into Eq. (3.2.7) leads to

$$\langle \sigma_{ij} \rangle_R = \sigma_{pj}^0 \frac{1}{V_R} \int_{\partial R} x_i n_p dS = \sigma_{pj}^0 \delta_{ip} = \sigma_{ij}^0. \quad (3.2.8)$$

Notice that Eq. (3.2.8) is valid without specific requirement of the material constants or the microstructure in R .

For the generalized strains, extension is made on the proposition of including the effect of generalized eigenstrains here. Suppose the generalized eigenstrain ε_{kl}^* is prescribed to an aggregate of inclusions Ω , giving rise to generalized stresses σ_{ij} and generalized strains ε_{kl} . $\varepsilon_{4l}^* = -E_l^*$ denotes the nonlinear part of electric fields. The volume average of the generalized strains ε_{kl} can be calculated by

$$\langle \varepsilon_{kl} \rangle_R = \frac{1}{V_R} \int_R \varepsilon_{kl} dR = \frac{1}{V_R} \int_R (e_{kl} + \varepsilon_{kl}^*) dR, \quad (3.2.9)$$

where e_{kl} is the generalized elastic strain. Eq. (3.2.9) can be transformed as

$$\langle \varepsilon_{kl} \rangle_R = \frac{1}{V_R} \left(\int_R C_{ijkl}^{-1} \sigma_{ij} dR + \int_R \varepsilon_{kl}^* dR \right) = C_{ijkl}^{-1} \sigma_{ij}^0 + \frac{1}{V_R} \int_R \varepsilon_{kl}^* dR, \quad (3.2.10)$$

where Eq. (3.2.8) is used. Since ε_{kl}^* is prescribed to Ω only, we have

$$\langle \varepsilon_{kl} \rangle_R = C_{ijkl}^{-1} \sigma_{ij}^0 + \frac{V_\Omega}{V_R} \langle \varepsilon_{kl}^* \rangle_\Omega. \quad (3.2.11)$$

For inhomogeneous inclusions, the generalized strains satisfy

$$\varepsilon_{kl} = e_{kl} + \varepsilon_{kl}^* + \varepsilon_{kl}^{**}, \quad (3.2.12)$$

where ε_{kl}^* is the prescribed eigenstrain in Ω and ε_{kl}^{**} is the equivalent eigenstrain defined by Eq. (2.5.14). Eq. (3.2.11) becomes

$$\langle \varepsilon_{kl} \rangle_R = C_{ijkl}^{-1} \sigma_{ij}^0 + \frac{V_\Omega}{V_R} \langle \varepsilon_{kl}^* + \varepsilon_{kl}^{**} \rangle_\Omega, \quad (3.2.13)$$

or can be rewritten for inhomogeneities according to Eq. (2.5.20) as

$$\langle \varepsilon_{kl} \rangle_R = C_{ijkl}^{-1} \sigma_{ij}^0 + c_j C_{ijkl}^{-1} \Delta C_{ijmn} \langle \varepsilon_{kl}^I \rangle_\Omega. \quad (3.2.14)$$

One should bear in mind that Eqs. (3.2.8), (3.2.11), (3.2.13) and (3.2.14) are obtained by using boundary conditions of the generalized stresses. If otherwise, the deformation is restricted on the boundary and the conclusions should be reconsidered.

- *Volume average of the generalized stresses and strains (displacement boundaries)*

If the boundary conditions on ∂R are given by restricted displacements, this means that (Benveniste and Dvorak, 1992)

$$u_i(S) = \varepsilon_{ij}^0 x_j, \quad u_4(S) = \phi(S) = \varepsilon_{4j}^0 x_j = -E_j^0 x_j. \quad (3.2.15)$$

According to the definition in Eq. (3.2.15), the volume average of the generalized strains satisfies

$$\langle \varepsilon_{kl} \rangle_R = \varepsilon_{kl}^0. \quad (3.2.16)$$

For inhomogeneous inclusions Eq. (3.2.16) is still valid.

For the volume average of the generalized stresses, Eq. (3.2.7) yields

$$\langle \sigma_{ij} \rangle_R = \frac{1}{V_R} \int_{\partial R} \sigma_{pj} x_i n_p dS, \quad (3.2.17)$$

which permits no further deduction. Yet, on the other hand, for a two-phase composite system we can write

$$\langle \sigma_{ij} \rangle_R = (1 - c_f) C_{ijkl} \langle \varepsilon_{kl}^M \rangle_{R-\Omega} + c_f C_{ijkl}^* \langle \varepsilon_{kl}^I \rangle_{\Omega}, \quad (3.2.18)$$

where C_{ijkl}^* is the material constant for the inhomogeneity phase and C_{ijkl} is the material constant for the matrix phase. Eq. (3.2.18) can be transformed as

$$\langle \sigma_{ij} \rangle_R = C_{ijkl} \varepsilon_{kl}^0 + c_f \Delta C_{ijkl} \langle \varepsilon_{kl}^I \rangle_{\Omega}. \quad (3.2.19)$$

• *Definition of effective constants of piezoelectric composites*

For a two-phase piezoelectric composite, where an aggregate of inclusions Ω is embedded in the matrix $R - \Omega$, the volume average of the generalized stresses and strains can be rewritten as

$$\begin{aligned} \langle \sigma_{ij} \rangle_R &= \frac{1}{V_R} \left[\int_{R-\Omega} \sigma_{ij}^M dv + \int_{\Omega} \sigma_{ij}^I dv \right] \\ &= (1 - c_f) \langle \sigma_{ij}^M \rangle_{R-\Omega} + c_f \langle \sigma_{ij}^I \rangle_{\Omega}, \end{aligned} \quad (3.2.20)$$

$$\langle \varepsilon_{kl} \rangle_R = (1 - c_f) \langle \varepsilon_{kl}^M \rangle_{R-\Omega} + c_f \langle \varepsilon_{kl}^I \rangle_{\Omega}, \quad (3.2.21)$$

where terms with superscript M denote physical quantities of the matrix phase and terms with superscript I denote physical quantities of the inhomogeneity phase, and

$$c_f = \frac{V_{\Omega}}{V_R}. \quad (3.2.22)$$

The effective constants of piezoelectric composites are defined by

$$\langle \sigma_{ij} \rangle_R = \bar{C}_{ijkl} \langle \varepsilon_{kl} \rangle_R, \quad (3.2.23)$$

or

$$\langle \varepsilon_{kl} \rangle_R = \bar{S}_{ijkl} \langle \sigma_{ij} \rangle_R. \quad (3.2.24)$$

From the definition Eq. (3.2.23) and Eq. (3.2.24), it is clear that the effective material constants linearly related the volume averaged generalized stresses with the volume averaged generalized strains. From the above deduction of the general

properties of the volume average quantities, the expressions for these effective material constants cannot be obtained. In the following section, several assumptions are introduced which leads to explicit expressions of the effective material constants.

3.3 Effective Material Constants of Piezoelectric Composites

Dunn and Taya (1993c) provided various micromechanics models for predicting the effective material constants of piezoelectric composites. The methods were based on linear constitutive equations and aimed at composites with inhomogeneities. For boundary conditions given by generalized strains, we have for Eqs. (3.2.16) and (3.2.19)

$$\langle \varepsilon_{kl} \rangle_R = \varepsilon_{kl}^0, \quad (3.3.1)$$

$$\langle \sigma_{ij} \rangle_R = C_{ijkl} \varepsilon_{kl}^0 - c_f (C_{ijkl} - C_{ijkl}^*) \langle \varepsilon_{kl}^I \rangle_\Omega, \quad (3.3.2)$$

if

$$\langle \varepsilon_{kl}^I \rangle_\Omega = A_{lKMn} \varepsilon_{Mn}^0. \quad (3.3.3)$$

From Eqs. (3.3.1), (3.3.2) and (3.3.3) we have

$$\bar{C}_{ijMn} = C_{ijMn} - c_f (C_{ijkl} - C_{ijkl}^*) A_{lKMn}. \quad (3.3.4)$$

The expression of Eq. (3.3.4) in matrix notation is given by

$$\bar{\mathbf{C}} = \mathbf{C} - c_f (\mathbf{C} - \mathbf{C}^*) \mathbf{A}. \quad (3.3.5)$$

If $\mathbf{A}$ is known, $\bar{\mathbf{C}}$ can be obtained from Eq. (3.3.5).

3.3.1 The Dilute Model

According to Eq. (3.3.3), $\mathbf{A}$ relates the average generalized strains inside the inhomogeneities phase to the given generalized strains at the boundary. For a simple and direct approximation, let the actual generalized strains of the inhomogeneities phase be replaced by the fields of one inhomogeneity embedded in the infinite matrix; the solution is then provided in Section 2.6.1. In this case, the generalized strains are proved to be constants for interior points of the inhomogeneity. Eq. (3.3.3) is transformed to matrix notation as

$$\langle \boldsymbol{\varepsilon}^I \rangle_{\Omega} = \boldsymbol{\varepsilon}^I = \mathbf{A}^{\text{dil}} \boldsymbol{\varepsilon}^0, \quad (3.3.6)$$

where

$$\mathbf{A}^{\text{dil}} = (\mathbf{I} - \mathbf{S}\mathbf{C}^{-1}(\mathbf{C} - \mathbf{C}^*))^{-1}. \quad (3.3.7)$$

In the deduction of Eq. (3.3.7), Eq. (2.5.2) is used. $\mathbf{S}$ is the 9×9 generalized Eshelby tensor, whose components are given by Eq. (2.5.4). Eq. (3.3.7), together with Eq. (3.3.5), determines the effective material constants of piezoelectric composites. The dilute model ignores the interaction between different inhomogeneities, and thus is valid only when the volume fraction of the inhomogeneities phase $c_f \ll 1$.

For piezoelectric composites with cylindrical-shaped (fiber-reinforced) inhomogeneities that are transversely isotropic, Wang (1992a) gave the explicit expressions of the effective material constants for various mechanical and electric conditions. The results are presented here. The matrix $\mathbf{A}^{\text{dil}}$ is solved from the following set of equations

$$\begin{aligned} \varepsilon_{11}^I = \varepsilon_{11}^0 - \frac{5C_{11} - 3C_{12}}{8C_{11}(C_{11} - C_{12})} (\Delta C_{11}\varepsilon_{11}^I + \Delta C_{12}\varepsilon_{22}^I + \Delta C_{13}\varepsilon_{33}^I) \\ + \frac{C_{11} + C_{12}}{8C_{11}(C_{11} - C_{12})} (\Delta C_{12}\varepsilon_{11}^I + \Delta C_{22}\varepsilon_{22}^I + \Delta C_{23}\varepsilon_{33}^I) + \frac{\Delta e_{31}}{2C_{11}} E_3^I, \end{aligned} \quad (3.3.8)$$

$$\begin{aligned} \varepsilon_{22}^I = \varepsilon_{22}^0 + \frac{C_{11} + C_{12}}{8C_{11}(C_{11} - C_{12})} (\Delta C_{11}\varepsilon_{11}^I + \Delta C_{12}\varepsilon_{22}^I + \Delta C_{13}\varepsilon_{33}^I) \\ - \frac{5C_{11} - 3C_{12}}{8C_{11}(C_{11} - C_{12})} (\Delta C_{12}\varepsilon_{11}^I + \Delta C_{22}\varepsilon_{22}^I + \Delta C_{23}\varepsilon_{33}^I) + \frac{\Delta e_{31}}{2C_{11}} E_3^I, \end{aligned} \quad (3.3.9)$$

$$\varepsilon_{33}^I = \varepsilon_{33}^0, \quad (3.3.10)$$

$$\varepsilon_{23}^I = \frac{4[C_{44}\alpha_{11} + (e_{15})^2]E_{23}^0 + (\alpha_{11}\Delta e_{24} - \Delta\alpha_{22}e_{15})E_2^I}{4[C_{44}\alpha_{11} + (e_{15})^2] + 2e_{15}\Delta e_{24} + 2\alpha_{11}\Delta C_{44}}, \quad (3.3.11)$$

$$\varepsilon_{13}^I = \frac{4[C_{44}\alpha_{11} + (e_{15})^2]E_{13}^0 + (\alpha_{11}\Delta e_{24} - \Delta\alpha_{22}e_{15})E_1^I}{4[C_{44}\alpha_{11} + (e_{15})^2] + 2e_{15}\Delta e_{24} + 2\alpha_{11}\Delta C_{44}}, \quad (3.3.12)$$

$$\varepsilon_{12}^I = [1 + \frac{3C_{11} - C_{12}}{2C_{11}(C_{11} - C_{12})} \Delta C_{66}]^{-1} \varepsilon_{12}^0, \quad (3.3.13)$$

$$E_1^I = \frac{2[C_{44}\alpha_{11} + (e_{15})^2]E_1^0 + 2(e_{15}\Delta C_{44} - \Delta e_{15}C_{44})\varepsilon_{13}^I}{2[C_{44}\alpha_{11} + (e_{15})^2] + e_{15}\Delta e_{15} + \Delta\alpha_{11}C_{44}}, \quad (3.3.14)$$

$$E_2^I = \frac{2[C_{44}\alpha_{11} + (e_{15})^2]E_2^0 + 2(e_{15}\Delta C_{44} - \Delta e_{24}C_{44})\varepsilon_{23}^I}{2[C_{44}\alpha_{11} + (e_{15})^2] + e_{15}\Delta e_{15} + \Delta\alpha_{11}C_{44}}, \quad (3.3.15)$$

$$E_3^I = E_3^0. \quad (3.3.16)$$

For the following specific boundary conditions, the effective material constants are obtained.

(a) Only $\varepsilon_{33}^0 \neq 0$

From Eqs. (3.3.8)–(3.3.10), one obtains

$$\varepsilon_{11}^I = \varepsilon_{22}^I = -\frac{\Delta C_{13}}{2C_{11} + \Delta C_{11} + \Delta C_{12}}\varepsilon_{33}^0. \quad (3.3.17)$$

According to Eq. (3.3.5), we obtain after manipulations

$$\bar{C}_{33} = C_{33} + c_f\Delta C_{33} - \frac{2c_f(\Delta C_{13})^2}{2C_{11} + \Delta C_{11} + \Delta C_{12}}, \quad (3.3.18)$$

$$\bar{C}_{13} = C_{13} + c_f\Delta C_{13} - \frac{c_f\Delta C_{13}(\Delta C_{11} + \Delta C_{12})}{2C_{11} + \Delta C_{11} + \Delta C_{12}}. \quad (3.3.19)$$

(b) $\varepsilon_{11}^0 = \varepsilon_{22}^0 \neq 0$ and $\varepsilon_{12}^0 \neq 0$

From Eqs. (3.3.8)–(3.3.10), one obtains

$$\varepsilon_{11}^I = \varepsilon_{22}^I = \frac{2C_{11}}{2C_{11} + \Delta C_{11} + \Delta C_{12}}\varepsilon_{11}^0. \quad (3.3.20)$$

According to Eq. (3.3.5), we obtain after manipulations

$$\bar{C}_{11} + \bar{C}_{12} = C_{11} + C_{12} + 2c_f(\Delta C_{11} + \Delta C_{12})\frac{C_{11}}{2C_{11} + \Delta C_{11} + \Delta C_{12}}, \quad (3.3.21)$$

and

$$\bar{C}_{11} - \bar{C}_{12} = 2\bar{C}_{66} = 2C_{66} + 2c_f\Delta C_{66}\left[1 + \frac{3C_{11} - C_{12}}{2C_{11}(C_{11} - C_{12})}\Delta C_{66}\right]^{-1}. \quad (3.3.22)$$

From Eqs. (3.3.21) and (3.3.22), $\bar{C}_{11}$ and $\bar{C}_{12}$ can be immediately solved.

(c) Only $\varepsilon_{23}^0 \neq 0$

From Eqs. (3.3.11) and (3.3.15), one obtains

$$\varepsilon_{23}^I = 4[\alpha_{11}C_{44} + (e_{15})^2] \left\{ 4[\alpha_{11}C_{44} + (e_{15})^2] + 2\alpha_{11}\Delta C_{44} + 2e_{15}\Delta e_{24} - \frac{2(\alpha_{11}\Delta e_{24} - \Delta\alpha_{22}e_{15})(\alpha_{11}\Delta e_{24} - \Delta\alpha_{22}e_{15})}{2[\alpha_{11}C_{44} + (e_{15})^2] + \Delta\alpha_{11}C_{44} + e_{15}\Delta e_{15}} \right\}^{-1} \varepsilon_{23}^0, \quad (3.3.23)$$

$$E_2^I = \frac{2(e_{15}\Delta C_{44} - C_{44}\Delta e_{15})}{2[\alpha_{11}C_{44} + (e_{15})^2] + \Delta\alpha_{11}C_{44} + e_{15}\Delta e_{15}} \varepsilon_{23}^I. \quad (3.3.24)$$

$\bar{C}_{44}$ can be obtained from the following equation

$$2\bar{C}_{44}\varepsilon_{23}^0 = 2C_{44}\varepsilon_{23}^0 + 2c_f\Delta C_{44}\varepsilon_{23}^I - c_f\Delta e_{24}E_2^I. \quad (3.3.25)$$

$\bar{e}_{24}$ can be obtained from the following equation

$$\bar{e}_{24}\varepsilon_{23}^0 = e_{24}\varepsilon_{23}^0 + c_f\Delta\alpha_{22}E_2^I + c_f\Delta e_{24}\varepsilon_{23}^I. \quad (3.3.26)$$

(d) Only $E_3^0 \neq 0$

From Eqs. (3.3.8) and (3.3.9), one obtains

$$\varepsilon_{11}^I = \varepsilon_{22}^I = \frac{\Delta e_{31}E_3^0}{2C_{11} + \Delta C_{11} + \Delta C_{12}}. \quad (3.3.27)$$

According to Eq. (3.3.5), we obtain after manipulations

$$\bar{e}_{33} = e_{33} + c_f\Delta e_{33} - \frac{2c_f\Delta C_{13}\Delta e_{31}}{2C_{11} + \Delta C_{11} + \Delta C_{12}}, \quad (3.3.28)$$

$$\bar{e}_{31} = e_{31} + c_f\Delta e_{31} - \frac{c_f(\Delta C_{11} + \Delta C_{12})\Delta e_{31}}{2C_{11} + \Delta C_{11} + \Delta C_{12}}, \quad (3.3.29)$$

$$\bar{\alpha}_{33} = \alpha_{33} + c_f\Delta\alpha_{33} + \frac{2c_f(\Delta e_{31})^2}{2C_{11} + \Delta C_{11} + \Delta C_{12}}. \quad (3.3.30)$$

(e) Only $E_1^0 \neq 0$

In this case, $\bar{\alpha}_{11}$ is determined by the following equation

$$\langle D_1 \rangle = \bar{\alpha}_{11} E_1^0 = \alpha_{11} E_1^0 + c_f \Delta \alpha_{11} E_1^I + c_f \Delta e_{13} \varepsilon_{13}^I, \quad (3.3.31)$$

where E_1^I and ε_{13}^I can be obtained through Eq. (3.3.12) and Eq. (3.3.14).

For transversely isotropic piezoelectric composites with spheroidal inhomogeneities, Eq. (3.3.5) can be used to obtain

$$\bar{\mathbf{C}} = \mathbf{C} + c_f \Delta \mathbf{C} \mathbf{B}, \quad (3.3.32)$$

where $\Delta \mathbf{C}$ is defined in Eq. (2.6.53) and $\mathbf{B}$ is defined in Eq. (2.6.45).

3.3.2 The Self-Consistent Model

The self-consistent model was first constructed by Bruggeman (1935) and applied to the prediction of effective constants of polycrystalline and composite materials (Hershev, 1954; Kroner, 1958; Budiansky, 1965; Hill, 1965). The basic assumption of the self-consistent model is that every single inhomogeneity is embedded in the equivalent homogeneous composites media, whose material constants are assumed to be the effective constant $\bar{\mathbf{C}}$. Using the self-consistent model, Eq. (3.3.7) is replaced by

$$\mathbf{A}^{\text{sc}} = (\mathbf{I} - \mathbf{S}^{\text{sc}} \bar{\mathbf{C}}^{-1} (\bar{\mathbf{C}} - \mathbf{C}^*))^{-1}. \quad (3.3.33)$$

Eq. (3.3.33) and Eq. (3.3.5) together form an implicit algebraic 9×9 matrix equation for $\bar{\mathbf{C}}$. Yet, in Eq. (3.3.33), the Eshelby tensor $\mathbf{S}^{\text{sc}}$ is also dependent on the unknown $\bar{\mathbf{C}}$, which makes the calculation procedure of the self-consistent model rather cumbersome. The solution process is an iteration procedure. It was pointed out by Dunn and Taya (1993c) that for the self-consistent model, $\bar{\mathbf{C}}$ can be determined without a priori knowledge of the symmetry of the effective electroelastic moduli as the material symmetry has no significant effect on the numerical computation of $\mathbf{S}^{\text{sc}}$.

3.3.3 The Mori-Tanaka Mean Field Model

The Mori-Tanaka mean field model derived from Mori and Tanaka's famous work

(Mori and Tanaka, 1973), which was based on Eshelby's equivalent inclusion method (Eshelby, 1957). A review of the early applications of this method was provided by Taya and Arsenault (1989). Due to its explicit form of expression in predicting the effective properties of two-phase composites, the method is widely accepted and used. Some representative work using the Mori-Tanaka mean field model was shown in the references (Weng, 1990; 1992; Benveniste et al., 1991; Ferrari, 1991).

The basic assumption for the Mori-Tanaka mean field model is that the average generalized strains of the inhomogeneities phase are equivalent to one inhomogeneity embedded in the infinite matrix subjected to an applied electroelastic field equal to the as yet unknown average generalized strains of the matrix phase. The assumption is given mathematically as

$$\langle \boldsymbol{\varepsilon}^I \rangle_{\Omega} = \mathbf{A}^{\text{dil}} \langle \boldsymbol{\varepsilon}^M \rangle_{R-\Omega}, \quad (3.3.34)$$

where $\mathbf{A}^{\text{dil}}$ is given by Eq. (3.3.7). According to Eq. (3.2.16) we have

$$c_f \langle \boldsymbol{\varepsilon}^I \rangle_{\Omega} + (1 - c_f) \langle \boldsymbol{\varepsilon}^M \rangle_{R-\Omega} = \boldsymbol{\varepsilon}_{KI}^0. \quad (3.3.35)$$

Substitution of Eq. (3.3.34) into Eq. (3.3.35) gives

$$\langle \boldsymbol{\varepsilon}^M \rangle_{R-\Omega} = [c_f \mathbf{A}^{\text{dil}} + (1 - c_f) \mathbf{I}]^{-1} \boldsymbol{\varepsilon}_{KI}^0. \quad (3.3.36)$$

Thus

$$\mathbf{A}^{\text{MT}} = \mathbf{A}^{\text{dil}} [c_f \mathbf{A}^{\text{dil}} + (1 - c_f) \mathbf{I}]^{-1}. \quad (3.3.37)$$

Using the Mori-Tanaka mean field model, the effective material constants are given by

$$\bar{\mathbf{C}} = \mathbf{C} - c_f (\mathbf{C} - \mathbf{C}^*) \mathbf{A}^{\text{MT}}. \quad (3.3.38)$$

For spheroidal inhomogeneities embedded in piezoelectric composites with transverse isotropy, the 10 independent components in $\bar{\mathbf{C}}$ are provided by Wang (1994) as

$$\begin{aligned} \bar{C}_{11} = & c_f (C_{11}^* I_{11}^1 + C_{12}^* I_{12}^1 + C_{13}^* I_{31}^1 - e_{31}^* I_{31}^3) \\ & + (1 - c_f) (C_{11} M_{11}^1 + C_{12} M_{12}^1 + C_{13} M_{31}^1 - e_{31} M_{31}^3), \end{aligned} \quad (3.3.39)$$

$$\begin{aligned} \bar{C}_{33} = & c_f (C_{13}^* I_{13}^1 + C_{13}^* I_{13}^1 + C_{33}^* I_{33}^1 - e_{33}^* I_{33}^3) \\ & + (1 - c_f) (C_{13} M_{13}^1 + C_{13} M_{13}^1 + C_{33} M_{33}^1 - e_{33} M_{33}^3), \end{aligned} \quad (3.3.40)$$

$$\begin{aligned}\bar{C}_{12} = & c_f(C_{11}^*I_{12}^1 + C_{13}^*I_{11}^1 + C_{13}^*I_{31}^1 - e_{31}^*I_{31}^3) \\ & + (1 - c_f)(C_{11}M_{12}^1 + C_{12}M_{11}^1 + C_{13}M_{31}^1 - e_{31}M_{33}^3),\end{aligned}\quad (3.3.41)$$

$$\begin{aligned}\bar{C}_{13} = & c_f(C_{11}^*I_{13}^1 + C_{12}^*I_{13}^1 + C_{13}^*I_{33}^1 - e_{31}^*I_{33}^3) \\ & + (1 - c_f)(C_{11}M_{13}^1 + C_{12}M_{13}^1 + C_{13}M_{33}^1 - e_{31}M_{33}^3),\end{aligned}\quad (3.3.42)$$

$$\bar{C}_{44} = C_{44}^* - (1 - c_f)\Delta C_{44}M_{44}^1 + \frac{1}{2}(1 - c_f)\Delta e_{24}M_{15}^3, \quad (3.3.43)$$

$$\begin{aligned}\bar{e}_{31} = & -c_f(C_{11}^*I_{13}^2 + C_{12}^*I_{13}^2 + C_{13}^*I_{33}^2 - e_{31}^*I_{33}^4) \\ & - (1 - c_f)(C_{11}M_{13}^2 + C_{12}M_{13}^2 + C_{13}M_{33}^2 - e_{31}M_{33}^4),\end{aligned}\quad (3.3.44)$$

$$\begin{aligned}\bar{e}_{33} = & -c_f(C_{13}^*I_{13}^2 + C_{13}^*I_{13}^2 + C_{33}^*I_{33}^2 - e_{33}^*I_{33}^4) \\ & - (1 - c_f)(C_{13}M_{13}^2 + C_{13}M_{13}^2 + C_{33}M_{33}^2 - e_{33}M_{33}^4),\end{aligned}\quad (3.3.45)$$

$$\bar{\alpha}_{11} = \alpha_{11} - (1 - c_f)\Delta\alpha_{11}M_{11}^4 - 2(1 - c_f)\Delta e_{24}M_{42}^2, \quad (3.3.46)$$

$$\begin{aligned}\bar{\alpha}_{33} = & c_f(e_{31}^*I_{13}^2 + e_{31}^*I_{13}^2 + e_{33}^*I_{33}^2 - \alpha_{33}^*I_{33}^4) \\ & + (1 - c_f)(e_{31}M_{13}^2 + e_{31}M_{13}^2 + e_{33}M_{33}^2 - \alpha_{33}M_{33}^4).\end{aligned}\quad (3.3.47)$$

M_{ij}^k and I_{ij}^k are defined by the following equations

$$\begin{bmatrix} \langle \varepsilon_{11}^M \rangle \\ \langle \varepsilon_{22}^M \rangle \\ \langle \varepsilon_{33}^M \rangle \\ \langle \varepsilon_{23}^M \rangle \\ \langle \varepsilon_{13}^M \rangle \\ \langle \varepsilon_{12}^M \rangle \\ \langle E_1^M \rangle \\ \langle E_2^M \rangle \\ \langle E_3^M \rangle \end{bmatrix} = \begin{bmatrix} M_{11}^1 & M_{12}^1 & M_{13}^1 & 0 & 0 & 0 & 0 & 0 & M_{13}^2 \\ M_{12}^1 & M_{11}^1 & M_{13}^1 & 0 & 0 & 0 & 0 & 0 & M_{13}^2 \\ M_{31}^1 & M_{31}^1 & M_{33}^1 & 0 & 0 & 0 & 0 & 0 & M_{33}^2 \\ 0 & 0 & 0 & M_{44}^1 & 0 & 0 & 0 & M_{42}^2 & 0 \\ 0 & 0 & 0 & 0 & M_{44}^1 & 0 & M_{42}^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & M_{66}^1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & M_{15}^3 & 0 & M_{11}^4 & 0 & 0 \\ 0 & 0 & 0 & M_{15}^3 & 0 & 0 & 0 & M_{11}^4 & 0 \\ M_{31}^3 & M_{31}^3 & M_{33}^3 & 0 & 0 & 0 & 0 & 0 & M_{33}^4 \end{bmatrix} \begin{bmatrix} \varepsilon_{11}^0 \\ \varepsilon_{22}^0 \\ \varepsilon_{33}^0 \\ \varepsilon_{23}^0 \\ \varepsilon_{13}^0 \\ \varepsilon_{12}^0 \\ E_1^0 \\ E_2^0 \\ E_3^0 \end{bmatrix}, \quad (3.3.48)$$

where

$$\begin{bmatrix} M_{11}^1 & M_{12}^1 & M_{13}^1 \\ M_{12}^1 & M_{11}^1 & M_{13}^1 \\ M_{31}^1 & M_{31}^1 & M_{33}^1 \end{bmatrix} = \begin{bmatrix} g_{11} & g_{12} & g_{13} \\ g_{12} & g_{11} & g_{13} \\ g_{31} & g_{31} & g_{33} \end{bmatrix}^{-1}, \quad (3.3.49)$$

and

$$\begin{bmatrix} \langle \varepsilon_{11}^I \rangle \\ \langle \varepsilon_{22}^I \rangle \\ \langle \varepsilon_{33}^I \rangle \\ \langle \varepsilon_{23}^I \rangle \\ \langle \varepsilon_{13}^I \rangle \\ \langle \varepsilon_{12}^I \rangle \\ \langle E_1^I \rangle \\ \langle E_2^I \rangle \\ \langle E_3^I \rangle \end{bmatrix} = \begin{bmatrix} I_{11}^1 & I_{12}^1 & I_{13}^1 & 0 & 0 & 0 & 0 & 0 & I_{13}^2 \\ I_{12}^1 & I_{11}^1 & I_{13}^1 & 0 & 0 & 0 & 0 & 0 & I_{13}^2 \\ I_{31}^1 & I_{31}^1 & I_{33}^1 & 0 & 0 & 0 & 0 & 0 & I_{33}^2 \\ 0 & 0 & 0 & I_{44}^1 & 0 & 0 & 0 & I_{42}^2 & 0 \\ 0 & 0 & 0 & 0 & I_{44}^1 & 0 & I_{42}^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & I_{66}^1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & I_{15}^3 & 0 & I_{11}^4 & 0 & 0 \\ 0 & 0 & 0 & I_{15}^3 & 0 & 0 & 0 & I_{11}^4 & 0 \\ I_{31}^3 & I_{31}^3 & I_{33}^3 & 0 & 0 & 0 & 0 & 0 & I_{33}^4 \end{bmatrix} \begin{bmatrix} \varepsilon_{11}^0 \\ \varepsilon_{22}^0 \\ \varepsilon_{33}^0 \\ \varepsilon_{23}^0 \\ \varepsilon_{13}^0 \\ \varepsilon_{12}^0 \\ E_1^0 \\ E_2^0 \\ E_3^0 \end{bmatrix}. \quad (3.3.50)$$

The components in Eqs. (3.3.48)–(3.3.50) are given by

$$g_{11} = 1 - c_f + c_f B_{11}^1 - c_f^2 \Xi_1 B_{13}^2 B_{31}^3, \quad (3.3.51)$$

$$g_{33} = 1 - c_f + c_f B_{33}^1 - c_f^2 \Xi_1 B_{33}^2 B_{33}^3, \quad (3.3.52)$$

$$g_{12} = c_f B_{12}^1 - c_f^2 \Xi_1 B_{13}^2 B_{31}^3, \quad (3.3.53)$$

$$g_{13} = c_f B_{13}^1 - c_f^2 \Xi_1 B_{13}^2 B_{33}^3, \quad (3.3.54)$$

$$g_{31} = c_f B_{31}^1 - c_f^2 \Xi_1 B_{33}^2 B_{31}^3, \quad (3.3.55)$$

$$M_{44}^1 = [1 - c_f + c_f B_{44}^1 - \frac{1}{2} c_f^2 (1 - c_f + c_f B_{11}^4)^{-1} B_{42}^2 B_{15}^3]^{-1}, \quad (3.3.56)$$

$$M_{66}^1 = (1 - c_f + c_f B_{66}^1)^{-1}, \quad (3.3.57)$$

$$M_{13}^2 = -c_f \Xi_1 (M_{11}^1 B_{13}^2 + M_{12}^1 B_{13}^2 + M_{13}^1 B_{33}^2), \quad (3.3.58)$$

$$M_{33}^2 = -c_f \Xi_1 (M_{31}^1 B_{13}^2 + M_{31}^1 B_{13}^2 + M_{33}^1 B_{33}^2), \quad (3.3.59)$$

$$M_{42}^2 = -\frac{1}{2} c_f (1 - c_f + c_f B_{11}^4)^{-1} M_{44}^1 B_{42}^2, \quad (3.3.60)$$

$$M_{31}^3 = -c_f \Xi_1 (M_{11}^1 B_{31}^3 + M_{12}^1 B_{31}^3 + M_{13}^1 B_{33}^3), \quad (3.3.61)$$

$$M_{33}^3 = -c_f \Xi_1 (M_{13}^1 B_{31}^3 + M_{13}^1 B_{31}^3 + M_{33}^1 B_{33}^3), \quad (3.3.62)$$

$$M_{15}^3 = -\frac{1}{2} c_f (1 - c_f + c_f B_{11}^4)^{-1} M_{44}^1 B_{15}^3, \quad (3.3.63)$$

$$M_{11}^4 = (1 - c_f + c_f B_{11}^4)^{-1} + \frac{1}{2} c_f^2 (1 - c_f + c_f B_{11}^4)^{-2} M_{44}^1 B_{15}^3 B_{42}^2, \quad (3.3.64)$$

$$M_{33}^4 = \bar{\mathcal{E}}_1 - c_f \bar{\mathcal{E}}_1 (B_{31}^3 M_{13}^2 + B_{31}^3 M_{13}^2 + B_{33}^3 M_{33}^2), \quad (3.3.65)$$

where

$$\bar{\mathcal{E}}_1 = (1 - c_f + c_f B_{11}^4)^{-1}. \quad (3.3.66)$$

And for I_{ij}^k we have

$$I_{11}^1 = B_{11}^1 M_{11}^1 + B_{12}^1 M_{12}^1 + B_{13}^1 M_{31}^1 + B_{13}^2 M_{31}^3, \quad (3.3.67)$$

$$I_{12}^1 = B_{11}^1 M_{12}^1 + B_{12}^1 M_{11}^1 + B_{13}^1 M_{31}^1 + B_{13}^2 M_{31}^3, \quad (3.3.68)$$

$$I_{13}^1 = B_{11}^1 M_{13}^1 + B_{12}^1 M_{13}^1 + B_{13}^1 M_{33}^1 + B_{13}^2 M_{33}^3, \quad (3.3.69)$$

$$I_{33}^1 = B_{31}^1 M_{13}^1 + B_{31}^1 M_{13}^1 + B_{33}^1 M_{33}^1 + B_{33}^2 M_{33}^3, \quad (3.3.70)$$

$$I_{44}^1 = \frac{1}{c_f} [1 - (1 - c_f) M_{44}^1], \quad (3.3.71)$$

$$I_{66}^1 = \frac{1}{c_f} [1 - (1 - c_f) M_{66}^1], \quad (3.3.72)$$

$$I_{13}^2 = B_{11}^1 M_{13}^2 + B_{12}^1 M_{13}^2 + B_{13}^1 M_{33}^3 + B_{13}^2 M_{33}^4, \quad (3.3.73)$$

$$I_{33}^2 = B_{31}^1 M_{13}^2 + B_{31}^1 M_{13}^2 + B_{33}^1 M_{33}^3 + B_{33}^2 M_{33}^4, \quad (3.3.74)$$

$$I_{42}^2 = -\frac{1}{c_f} (1 - c_f) M_{42}^1, \quad (3.3.75)$$

$$I_{31}^3 = B_{31}^3 M_{11}^1 + B_{31}^3 M_{12}^1 + B_{33}^3 M_{31}^1 + B_{33}^4 M_{31}^3, \quad (3.3.76)$$

$$I_{33}^3 = B_{31}^3 M_{13}^1 + B_{31}^3 M_{13}^1 + B_{33}^3 M_{33}^1 + B_{33}^4 M_{33}^3, \quad (3.3.77)$$

$$I_{15}^3 = -\frac{1}{c_f} (1 - c_f) M_{15}^3, \quad (3.3.78)$$

$$I_{33}^4 = B_{31}^3 M_{13}^2 + B_{31}^3 M_{13}^2 + B_{33}^3 M_{33}^2 + B_{33}^4 M_{33}^4, \quad (3.3.79)$$

$$I_{11}^4 = \frac{1}{c_f} [1 - (1 - c_f) M_{11}^4]. \quad (3.3.80)$$

3.3.4 The Differential Model

The differential model is an alternative method for estimation of the effective material properties of composites. A thorough explanation of its early application was provided by McLaughlin (1977). The method is realized by successively replacing the matrix material by an incremental volume of the reinforcement which finally approaches the composite material.

For the instantaneous configuration of the composite (the volume fraction of the inhomogeneity is c_f), if we remove a volume increment dV , and replace the cavity by pure inhomogeneity, we have

$$dc_f = \frac{dV}{V}(1 - c_f), \quad (3.3.81)$$

where V is the volume of the composite. Denoting $\bar{\mathbf{C}}(c_f + dc_f)$ as the effective material constants at volume fraction $c_f + dc_f$, using Eq. (3.3.81) and Eq. (3.3.5), one obtains

$$\frac{d\bar{\mathbf{C}}}{dc_f} = \frac{1}{1 - c_f}(\mathbf{C}^* - \bar{\mathbf{C}})\mathbf{A}^{\text{dif}}, \quad (3.3.82)$$

where

$$\mathbf{A}^{\text{dif}} = [\mathbf{I} + \mathbf{S}^{\text{dif}}\bar{\mathbf{C}}^{-1}(\mathbf{C}^* - \bar{\mathbf{C}})]^{-1}. \quad (3.3.83)$$

$\mathbf{S}^{\text{dif}}$ in Eq. (3.3.83) is a function of the unknown $\bar{\mathbf{C}}$ of the instantaneous material. Generally speaking, Eq. (3.3.82) represents a set of 9×9 coupled nonlinear ordinary differential equations which can be solved with the initial condition

$$\bar{\mathbf{C}}(c_f = 0) = \mathbf{C}. \quad (3.3.84)$$

3.4 Energy Formulation of Ferroelectric Composites

In this section, the overall properties of ferroelectric composites are considered. From the above discussion, it is shown that through various micromechanics models, the overall properties for piezoelectric composites can be obtained, once the linear constitutive equation holds:

$$\begin{aligned}\sigma_{ij} &= C_{ijkl}^E \epsilon_{kl} - e_{kij} E_k, \\ D_i &= e_{ikl} \epsilon_{kl} + \alpha_{ik}^E E_k\end{aligned}\quad (3.4.1)$$

This linear relation describes the physical properties for electroelastic materials in three parts: a) the mechanical properties, described in Eq. (3.4.1) by the term $C_{ijkl}^E \epsilon_{kl}$; b) the electrical properties, described in Eq. (3.4.1) by the term $\alpha_{ik}^E E_k$; c) the electromechanical properties, described in Eq. (3.4.1) by the term $-e_{kij} E_k$ or $e_{ikl} \epsilon_{kl}$. For ferroelectric materials, the electric properties and the electromechanical properties no longer satisfy the linear relation. When nonlinear effects are included, the micromechanics models introduced in Section 3.2, which are based on linear constitutive equations, are not useful. The nonlinear contribution in the free energy of the ferroelectric composite system plays an important role and determines the macroscopic properties, such as the phase diagram. If one uses the spontaneous polarizations of the ferroelectric material as their order parameters to describe the state, the Helmholtz free energy of the system consists of two parts: one is the intrinsic free energy related to the ferroelectric material, which includes the Landau free energy expansion, the effect of depolarization fields, etc.; the other is the elastic energy induced by the interaction of the inclusions and matrix accompanying the phase transformation. For a further explanation, the inclusions and matrix are composed of different ferroelectric materials; they will have different transformation strains, and thus will induce a misfit strain field and increase the elastic energy of the system. Thus the Helmholtz free energy of the system can be written in the form

$$\psi(\epsilon_{ij}, P_i) = F^{\text{ela}}(\epsilon_{ij}, P_i) + F^{\text{ele}}(\epsilon_{ij}, P_i), \quad (3.4.2)$$

where F^{ela} and F^{ele} denote the elastic strain energy density and the intrinsic free energy density of the system, respectively. Now we try to obtain the expressions for F^{ela} and F^{ele} for ferroelectric composites.

3.4.1 Elastic Strain Energy Density for Ferroelectric Composites

The elastic strain energy describes the total elastic energy of the ferroelectric composites, including the elastic strain energy for each ferroelectric phase, and the elastic interaction energy between the two phases. For a ferroelectric composite with complicated microstructures, the elastic strain energy density is a function of the location $\mathbf{x}$, whose analytical expression is not available. Yet, by using the micromechanics models introduced in Section 3.2, it is possible to obtain the average stresses and strains for composites with eigenstrains. Define the average elastic strain energy density by

$$\bar{F}^{\text{ela}} = \frac{1}{2} c_f < \boldsymbol{\sigma}^I >_{\Omega}^T < \mathbf{e}^I >_{\Omega} + \frac{1}{2} (1 - c_f) < \boldsymbol{\sigma}^M >_{R-\Omega}^T < \mathbf{e}^M >_{R-\Omega}, \quad (3.4.3)$$

where the inhomogeneity phase occupies Ω in R and c_f is the volume fraction of Ω , physical quantities with superscript I refer to the inhomogeneity phase and physical quantities with superscript M refer to the matrix phase. Before obtaining the exact expressions for the average stresses and strains, we have to determine the source of the eigenstrains in the first place. For ferroelectric materials, consider the non-elastic deformation caused by the electrostriction effect. The eigenstrains are then given by

$$\varepsilon_{ij}^* = Q_{kij} P_k P_l, \quad (3.4.4)$$

where Q_{kij} denotes the electrostriction coefficient. For a two-phase ferroelectric composite, both phases undergo eigenstrains, which differ from each other by the electrostriction coefficients and electric polarizations. Denote the eigenstrains for the two phases in the ferroelectric composite as

$$\begin{aligned} \varepsilon_{ij}^{I*} &= Q_{kij}^I P_k^I P_l^I, \\ \varepsilon_{ij}^{M*} &= Q_{kij}^M P_k^M P_l^M. \end{aligned} \quad (3.4.5)$$

The micromechanical problem of ellipsoidal inhomogeneities embedded in the infinite matrix when both phases undergo eigenstrains is solved in Section 2.6.3. The solution Eq. (2.6.140) relies on the assumption that the eigenstrains in the two phases are uniform in the domains where they are defined. In a mean field approximation, we assume P_k^I and P_k^M to be the statistically unified electric polarizations, so that the eigenstrains ε_{ij}^{I*} and ε_{ij}^{M*} are uniform throughout the inhomogeneity phase and the matrix phase. The physical meaning of this approximation can also be expressed by saying that the two phases of the composite are assumed to be made up of a single crystalline single domain ferroelectric material. The average stresses and strains are obtained by using the Mori-Tanaka mean field model. When the total strains are given at the boundary by $\boldsymbol{\varepsilon}^0$, the average total strains in the inhomogeneity phase $< \boldsymbol{\varepsilon}^I >_{\Omega}$ and the matrix phase $< \boldsymbol{\varepsilon}^M >_{R-\Omega}$ are related by

$$c_f < \boldsymbol{\varepsilon}^I >_{\Omega} + (1 - c_f) < \boldsymbol{\varepsilon}^M >_{R-\Omega} = \boldsymbol{\varepsilon}^0, \quad (3.4.6)$$

where

$$\begin{aligned} < \boldsymbol{\varepsilon}^I >_{\Omega} &= < \mathbf{e}^I >_{\Omega} + \varepsilon_{ij}^{I*}, \\ < \boldsymbol{\varepsilon}^M >_{R-\Omega} &= < \mathbf{e}^M >_{R-\Omega} + \varepsilon_{ij}^{M*}. \end{aligned} \quad (3.4.7)$$

Every single inclusion is considered to be embedded in the infinite matrix suffering average elastic strains $\langle \mathbf{e}^M \rangle_{R-\Omega}$ at infinity. According to the solution of the disturbance field obtained for one inclusion in Eq. (2.6.134) and Eq. (2.6.140), we have

$$\langle \mathbf{e}^I \rangle_{\Omega} = \langle \mathbf{e}^M \rangle_{R-\Omega} - \boldsymbol{\varepsilon}^R + \mathbf{S}(\boldsymbol{\varepsilon}^R + \boldsymbol{\varepsilon}^{**}). \quad (3.4.8)$$

From Eq. (2.6.137) we have

$$\mathbf{C}(\langle \mathbf{e}^I \rangle_{\Omega} - \boldsymbol{\varepsilon}^{**}) = \mathbf{C}^* \langle \mathbf{e}^I \rangle_{\Omega} \quad \text{in } \Omega. \quad (3.4.9)$$

From Eqs. (3.4.6)–(3.4.8) we have

$$\begin{aligned} \langle \mathbf{e}^M \rangle_{R-\Omega} &= \boldsymbol{\varepsilon}^{\infty} - c_f \mathbf{S}(\boldsymbol{\varepsilon}^R + \boldsymbol{\varepsilon}^{**}), \\ \langle \mathbf{e}^I \rangle_{\Omega} &= \boldsymbol{\varepsilon}^{\infty} + (1 - c_f) \mathbf{S}(\boldsymbol{\varepsilon}^R + \boldsymbol{\varepsilon}^{**}) - \boldsymbol{\varepsilon}^R. \end{aligned} \quad (3.4.10)$$

Substitution of Eq. (3.4.10) into Eq. (3.4.9) gives

$$\boldsymbol{\varepsilon}^{**} = \mathbf{A}^I \boldsymbol{\varepsilon}^R + \mathbf{B}^I \boldsymbol{\varepsilon}^{\infty}, \quad (3.4.11)$$

where

$$\begin{aligned} \mathbf{A}^I &= [(1 - c_f)(\mathbf{S} - \mathbf{I})(\mathbf{C}^* - \mathbf{C}) + \mathbf{C}]^{-1} \mathbf{C}^* - \mathbf{I}, \\ \mathbf{B}^I &= [(1 - c_f)(\mathbf{S} - \mathbf{I})(\mathbf{C}^* - \mathbf{C}) + \mathbf{C}]^{-1} (\mathbf{C} - \mathbf{C}^*). \end{aligned} \quad (3.4.12)$$

Eq. (3.4.10) is finally transformed as

$$\begin{aligned} \langle \mathbf{e}^M \rangle_{R-\Omega} &= \mathbf{E} \boldsymbol{\varepsilon}^{\infty} + \mathbf{F} \boldsymbol{\varepsilon}^R, \\ \langle \mathbf{e}^I \rangle_{\Omega} &= \mathbf{G} \boldsymbol{\varepsilon}^{\infty} + \mathbf{H} \boldsymbol{\varepsilon}^R, \end{aligned} \quad (3.4.13)$$

where

$$\begin{aligned} \mathbf{E} &= \mathbf{I} - c_f \mathbf{S} \mathbf{B}^I, \quad \mathbf{F} = -c_f \mathbf{S}(\mathbf{I} + \mathbf{A}^I), \\ \mathbf{G} &= \mathbf{I} + (1 - c_f) \mathbf{S} \mathbf{B}^I, \quad \mathbf{H} = (1 - c_f) \mathbf{S}(\mathbf{I} + \mathbf{A}^I) - \mathbf{I}. \end{aligned} \quad (3.4.14)$$

Substitution of Eq. (3.4.13) into Eq. (3.4.3) leads to

$$\bar{F}^{\text{cla}} = (\boldsymbol{\varepsilon}^{\infty})^T \mathbf{J} \boldsymbol{\varepsilon}^{\infty} + (\boldsymbol{\varepsilon}^{\infty})^T \mathbf{K} \boldsymbol{\varepsilon}^R + (\boldsymbol{\varepsilon}^R)^T \mathbf{L} \boldsymbol{\varepsilon}^R, \quad (3.4.15)$$

where

$$\begin{aligned}\mathbf{J} &= \frac{1}{2}c_f \mathbf{G} \mathbf{C}^* \mathbf{G} + \frac{1}{2}(1-c_f) \mathbf{E} \mathbf{C} \mathbf{E}, \\ \mathbf{K} &= c_f \mathbf{G} \mathbf{C}^* \mathbf{H} + (1-c_f) \mathbf{E} \mathbf{C} \mathbf{F}, \\ \mathbf{L} &= \frac{1}{2}c_f \mathbf{H} \mathbf{C}^* \mathbf{H} + \frac{1}{2}(1-c_f) \mathbf{F} \mathbf{C} \mathbf{F}.\end{aligned}\quad (3.4.16)$$

According to Eqs. (2.6.66) and (2.6.69), Eq. (3.4.15) is finally given as

$$\begin{aligned}\bar{F}^{\text{ela}} &= (\boldsymbol{\varepsilon}^0)^T \mathbf{N}^{01} \boldsymbol{\varepsilon}^0 + (\boldsymbol{\varepsilon}^0)^T \mathbf{N}^{11} \boldsymbol{\varepsilon}^{M*} + (\boldsymbol{\varepsilon}^0)^T \mathbf{N}^{12} \boldsymbol{\varepsilon}^{I*} + (\boldsymbol{\varepsilon}^{M*})^T \mathbf{N}^{Ml} \boldsymbol{\varepsilon}^{I*} \\ &\quad + (\boldsymbol{\varepsilon}^{M*})^T \mathbf{N}^{21} \boldsymbol{\varepsilon}^{M*} + (\boldsymbol{\varepsilon}^{I*})^T \mathbf{N}^{22} \boldsymbol{\varepsilon}^{I*},\end{aligned}\quad (3.4.17)$$

where

$$\begin{aligned}\mathbf{N}^{01} &= \mathbf{J}, \quad \mathbf{N}^{11} = -2\mathbf{J} - \mathbf{K}, \quad \mathbf{N}^{12} = \mathbf{K}, \\ \mathbf{N}^{Ml} &= -\mathbf{K} - 2\mathbf{L}, \quad \mathbf{N}^{21} = \mathbf{J} + \mathbf{K} + \mathbf{L}, \quad \mathbf{N}^{22} = \mathbf{L}.\end{aligned}\quad (3.4.18)$$

In Eq. (3.4.17), the average elastic energy density is expressed in terms of the external generalized strain $\boldsymbol{\varepsilon}^0$, and the eigenstrains $\boldsymbol{\varepsilon}^{I*}$, $\boldsymbol{\varepsilon}^{M*}$. For coherent interfaces, the misfit strains should be included in $\boldsymbol{\varepsilon}^{I*}$.

3.4.2 Intrinsic Free Energy Density for Ferroelectric Composites

The intrinsic free energy describes the total electrostatic energy of the ferroelectric composites, including the energy for each ferroelectric phase without interior electric fields, and the electrostatic interaction energy between the two phases. The latter part can be explained as follows: in a general condition, for a ferroelectric composite, the matrix phase and the inclusion phase both possess electric polarizations, denoted by P_i^M and P_i^I . Due to the existence of interfaces and surfaces, bound charges will generate depolarization fields. Since the two phases are connected in the composite, the solutions for the interior fields are correlated. The electrostatic interaction energy describes the contribution of these interior fields.

For a single phase bulk ferroelectric crystal, after the phase transition from paraelectric phase to ferroelectric phase, the intrinsic free energy density f^{ele} can be obtained by taking a Landau expansion (Muller, 1935; 1940a; 1940b; 1940c; Ginsburg, 1945; Devonshire, 1949) as

$$\begin{aligned}
f^{\text{ele}} = & a_1(P_1^2 + P_2^2 + P_3^2) + a_{11}(P_1^4 + P_2^4 + P_3^4) + a_{12}(P_1^2 P_2^2 + P_1^2 P_3^2 + P_2^2 P_3^2) \\
& + a_{111}(P_1^6 + P_2^6 + P_3^6) + a_{112}[P_1^4(P_2^2 + P_3^2) + P_2^4(P_1^2 + P_3^2) \\
& + P_3^4(P_1^2 + P_2^2)] + a_{123}P_1^2 P_2^2 P_3^2.
\end{aligned} \quad (3.4.19)$$

Notice that in Eq. (3.4.19) the derivatives of the electric polarizations are not included. The electrostatic energy density for the two-phase composite is given by

$$F^{\text{bulk}} = c_f f^I + (1 - c_f) f^M, \quad (3.4.20)$$

where

$$\begin{aligned}
f^I = & a_1^I[(P_1^I)^2 + (P_2^I)^2 + (P_3^I)^2] + a_{11}^I[(P_1^I)^4 + (P_2^I)^4 + (P_3^I)^4] \\
& + a_{12}^I[(P_1^I)^2(P_2^I)^2 + (P_1^I)^2(P_3^I)^2 + (P_2^I)^2(P_3^I)^2] + a_{111}^I[(P_1^I)^6 \\
& + (P_2^I)^6 + (P_3^I)^6] + a_{112}^I\{(P_1^I)^4[(P_2^I)^2 + (P_3^I)^2] + (P_2^I)^4[(P_1^I)^2 \\
& + (P_3^I)^2] + (P_3^I)^4[(P_1^I)^2 + (P_2^I)^2]\} + a_{123}^I(P_1^I)^2(P_2^I)^2(P_3^I)^2,
\end{aligned} \quad (3.4.21)$$

$$\begin{aligned}
f^M = & a_1^M[(P_1^M)^2 + (P_2^M)^2 + (P_3^M)^2] + a_{11}^M[(P_1^M)^4 + (P_2^M)^4 + (P_3^M)^4] \\
& + a_{12}^M[(P_1^M)^2(P_2^M)^2 + (P_1^M)^2(P_3^M)^2 + (P_2^M)^2(P_3^M)^2] + a_{111}^M[(P_1^M)^6 \\
& + (P_2^M)^6 + (P_3^M)^6] + a_{112}^M\{(P_1^M)^4[(P_2^M)^2 + (P_3^M)^2] + (P_2^M)^4[(P_1^M)^2 \\
& + (P_3^M)^2] + (P_3^M)^4[(P_1^M)^2 + (P_2^M)^2]\} + a_{123}^M(P_1^M)^2(P_2^M)^2(P_3^M)^2.
\end{aligned} \quad (3.4.22)$$

Next, we discuss the electrostatic interaction energy between the two phases. For one dielectric ellipsoid embedded in another dielectric, linear electrostatic theory permits the following relation (Landau and Lifshitz, 1960)

$$(1 - d_i)\alpha_i^M E_i^I + d_i D_i^I = \alpha_i^M E_i^0 \quad (\text{no sum on } i), \quad (3.4.23)$$

where α_i^M denotes the diagonal components of the matrix of linear permittivity of the matrix material (it is assumed that non-diagonal components of the matrix of linear permittivity are zero), E_i^I , D_i^I denote components of the electric field vector and the electric displacement vector for the inclusion phase respectively, and E_i^0 denotes components of the applied electric field vector. In Eq. (3.4.23), the coefficients d_i are called the depolarization factors for the inclusion phase. They describe the influence of the configuration of the inclusion on the electric field. They resemble Eshelby's tensor in anisotropic elasticity, only with less complexity. For elliptical shape of inclusions, d_i can be written as

$$\begin{aligned}
d_1 &= \frac{1}{2} abc \int_0^\infty \frac{ds}{(s+a^2)R_s}, \\
d_2 &= \frac{1}{2} abc \int_0^\infty \frac{ds}{(s+b^2)R_s}, \\
d_3 &= \frac{1}{2} abc \int_0^\infty \frac{ds}{(s+c^2)R_s},
\end{aligned} \tag{3.4.24}$$

where

$$R_s = \sqrt{(s+a^2)(s+b^2)(s+c^2)}, \tag{3.4.25}$$

and a, b and c are the three semi-axes for the ellipsoidal inhomogeneity. For a spherical inhomogeneity, Eq. (3.4.24) reduces to

$$d_1 = d_2 = d_3 = \frac{1}{3}. \tag{3.4.26}$$

The electric displacements D_i^I are defined as

$$D_i^I = \alpha_d^I E_i^I + P_i^I, \tag{3.4.27}$$

where α_d^I denotes the dielectric permittivity of the reference state far from any instability (Woo and Zheng, 2008). Substitution of Eq. (3.4.27) into Eq. (3.4.23) gives

$$E_i^I = \frac{\alpha_i^M}{[(1-d_i)\alpha_i^M + d_i\alpha_d^I]} E_i^0 - \frac{d_i}{[(1-d_i)\alpha_i^M + d_i\alpha_d^I]} P_i^I \quad (\text{no sum on } i). \tag{3.4.28}$$

When external electric fields are absent, Eq. (3.4.28) reduces to

$$E_i^I = -\frac{d_i}{[(1-d_i)\alpha_i^M + d_i\alpha_d^I]} P_i^I \quad (\text{no sum on } i). \tag{3.4.29}$$

For a two-phase ferroelectric composite occupying R with randomly distributed ellipsoidal inclusions occupying Ω , the average electric fields in both phases can be solved by using the Mori-Tanaka method. Assume that every ellipsoidal inclusion is embedded in the infinite matrix with applied electric fields $\langle E_i^M \rangle_{R-\Omega}$, which is equivalent to the average electric fields in the matrix phase. Then, from Eq. (3.4.28) we have

$$\langle E_i^I \rangle_{\Omega} = \frac{\alpha_i^M}{[(1-d_i)\alpha_i^M + d_i\alpha_d^I]} \langle E_i^M \rangle_{R-\Omega} - \frac{d_i}{[(1-d_i)\alpha_i^M + d_i\alpha_d^I]} P_i^I \quad (\text{no sum on } i). \quad (3.4.30)$$

When no external electric field is applied, the electric boundary condition for the composite is given by Eq. (2.2.10) as

$$D_i n_i = 0. \quad (3.4.31)$$

From Eqs. (3.2.5)–(3.2.8), we have

$$\langle D_i \rangle_R = 0. \quad (3.4.32)$$

Eq. (3.4.32) can be expanded as

$$c_f \langle D_i^I \rangle_{\Omega} + (1-c_f) \langle D_i^M \rangle_{R-\Omega} = 0. \quad (3.4.33)$$

The average electric displacements satisfy

$$\begin{aligned} \langle D_i^I \rangle_{\Omega} &= \alpha_i^I \langle E_i^I \rangle_{\Omega}, \\ \langle D_i^M \rangle_{R-\Omega} &= \alpha_i^M \langle E_i^M \rangle_{R-\Omega}. \end{aligned} \quad (3.4.34)$$

Substitution of Eqs. (3.4.30) and (3.4.34) into Eq. (3.4.33) yields

$$\langle E_i^M \rangle_{R-\Omega} = \frac{c_f d_i \alpha_i^I}{\alpha_i^M [c_f \alpha_i^I + (1-d_i)(1-c_f) \alpha_i^M + d_i(1-c_f) \alpha_d^I]} P_i^I \quad (\text{no sum on } i), \quad (3.4.35)$$

$$\langle E_i^I \rangle_{\Omega} = - \frac{(1-c_f) d_i}{[c_f \alpha_i^I + (1-d_i)(1-c_f) \alpha_i^M + d_i(1-c_f) \alpha_d^I]} P_i^I \quad (\text{no sum on } i). \quad (3.4.36)$$

For spherical inclusions, Eqs. (3.4.35) and (3.4.36) reduce to

$$\langle E_i^M \rangle_{R-\Omega} = \frac{c_f \alpha_i^I}{\alpha_i^M [3c_f \alpha_i^I + 2(1-c_f) \alpha_i^M + (1-c_f) \alpha_d^I]} P_i^I \quad (\text{no sum on } i), \quad (3.4.37)$$

$$\langle E_i^I \rangle_{\Omega} = - \frac{(1-c_f)}{[3c_f \alpha_i^I + 2(1-c_f) \alpha_i^M + (1-c_f) \alpha_d^I]} P_i^I \quad (\text{no sum on } i). \quad (3.4.38)$$

In Eqs. (3.4.35)–(3.4.38), the average electric fields in each phase are expressed as linear functions of the electric polarizations of the inclusion phase. One should always bear in mind that, strictly speaking, α_i^I and α_i^M both depend on the electric field and the polarizations. Yet, in the calculation process, they are kept as constants for simplicity.

The intrinsic free energy density for a two-phase ferroelectric composite is finally given by

$$\begin{aligned}\bar{F}^{\text{ele}} &= F^{\text{bulk}} + \bar{F}^{\text{field}} \\ &= c_f f^I + (1 - c_f) f^M - \frac{1}{2} c_f \langle E_i^I \rangle_{\Omega} P_i^I - \frac{1}{2} (1 - c_f) \langle E_i^M \rangle_{R-\Omega} P_i^M.\end{aligned}\quad (3.4.39)$$

Expanding Eq. (3.4.39), we finally have

$$\begin{aligned}\bar{F}^{\text{ele}} &= c_f \{ \alpha_1^I [(P_1^I)^2 + (P_2^I)^2 + (P_3^I)^2] + \alpha_{11}^I [(P_1^I)^4 + (P_2^I)^4 + (P_3^I)^4] \\ &\quad + \alpha_{12}^I [(P_1^I)^2 (P_2^I)^2 + (P_1^I)^2 (P_3^I)^2 + (P_2^I)^2 (P_3^I)^2] + \alpha_{111}^I [(P_1^I)^6 \\ &\quad + (P_2^I)^6 + (P_3^I)^6] + \alpha_{112}^I \{ (P_1^I)^4 [(P_2^I)^2 + (P_3^I)^2] + (P_2^I)^4 [(P_1^I)^2 + (P_3^I)^2] \\ &\quad + (P_3^I)^4 [(P_1^I)^2 + (P_2^I)^2] \} + \alpha_{123}^I (P_1^I)^2 (P_2^I)^2 (P_3^I)^2 \} \\ &\quad + (1 - c_f) \{ \alpha_1^M [(P_1^M)^2 + (P_2^M)^2 + (P_3^M)^2] + \alpha_{11}^M [(P_1^M)^4 + (P_2^M)^4 + (P_3^M)^4] \\ &\quad + \alpha_{12}^M [(P_1^M)^2 (P_2^M)^2 + (P_1^M)^2 (P_3^M)^2 + (P_2^M)^2 (P_3^M)^2] \\ &\quad + \alpha_{111}^M [(P_1^M)^6 + (P_2^M)^6 + (P_3^M)^6] + \alpha_{112}^M \{ (P_1^M)^4 [(P_2^M)^2 + (P_3^M)^2] \\ &\quad + (P_2^M)^4 [(P_1^M)^2 + (P_3^M)^2] + (P_3^M)^4 [(P_1^M)^2 + (P_2^M)^2] \} \\ &\quad + \alpha_{123}^M (P_1^M)^2 (P_2^M)^2 (P_3^M)^2 \} \\ &\quad + \frac{1}{2} \sum_{i=1}^3 c_f \left[\frac{(1 - c_f) d_i}{[c_f \alpha_i^I + (1 - d_i)(1 - c_f) \alpha_i^M + d_i(1 - c_f) \alpha_d^I]} P_i^I \right] P_i^I \\ &\quad - \frac{1}{2} \sum_{i=1}^3 (1 - c_f) \left[\frac{c_f d_i \alpha_i^I}{\alpha_i^M [c_f \alpha_i^I + (1 - d_i)(1 - c_f) \alpha_i^M + d_i(1 - c_f) \alpha_d^I]} P_i^I \right] P_i^M.\end{aligned}\quad (3.4.40)$$

3.4.3 Total Free Energy for Ferroelectric Composites with Spherical Inclusions

Substitution of Eq. (3.4.39) and Eq. (3.4.17) into Eq. (3.4.2) yields the average Helmholtz free energy of the two-phase ferroelectric composite. The eigenstrains

for the matrix and the inclusion phases are respectively given by

$$\begin{aligned}
 \varepsilon_{11}^{I*} &= Q_{11}^I (P_1^I)^2 + Q_{12}^I [(P_2^I)^2 + (P_3^I)^2], \\
 \varepsilon_{22}^{I*} &= Q_{11}^I (P_2^I)^2 + Q_{12}^I [(P_1^I)^2 + (P_3^I)^2], \\
 \varepsilon_{33}^{I*} &= Q_{11}^I (P_3^I)^2 + Q_{12}^I [(P_1^I)^2 + (P_2^I)^2], \\
 \varepsilon_{12}^{I*} &= Q_{44}^I P_1^I P_2^I, \quad \varepsilon_{13}^{I*} = Q_{44}^I P_1^I P_3^I, \quad \varepsilon_{23}^{I*} = Q_{44}^I P_2^I P_3^I; \\
 \varepsilon_{11}^{M*} &= Q_{11}^M (P_1^M)^2 + Q_{12}^M [(P_2^M)^2 + (P_3^M)^2], \\
 \varepsilon_{22}^{M*} &= Q_{11}^M (P_2^M)^2 + Q_{12}^M [(P_1^M)^2 + (P_3^M)^2], \\
 \varepsilon_{33}^{M*} &= Q_{11}^M (P_3^M)^2 + Q_{12}^M [(P_1^M)^2 + (P_2^M)^2], \\
 \varepsilon_{12}^{I*} &= Q_{44}^M P_1^M P_2^M, \quad \varepsilon_{13}^{M*} = Q_{44}^M P_1^M P_3^M, \quad \varepsilon_{23}^{M*} = Q_{44}^M P_2^M P_3^M.
 \end{aligned} \tag{3.4.41}$$

The expression for the average elastic strain energy density can be obtained by substitution of Eq. (3.4.41) into Eq. (3.4.17), which is rather cumbersome and not presented here. For some specified configurations of the shape of inclusion, however, the expression can be further simplified. Here we present the results for spherical inclusions ($a = b = c$). For materials with cubic symmetry, the elastic moduli tensor $\mathbf{C}$ has the form

$$\mathbf{C} = \begin{bmatrix} C_{11} & C_{12} & C_{12} & 0 & 0 & 0 \\ C_{12} & C_{11} & C_{12} & 0 & 0 & 0 \\ C_{12} & C_{12} & C_{11} & 0 & 0 & 0 \\ 0 & 0 & 0 & C_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & C_{44} & 0 \\ 0 & 0 & 0 & 0 & 0 & C_{44} \end{bmatrix}, \tag{3.4.42}$$

so has the elastic moduli tensor for the inhomogeneous inclusions $\mathbf{C}^*$. According to Chiang (2006), the Eshelby tensor $\mathbf{S}$ for spherical inclusion embedded in the cubic matrix also has the same structure as $\mathbf{C}$ in Eq. (3.4.42), where

$$\begin{aligned}
 S_{11} &= S_{22} = S_{33} = S_{1111}, \\
 S_{12} &= S_{13} = S_{23} = S_{1122}, \\
 S_{44} &= S_{55} = S_{66} = 2S_{1212}.
 \end{aligned} \tag{3.4.43}$$

The three coefficients appearing in Eq. (3.4.43) are not independent, since

$$S_{1111} + 2S_{1212} = 1. \tag{3.4.44}$$

The two independent coefficients S_{1111} and S_{1122} cannot be explicitly expressed. Yet we can obtain them by calculating the following integral

$$S_{ijmn} = \frac{1}{8\pi} C_{pqmn} (\bar{G}_{ipjq} + \bar{G}_{jpiq}), \quad (3.4.45)$$

$$\bar{G}_{ipjq} = \int_0^\pi \sin \theta d\theta \int_0^{2\pi} d\phi (D_{ip}^{-1}(\xi) \xi_j \xi_q),$$

where $D_{ip}^{-1}(\xi)$ denotes the components of the inverse tensor of $D_{ip}(\xi) = C_{ijpq} \xi_j \xi_q$. θ and ϕ denote the spherical coordinates in the ξ -space. For two tensors **A** and **B** having the same structure of **C** in Eq. (3.4.45), it is easy to verify that for the following manipulations the resulting tensor maintains the structure of **C**: 1) $a\mathbf{A} + b\mathbf{B}$, where a and b are arbitrary constants; 2) $\mathbf{AB}$; 3) $\mathbf{A}^{-1}$ when **A** is nonsingular. Since tensors in Eq. (3.4.18) all derive from manipulations falling into the three categories above of **C**, $\mathbf{C}^*$ and **S**, we can conclude that the tensors $\mathbf{N}^{ij}$ in Eq. (3.4.18) are of the same form as **C**, $\mathbf{C}^*$ and **S**. Thus the average elastic energy density can be expressed as

$$\begin{aligned} \bar{F}^{\text{cla}} = & \bar{F}^0 + b_1^I P_2^I P_3^I + b_2^I P_1^I P_3^I + b_3^I P_1^I P_2^I + b_1^I (P_1^I)^2 + b_2^I (P_2^I)^2 + b_3^I (P_3^I)^2 \\ & + b_{11}^I [(P_1^I)^4 + (P_2^I)^4 + (P_3^I)^4] + b_{12}^I [(P_1^I)^2 (P_2^I)^2 + (P_1^I)^2 (P_3^I)^2 \\ & + (P_2^I)^2 (P_3^I)^2] + b_{01}^M P_2^M P_3^M + b_{02}^M P_1^M P_3^M + b_{03}^M P_1^M P_2^M + b_1^M (P_1^M)^2 \\ & + b_2^M (P_2^M)^2 + b_3^M (P_3^M)^2 + b_{11}^M [(P_1^M)^4 + (P_2^M)^4 + (P_3^M)^4] \\ & + b_{12}^M [(P_1^M)^2 (P_2^M)^2 + (P_1^M)^2 (P_3^M)^2 + (P_2^M)^2 (P_3^M)^2] + \gamma_{11} [(P_1^I)^2 (P_1^M)^2 \\ & + (P_2^I)^2 (P_2^M)^2 + (P_3^I)^2 (P_3^M)^2] + \gamma_{12} [(P_1^I)^2 (P_2^M)^2 + (P_2^I)^2 (P_1^M)^2 \\ & + (P_1^I)^2 (P_3^M)^2 + (P_3^I)^2 (P_1^M)^2 + (P_2^I)^2 (P_3^M)^2 + (P_3^I)^2 (P_2^M)^2] \\ & + \gamma_{44} (P_2^I P_3^I P_2^M P_3^M + P_1^I P_3^I P_1^M P_3^M + P_1^I P_2^I P_1^M P_2^M), \end{aligned} \quad (3.4.46)$$

where

$$\begin{aligned} \bar{F}^0 = & N_{11}^{01} [(\varepsilon_{11}^0)^2 + (\varepsilon_{22}^0)^2 + (\varepsilon_{33}^0)^2] + 2N_{12}^{01} (\varepsilon_{11}^0 \varepsilon_{22}^0 + \varepsilon_{11}^0 \varepsilon_{33}^0 + \varepsilon_{22}^0 \varepsilon_{33}^0) \\ & + N_{44}^{01} [(\varepsilon_{23}^0)^2 + (\varepsilon_{13}^0)^2 + (\varepsilon_{12}^0)^2], \end{aligned} \quad (3.4.47)$$

and

$$\begin{aligned}
b_{01}^I &= Q_{44}^I N_{44}^{12} \epsilon_{23}^0, \quad b_{02}^I = Q_{44}^I N_{44}^{12} \epsilon_{13}^0, \quad b_{03}^I = Q_{44}^I N_{44}^{12} \epsilon_{12}^0, \\
b_1^I &= (Q_{11}^I N_{11}^{12} + 2Q_{12}^I N_{12}^{12}) \epsilon_{11}^0 + [(Q_{11}^I + Q_{12}^I) N_{12}^{12} + Q_{12}^I N_{11}^{12}] (\epsilon_{22}^0 + \epsilon_{33}^0), \\
b_2^I &= (Q_{11}^I N_{11}^{12} + 2Q_{12}^I N_{12}^{12}) \epsilon_{22}^0 + [(Q_{11}^I + Q_{12}^I) N_{12}^{12} + Q_{12}^I N_{11}^{12}] (\epsilon_{11}^0 + \epsilon_{33}^0), \\
b_3^I &= (Q_{11}^I N_{11}^{12} + 2Q_{12}^I N_{12}^{12}) \epsilon_{33}^0 + [(Q_{11}^I + Q_{12}^I) N_{12}^{12} + Q_{12}^I N_{11}^{12}] (\epsilon_{11}^0 + \epsilon_{22}^0), \\
b_{11}^I &= (Q_{11}^I)^2 N_{11}^{22} + 4Q_{11}^I Q_{12}^I N_{12}^{22} + 2(Q_{12}^I)^2 (N_{11}^{22} + N_{12}^{22}), \\
b_{12}^I &= 4Q_{12}^I (Q_{12}^I + Q_{11}^I) N_{11}^{22} + 2[(Q_{11}^I)^2 + Q_{11}^I Q_{12}^I + 3(Q_{12}^I)^2] N_{12}^{22} + (Q_{44}^I)^2 N_{44}^{22}, \\
b_{01}^M &= Q_{44}^M N_{44}^{11} \epsilon_{23}^0, \quad b_{02}^M = Q_{44}^M N_{44}^{11} \epsilon_{13}^0, \quad b_{03}^M = Q_{44}^M N_{44}^{11} \epsilon_{12}^0, \\
b_1^M &= (Q_{11}^M N_{11}^{11} + 2Q_{12}^M N_{12}^{11}) \epsilon_{11}^0 + [(Q_{11}^M + Q_{12}^M) N_{12}^{11} + Q_{12}^M N_{11}^{11}] (\epsilon_{22}^0 + \epsilon_{33}^0), \\
b_2^M &= (Q_{11}^M N_{11}^{11} + 2Q_{12}^M N_{12}^{11}) \epsilon_{22}^0 + [(Q_{11}^M + Q_{12}^M) N_{12}^{11} + Q_{12}^M N_{11}^{11}] (\epsilon_{11}^0 + \epsilon_{33}^0), \\
b_3^M &= (Q_{11}^M N_{11}^{11} + 2Q_{12}^M N_{12}^{11}) \epsilon_{33}^0 + [(Q_{11}^M + Q_{12}^M) N_{12}^{11} + Q_{12}^M N_{11}^{11}] (\epsilon_{11}^0 + \epsilon_{22}^0), \\
b_{11}^M &= (Q_{11}^M)^2 N_{11}^{21} + 4Q_{11}^M Q_{12}^M N_{12}^{21} + 2(Q_{12}^M)^2 (N_{11}^{21} + N_{12}^{21}), \\
b_{12}^M &= 4Q_{12}^M (Q_{12}^M + Q_{11}^M) N_{11}^{21} + 2[(Q_{11}^M)^2 + Q_{11}^M Q_{12}^M + 3(Q_{12}^M)^2] N_{12}^{21} + (Q_{44}^M)^2 N_{44}^{21}, \\
\gamma_{11} &= (Q_{11}^I Q_{11}^M + 2Q_{12}^I Q_{12}^M) N_{11}^{ME} + 2(Q_{12}^I Q_{11}^M + Q_{11}^I Q_{12}^M + Q_{12}^I Q_{12}^M) N_{12}^{ME}, \\
\gamma_{44} &= Q_{44}^I Q_{44}^M N_{44}^{ME}, \\
\gamma_{12} &= (Q_{11}^I Q_{12}^M + Q_{12}^I Q_{11}^M + Q_{12}^I Q_{12}^M) N_{11}^{ME} + (Q_{11}^I Q_{11}^M + Q_{11}^I Q_{12}^M + Q_{12}^I Q_{11}^M + 3Q_{12}^I Q_{12}^M) N_{12}^{ME}.
\end{aligned} \tag{3.4.48}$$

The coefficient N_{ij}^{kl} is a corresponding component in the matrix $\mathbf{N}^{kl}$ from Eq. (3.3.38), which is determined ultimately by $\mathbf{S}$, $\mathbf{C}$ and $\mathbf{C}^*$. Eq. (3.4.46) gives the general expression of the elastic strain energy including all the coupling effects for cubic materials.

The total free energy for a ferroelectric composite expressed by polarizations is then given by (Hu and Wang, submitted)

$$\begin{aligned}
\psi &= \bar{F}^0 + b_{01}^I P_2^I P_3^I + b_{02}^I P_1^I P_3^I + b_{03}^I P_1^I P_2^I + b_1^{I*} (P_1^I)^2 + b_2^{I*} (P_2^I)^2 + b_3^{I*} (P_3^I)^2 \\
&\quad + b_{11}^{I*} [(P_1^I)^4 + (P_2^I)^4 + (P_3^I)^4] + b_{12}^{I*} [(P_1^I)^2 (P_2^I)^2 + (P_1^I)^2 (P_3^I)^2 + (P_2^I)^2 (P_3^I)^2] \\
&\quad + c_f a_{111}^I [(P_1^I)^6 + (P_2^I)^6 + (P_3^I)^6] + c_f a_{112}^I \{ (P_1^I)^4 [(P_2^I)^2 + (P_3^I)^2] \\
&\quad + (P_2^I)^4 [(P_1^I)^2 + (P_3^I)^2] + (P_3^I)^4 [(P_1^I)^2 + (P_2^I)^2] \} + c_f a_{123}^I (P_1^I)^2 (P_2^I)^2 (P_3^I)^2 \\
&\quad + b_{01}^M P_2^M P_3^M + b_{02}^M P_1^M P_3^M + b_{03}^M P_1^M P_2^M + b_1^{M*} (P_1^M)^2 + b_2^{M*} (P_2^M)^2 + b_3^{M*} (P_3^M)^2 \\
&\quad + b_{11}^{M*} [(P_1^M)^4 + (P_2^M)^4 + (P_3^M)^4] + b_{12}^{M*} [(P_1^M)^2 (P_2^M)^2 + (P_1^M)^2 (P_3^M)^2 \\
&\quad + (P_2^M)^2 (P_3^M)^2] + (1 - c_f) a_{111}^M [(P_1^M)^6 + (P_2^M)^6 + (P_3^M)^6] \\
&\quad + (1 - c_f) a_{112}^M \{ (P_1^M)^4 [(P_2^M)^2 + (P_3^M)^2] + (P_2^M)^4 [(P_1^M)^2 + (P_3^M)^2] + (P_3^M)^4 [(P_1^M)^2 \\
&\quad + (P_2^M)^2] \} + (1 - c_f) a_{123}^M (P_1^M)^2 (P_2^M)^2 (P_3^M)^2 + \gamma_{01} (P_1^I P_1^M + P_2^I P_2^M + P_3^I P_3^M) \\
&\quad + \gamma_{11} [(P_1^I)^2 (P_1^M)^2 + (P_2^I)^2 (P_2^M)^2 + (P_3^I)^2 (P_3^M)^2] + \gamma_{12} [(P_1^I)^2 (P_2^M)^2 + (P_2^I)^2 (P_1^M)^2 \\
&\quad + (P_1^I)^2 (P_3^M)^2 + (P_3^I)^2 (P_1^M)^2 + (P_2^I)^2 (P_3^M)^2 + (P_3^I)^2 (P_2^M)^2] + \gamma_{44} (P_2^I P_3^I P_2^M P_3^M \\
&\quad + P_1^I P_3^I P_1^M P_3^M + P_1^I P_2^I P_1^M P_2^M),
\end{aligned} \tag{3.4.49}$$

where

$$\begin{aligned}
b_1^{I*} &= b_1^I + c_f a_1^I + \frac{c_f(1-c_f)d_1}{2[c_f\alpha_1^I + (1-d_1)(1-c_f)\alpha_1^M + d_1(1-c_f)\alpha_d^I]}, \\
b_2^{I*} &= b_2^I + c_f a_1^I + \frac{c_f(1-c_f)d_2}{2[c_f\alpha_2^I + (1-d_2)(1-c_f)\alpha_2^M + d_2(1-c_f)\alpha_d^I]}, \\
b_3^{I*} &= b_3^I + c_f a_1^I + \frac{c_f(1-c_f)d_3}{2[c_f\alpha_3^I + (1-d_3)(1-c_f)\alpha_3^M + d_3(1-c_f)\alpha_d^I]}, \\
b_{11}^{I*} &= b_{11}^I + c_f a_{11}^I, \quad b_{12}^{I*} = b_{12}^I + c_f a_{12}^I, \\
b_1^{M*} &= b_1^M + (1-c_f)\alpha_1^M, \quad b_2^{M*} = b_2^M + (1-c_f)\alpha_1^M, \\
b_3^{M*} &= b_3^M + (1-c_f)\alpha_1^M, \quad b_{11}^{M*} = b_{11}^M + (1-c_f)\alpha_{11}^M, \\
b_{12}^{M*} &= b_{12}^M + (1-c_f)\alpha_{12}^M, \\
\gamma_{01} &= -\frac{c_f d_1 \alpha_1^I (1-c_f)}{2\alpha_1^M [c_f \alpha_1^I + (1-d_1)(1-c_f)\alpha_1^M + d_1(1-c_f)\alpha_d^I]}, \\
\gamma_{02} &= -\frac{c_f d_2 \alpha_2^I (1-c_f)}{2\alpha_2^M [c_f \alpha_2^I + (1-d_2)(1-c_f)\alpha_2^M + d_2(1-c_f)\alpha_d^I]}, \\
\gamma_{03} &= -\frac{c_f d_3 \alpha_3^I (1-c_f)}{2\alpha_3^M [c_f \alpha_3^I + (1-d_3)(1-c_f)\alpha_3^M + d_3(1-c_f)\alpha_d^I]}.
\end{aligned} \tag{3.4.50}$$

Eq. (3.4.49) provides the general expression of the Helmholtz free energy of any two-phase ferroelectric composites. For specified boundary conditions, the expression can be simplified.

3.5 Phase Diagrams

3.5.1 Total Free Energy for Ferroelectric Composites with Spherical Inclusions and Equiaxed Strains

Eq. (3.4.49) provides a way of including the micromechanical effects into the analyzing process of the physical properties of ferroelectric composites. In more detail, these micromechanical effects include the material constants, shape, volume fraction, orientation of the inhomogeneities, external mechanical loads (Eq. (3.4.49) is suitable only for analyzing deformation boundary conditions). The combination with the Landau theory of phase transition provides one with the most concise yet macroscopically reasonable method of treating such a complex

problem as ferroelectric composites with microstructures.

By using the energy expression obtained in Eq. (3.4.49), a lot of problems can be treated, e.g., the temperature-eigenstrains phase diagram of ferroelectric composites. Yet, due to the complexity of the problem, the general expression Eq. (3.4.49) should be simplified when applied to any specific physical problems. For example, consider a ferroelectric composite with spherical shapes of inclusions with a set of equiaxed strains $\varepsilon_{11}^0 = \varepsilon_{22}^0 = \varepsilon_{33}^0 = \varepsilon^0$, and $\varepsilon_{12}^0 = \varepsilon_{23}^0 = \varepsilon_{13}^0 = 0$ applied at the boundary. Eq. (3.4.49) becomes

$$\begin{aligned}
 \psi = & \bar{F}^0 + b_1^{I*} [(P_1^I)^2 + (P_2^I)^2 + (P_3^I)^2] + b_{11}^{I*} [(P_1^I)^4 + (P_2^I)^4 + (P_3^I)^4] \\
 & + b_{12}^{I*} [(P_1^I)^2 (P_2^I)^2 + (P_1^I)^2 (P_3^I)^2 + (P_2^I)^2 (P_3^I)^2] + c_f a_{111}^I [(P_1^I)^6 + (P_2^I)^6 + (P_3^I)^6] \\
 & + c_f a_{112}^I \{ (P_1^I)^4 [(P_2^I)^2 + (P_3^I)^2] + (P_2^I)^4 [(P_1^I)^2 + (P_3^I)^2] + (P_3^I)^4 [(P_1^I)^2 + (P_2^I)^2] \} \\
 & + c_f a_{123}^I (P_1^I)^2 (P_2^I)^2 (P_3^I)^2 + b_1^{M*} [(P_1^M)^2 + (P_2^M)^2 + (P_3^M)^2] + b_{11}^{M*} [(P_1^M)^4 + (P_2^M)^4 \\
 & + (P_3^M)^4] + b_{12}^{M*} [(P_1^M)^2 (P_2^M)^2 + (P_1^M)^2 (P_3^M)^2 + (P_2^M)^2 (P_3^M)^2] + (1 - c_f) a_{111}^M [(P_1^M)^6 \\
 & + (P_2^M)^6 + (P_3^M)^6] + (1 - c_f) a_{112}^M \{ (P_1^M)^4 [(P_2^M)^2 + (P_3^M)^2] + (P_2^M)^4 [(P_1^M)^2 + (P_3^M)^2] \\
 & + (P_3^M)^4 [(P_1^M)^2 + (P_2^M)^2] \} + (1 - c_f) a_{123}^M (P_1^M)^2 (P_2^M)^2 (P_3^M)^2 + \gamma_{01} (P_1^I P_1^M + P_2^I P_2^M \\
 & + P_3^I P_3^M) + \gamma_{11} [(P_1^I)^2 (P_1^M)^2 + (P_2^I)^2 (P_2^M)^2 + (P_3^I)^2 (P_3^M)^2] + \gamma_{12} [(P_1^I)^2 (P_2^M)^2 \\
 & + (P_2^I)^2 (P_1^M)^2 + (P_1^I)^2 (P_3^M)^2 + (P_3^I)^2 (P_1^M)^2 + (P_2^I)^2 (P_3^M)^2 + (P_3^I)^2 (P_2^M)^2] \\
 & + \gamma_{44} (P_2^I P_3^I P_2^M P_3^M + P_1^I P_3^I P_1^M P_3^M + P_1^I P_2^I P_1^M P_2^M),
 \end{aligned} \tag{3.5.1}$$

where

$$\begin{aligned}
 b_1^{I*} &= b_1^I + c_f a_1^I + \frac{c_f (1 - c_f)}{2[3c_f \alpha^I + 2(1 - c_f) \alpha^M + (1 - c_f) \alpha_d^I]}, \\
 b_{11}^{I*} &= b_{11}^I + c_f a_{11}^I, \quad b_{12}^{I*} = b_{12}^I + c_f a_{12}^I, \\
 b_1^{M*} &= b_1^M + (1 - c_f) a_1^M, \\
 b_{11}^{M*} &= b_{11}^M + (1 - c_f) a_{11}^M, \quad b_{12}^{M*} = b_{12}^M + (1 - c_f) a_{12}^M.
 \end{aligned} \tag{3.5.2}$$

and

$$\begin{aligned}
b_1^I &= (Q_{11}^I + 2Q_{12}^I)(N_{11}^{12} + 2N_{12}^{12})\epsilon^0, \\
b_{11}^I &= (Q_{11}^I)^2 N_{11}^{22} + 4Q_{11}^I Q_{12}^I N_{12}^{22} + 2(Q_{12}^I)^2 (N_{11}^{22} + N_{12}^{22}), \\
b_{12}^I &= 4Q_{12}^I (Q_{12}^I + Q_{11}^I) N_{11}^{22} + 2[(Q_{11}^I)^2 + Q_{11}^I Q_{12}^I + 3(Q_{12}^I)^2] N_{12}^{22} + (Q_{44}^I)^2 N_{44}^{22}, \\
b_1^M &= (Q_{11}^M + 2Q_{12}^M)(N_{11}^{11} + 2N_{12}^{11})\epsilon^0, \\
b_{11}^M &= (Q_{11}^M)^2 N_{11}^{21} + 4Q_{11}^M Q_{12}^M N_{12}^{21} + 2(Q_{12}^M)^2 (N_{11}^{21} + N_{12}^{21}), \\
b_{12}^M &= 4Q_{12}^M (Q_{12}^M + Q_{11}^M) N_{11}^{21} + 2[(Q_{11}^M)^2 + Q_{11}^M Q_{12}^M + 3(Q_{12}^M)^2] N_{12}^{21} + (Q_{44}^M)^2 N_{44}^{21}, \quad (3.5.3) \\
\gamma_{01} &= -\frac{c_f(1-c_f)\alpha^I}{2\alpha_i^M[3c_f\alpha^I + 2(1-c_f)\alpha^M + (1-c_f)\alpha_d^I]}, \quad \gamma_{44} = Q_{44}^I Q_{44}^M N_{44}^{ME}, \\
\gamma_{11} &= (Q_{11}^I Q_{11}^M + 2Q_{12}^I Q_{12}^M) N_{11}^{ME} + 2(Q_{12}^I Q_{11}^M + Q_{11}^I Q_{12}^M + Q_{12}^I Q_{12}^M) N_{12}^{ME}, \\
\gamma_{12} &= (Q_{11}^I Q_{12}^M + Q_{12}^I Q_{11}^M + Q_{12}^I Q_{12}^M) N_{11}^{ME} + (Q_{11}^I Q_{12}^M + Q_{12}^I Q_{11}^M + Q_{12}^I Q_{12}^M) N_{12}^{ME} \\
&\quad + 3Q_{12}^I Q_{12}^M N_{12}^{ME}.
\end{aligned}$$

In the deduction of Eq. (3.5.1), it is assumed that $\alpha_1^M = \alpha_2^M = \alpha_3^M = \alpha^M$, and $\alpha_1^I = \alpha_2^I = \alpha_3^I = \alpha^I$.

3.5.2 Phase Diagrams and Total Polarizations

A specific calculation is provided using Eq. (3.5.1) for a 0-3 type ferroelectric composite with PbTiO_3 as its matrix containing randomly distributed spherical ferroelectric BaTiO_3 inclusions. The total volume fraction of the inclusion phase is denoted by c_f . Corresponding material parameters are listed in Table 3.1. The

electric polarizations that minimize the energy in Eq. (3.5.1) satisfy $\frac{\partial \psi}{\partial P_i^I} = 0$ ($i = 1, 2, 3$)

and $\frac{\partial \psi}{\partial P_i^M} = 0$ ($i = 1, 2, 3$), which gives for $i = 1$

$$\begin{aligned}
&P_1^I \{2b_1^{I*} + 4b_{11}^{I*}(P_1^I)^2 + 2b_{12}^{I*}[(P_2^I)^2 + (P_3^I)^2] + 6c_f a_{111}^I (P_1^I)^4 \\
&+ c_f a_{112}^I \{4(P_1^I)^2[(P_2^I)^2 + (P_3^I)^2] + 2[(P_2^I)^4 + (P_3^I)^4]\} \\
&+ 2c_f a_{123}^I (P_2^I)^2 (P_3^I)^2 + 2\gamma_{11}(P_1^M)^2 + 2\gamma_{12}[(P_2^M)^2 + (P_3^M)^2]\} \\
&+ \gamma_{44}(P_3^I P_1^M P_3^M + P_2^I P_1^M P_2^M) = 0,
\end{aligned} \quad (3.5.4)$$

and

$$\begin{aligned}
& P_1^M \{2b_1^{M*} + 4b_{11}^{M*} (P_1^M)^2 + 2b_{12}^{M*} [(P_2^M)^2 + (P_3^M)^2] + 6(1-c_f)a_{111}^M (P_1^M)^4 \\
& + (1-c_f)a_{112}^M [4(P_1^M)^2 [(P_2^M)^2 + (P_3^M)^2] + 2[(P_2^M)^4 + (P_3^M)^4]] \} \\
& + 2(1-c_f)a_{123}^M (P_2^M)^2 (P_3^M)^2 + 2\gamma_{11} (P_1^I)^2 + 2\gamma_{12} [(P_2^I)^2 + (P_3^I)^2] \} \\
& + \gamma_{44} (P_1^I P_3^I P_3^M + P_1^I P_2^I P_2^M) = 0.
\end{aligned} \tag{3.5.5}$$

The equations for $i=2$ and $i=3$ take similar forms to Eqs. (3.5.4) and Eq. (3.5.5). Due to the rotational properties of the equations, the following 20 phase states are possible for minimizing the energy expression in Eq. (3.5.1), listed as:

- 1) $P^I = 0, P^M = 0$; (p phase)
- 2) $P_1^I \neq 0, P_2^I = P_3^I = 0, P^M = 0$; (a_1 phase)
- 3) $P_1^I = P_2^I \neq 0, P_3^I = 0, P^M = 0$; (a_{11} phase)
- 4) $P_1^I = P_2^I = P_3^I \neq 0, P^M = 0$; (r phase)
- 5) $P^I = 0, P_1^M \neq 0, P_2^M = P_3^M = 0$; (b_1 phase)
- 6) $P_1^I \neq 0, P_2^I = P_3^I = 0, P_1^M \neq 0, P_2^M = P_3^M = 0$; ($a_1 - b_1$ phase)
- 7) $P_1^I \neq 0, P_2^I = P_3^I = 0, P_2^M \neq 0, P_1^M = P_3^M = 0$; ($a_1 - b_2$ phase)
- 8) $P_1^I \neq 0, P_2^I \neq 0, P_3^I = 0, P_1^M \neq 0, P_2^M = P_3^M = 0$; ($a_{12} - b_1$ phase)
- 9) $P_1^I = P_2^I \neq 0, P_3^I = 0, P_3^M \neq 0, P_1^M = P_2^M = 0$; ($a_{11} - b_3$ phase)
- 10) $P_1^I \neq 0, P_2^I = P_3^I \neq 0, P_1^M \neq 0, P_2^M = P_3^M = 0$; ($r - b_1$ phase)
- 11) $P^I = 0, P_1^M = P_2^M \neq 0, P_3^M = 0$; (b_{11} phase)
- 12) $P_1^I \neq 0, P_2^I = P_3^I = 0, P_1^M \neq 0, P_2^M \neq 0, P_3^M = 0$; ($a_1 - b_{12}$ phase)
- 13) $P_3^I \neq 0, P_1^I = P_2^I = 0, P_1^M = P_2^M \neq 0, P_3^M = 0$; ($a_3 - b_{11}$ phase)
- 14) $P_1^I = P_2^I \neq 0, P_3^I = 0, P_1^M = P_2^M \neq 0, P_3^M = 0$; ($a_{11} - b_{11}$ phase)
- 15) $P_1^I \neq 0, P_2^I \neq 0, P_3^I = 0, P_1^M \neq 0, P_3^M \neq 0, P_2^M = 0$; ($a_{12} - b_{13}$ phase)
- 16) $P_1^I = P_2^I \neq 0, P_3^I \neq 0, P_1^M = P_2^M \neq 0, P_3^M = 0$; ($r - b_{11}$ phase)
- 17) $P^I = 0, P_1^M = P_2^M = P_3^M \neq 0$; ($\tilde{r}$ phase)
- 18) $P_1^I \neq 0, P_2^I = P_3^I = 0, P_1^M \neq 0, P_2^M = P_3^M \neq 0$; ($a_1 - \tilde{r}$ phase)
- 19) $P_1^I = P_2^I \neq 0, P_3^I = 0, P_1^M = P_2^M \neq 0, P_3^M \neq 0$; ($a_{11} - \tilde{r}$ phase)
- 20) $P_1^I = P_2^I = P_3^I \neq 0, P_1^M = P_2^M = P_3^M \neq 0$; ($r - \tilde{r}$ phase)

One should bear in mind that for phases such as the $a_1 - b_2$ phase, the name describes an orthogonal relation between the matrix polarization $\mathbf{P}^M$ and the inclusion polarization $\mathbf{P}^I$, and the fact that only one non-zero component exists for $\mathbf{P}^I$ and for $\mathbf{P}^M$. Yet $\mathbf{P}^I$ is not restricted to x_1 and $\mathbf{P}^M$ is not restricted to x_2 .

In order to check the effect of the depolarization fields (i.e. to check the effect

of the term $\bar{F}^{\text{field}}$ on the electric polarizations), two calculations are done: one without the part $\bar{F}^{\text{field}}$ in total energy and the other with the part $\bar{F}^{\text{field}}$ in total energy.

In Fig. 3.1a, the $\varepsilon^0 - T$ phase diagram for PTO-BTO composite with $c_f = 0.5$ is plotted in a range of ε^0 ($-0.02 \sim 0.06$). In Fig. 3.1b, $\varepsilon^0 - T$ phase diagrams for PTO-BTO composite with varying c_f (0.2, 0.5, 0.7) are plotted in a range of ε^0 ($-0.02 \sim 0.06$). The effect of the depolarization fields is neglected in this calculation. It is shown that only six phases appear in the diagram, corresponding to

- 1) $P^I = 0, P^M = 0$; (p phase)
- 2) $P^I = 0, P_1^M \neq 0, P_2^M = P_3^M = 0$; (b_1 phase)
- 3) $P_1^I \neq 0, P_2^I = P_3^I = 0, P_2^M \neq 0, P_1^M = P_3^M = 0$; ($a_1 - b_2$ phase)
- 4) $P_1^I = P_2^I \neq 0, P_3^M = 0, P_3^M \neq 0, P_1^M = P_2^M = 0$; ($a_{11} - b_3$ phase)
- 5) $P_1^I = P_2^I \neq 0, P_3^I = 0, P_1^M = P_2^M \neq 0, P_3^M \neq 0$; ($a_{11} - \tilde{r}$ phase)
- 6) $P_1^I = P_2^I = P_3^I \neq 0, P_1^M = P_2^M = P_3^M \neq 0$. ($r - \tilde{r}$ phase)

Table 3.1 Parameters for the PbTiO_3 and BaTiO_3

BaTiO ₃	PbTiO ₃
Landau coefficients (Pertsev, 1998):	Landau coefficients (Pertsev, 1998):
$a_1 = 3.3 \times 10^5 (T - 110) \text{ (C}^{-2} \cdot \text{m}^2 \cdot \text{N)}$	$a_1 = 3.8 \times 10^5 (T - 479) \text{ (C}^{-2} \cdot \text{m}^2 \cdot \text{N)}$
$a_{11} = 3.6 \times 10^6 (T - 175) \text{ (C}^{-4} \cdot \text{m}^6 \cdot \text{N)}$	$a_{11} = -7.3 \times 10^7 \text{ (C}^{-4} \cdot \text{m}^6 \cdot \text{N)}$
$a_{12} = 4.9 \times 10^8 \text{ (C}^{-4} \cdot \text{m}^6 \cdot \text{N)}$	$a_{12} = 7.5 \times 10^8 \text{ (C}^{-4} \cdot \text{m}^6 \cdot \text{N)}$
$a_{111} = 6.6 \times 10^9 \text{ (C}^{-6} \cdot \text{m}^{10} \cdot \text{N)}$	$a_{111} = 2.6 \times 10^8 \text{ (C}^{-6} \cdot \text{m}^{10} \cdot \text{N)}$
$a_{112} = 2.9 \times 10^9 \text{ (C}^{-6} \cdot \text{m}^{10} \cdot \text{N)}$	$a_{112} = 6.1 \times 10^8 \text{ (C}^{-6} \cdot \text{m}^{10} \cdot \text{N)}$
$a_{123} = 7.6(T - 120) \times 10^7 + 4.4 \times 10^{10} \text{ (C}^{-6} \cdot \text{m}^{10} \cdot \text{N)}$	$a_{123} = -3.7 \times 10^9 \text{ (C}^{-6} \cdot \text{m}^{10} \cdot \text{N)}$
Electrostrictive coefficients (Li et al., 2005):	Electrostrictive coefficients (Li et al., 2005):
$Q_{11} = 0.10 \text{ (C}^{-2} \cdot \text{m}^4), Q_{12} = -0.034 \text{ (C}^{-2} \cdot \text{m}^4),$ $Q_{44} = 0.029 \text{ (C}^{-2} \cdot \text{m}^4)$	$Q_{11} = 0.089 \text{ (C}^{-2} \cdot \text{m}^4), Q_{12} = -0.026 \text{ (C}^{-2} \cdot \text{m}^4),$ $Q_{44} = 0.0675 \text{ (C}^{-2} \cdot \text{m}^4)$
Elastic constants:	Elastic constants:
$C_{11}^* = 1.755 \times 10^{11} \text{ (N} \cdot \text{m}^{-2})$	$C_{11} = 1.746 \times 10^{11} \text{ (N} \cdot \text{m}^{-2})$
$C_{12}^* = 0.846 \times 10^{11} \text{ (N} \cdot \text{m}^{-2})$	$C_{12} = 0.794 \times 10^{11} \text{ (N} \cdot \text{m}^{-2})$
$C_{44}^* = 1.08 \times 10^{11} \text{ (N} \cdot \text{m}^{-2})$	$C_{44} = 1.11 \times 10^{11} \text{ (N} \cdot \text{m}^{-2})$
Dielectric constants (Wang and Woo, 2005):	Dielectric constants (Wang and Woo, 2005):
$\alpha^I = 200\alpha_0$	$\alpha^M = 210\alpha_0$
$\alpha_d^I = 50\alpha_0$	

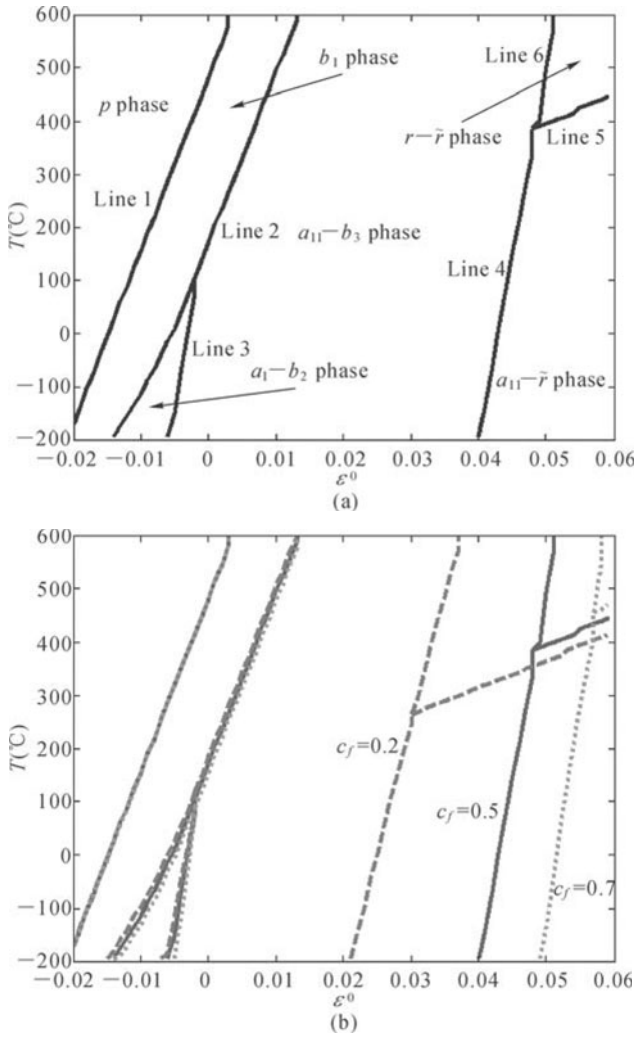


Fig. 3.1 ϵ^0 - T phase diagram for the 0-3 type ferroelectric composite made of PbTiO_3 matrix and spherical BaTiO_3 inclusions. Effect of the depolarization fields is neglected in this calculation. In (a), the phase diagram is plotted in a range of ϵ^0 (-0.02~0.06), and the volume fraction of PbTiO_3 inclusions is $c_f = 0.5$. All possible phases for the composite are tagged. In (b), the phase diagrams are plotted in a range of ϵ^0 (-0.02~0.06), but for different values of c_f ($c_f = 0.2$ (dashed lines), $c_f = 0.5$ (lines), and $c_f = 0.7$ (dotted lines))

In Fig. 3.2a, the ϵ^0 - T phase diagram for PTO-BTO composite with $c_f = 0.5$ is plotted in a range of ϵ^0 (-0.02~0.06). In Fig. 3.2b, ϵ^0 - T phase diagrams for PTO-BTO composite with varying c_f (0.2, 0.5, 0.7) are plotted in a range of ϵ^0 (-0.02~0.06). The effect of the depolarization fields is considered in this calculation. It is shown that only six phases appear in the diagram, corresponding to

- 1) $P^I = 0, P^M = 0$; (p phase)
- 2) $P_1^I \neq 0, P_2^I = P_3^I = 0, P_1^M \neq 0, P_2^M = P_3^M = 0$; ($a_1 - b_1$ phase)
- 3) $P_1^I = P_2^I \neq 0, P_3^I = 0, P_1^M \neq 0, P_2^M = P_3^M = 0$; ($a_{12} - b_1$ phase)
- 4) $P_1^I = P_2^I = P_3^I \neq 0, P_1^M \neq 0, P_2^M = P_3^M = 0$; ($r - b_1$ phase)
- 5) $P_1^I = P_2^I \neq 0, P_3^I = 0, P_1^M = P_2^M = P_3^M \neq 0$; ($a_{11} - \tilde{r}$ phase)
- 6) $P_1^I = P_2^I = P_3^I \neq 0, P_1^M = P_2^M = P_3^M \neq 0$. ($r - \tilde{r}$ phase)

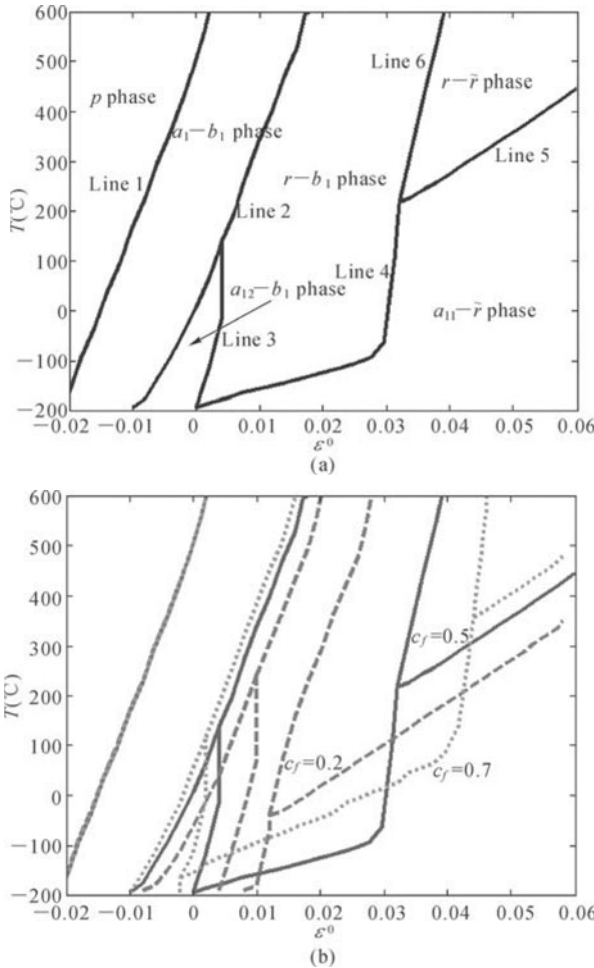


Fig. 3.2 ε^0 - T phase diagram for the 0-3 type ferroelectric composite made of PbTiO_3 matrix and spherical BaTiO_3 inclusions. Effect of the depolarization fields is considered in this calculation. In (a), the phase diagram is plotted in a range of ε^0 (-0.02~0.06), and the volume fraction of PbTiO_3 inclusions is $c_f = 0.5$. All possible phases for the composite are tagged. In (b), the phase diagrams are plotted in a range of ε^0 (-0.02~0.06), but for different values of c_f ($c_f = 0.2$ (dashed lines), $c_f = 0.5$ (lines), and $c_f = 0.7$ (dotted lines))

To explain the transition lines and phase patterns in Fig. 3.1 and Fig. 3.2, Eq. (3.5.1) is studied from several different aspects:

i) Neglecting effects of depolarization fields, interaction between $\mathbf{P}^I$ and $\mathbf{P}^M$ is described by the three terms:

$$\begin{aligned} & \gamma_{11}[(P_1^I)^2(P_1^M)^2 + (P_2^I)^2(P_2^M)^2 + (P_3^I)^2(P_3^M)^2], \\ & \gamma_{44}(P_2^I P_3^I P_2^M P_3^M + P_1^I P_3^I P_1^M P_3^M + P_1^I P_2^I P_1^M P_2^M), \\ & \gamma_{12}[(P_1^I)^2(P_2^M)^2 + (P_2^I)^2(P_1^M)^2 + (P_1^I)^2(P_3^M)^2 + (P_3^I)^2(P_1^M)^2 + (P_2^I)^2(P_3^M)^2 + (P_3^I)^2(P_2^M)^2]. \end{aligned}$$

For PTO-BTO composites with $c_f = 0.5$, the three corresponding coefficients are given by $\gamma_{11} = 1.813 \times 10^8$, $\gamma_{12} = -5.2958 \times 10^7$, $\gamma_{44} = 4.2814 \times 10^7$. Since γ_{11} is positive and γ_{12} is negative, the inclusion polarization vector $\mathbf{P}^I$ and the matrix polarization vector $\mathbf{P}^M$ tend to take directions orthogonal to each other. Thus, for $c_f = 0.5$, $a_1 - b_2$ phase is always more stable than $a_1 - b_1$ phase, and $a_{11} - b_3$ phase is always more stable than $a_{12} - b_1$ phase. The proof is simple: if a set of $\{P_1^I, P_1^M\}$ minimizes the energy expression in $a_1 - b_1$ phase, letting the direction of $\mathbf{P}^M$ switch from x_1 to x_3 , the total energy of the composites is changed by $(\gamma_{12} - \gamma_{11})(P_1^I)^2(P_1^M)^2$, which is negative according to the values of γ_{11} and γ_{12} given above. Therefore, the state after the switching is more stable than the original one. The term with γ_{44} influences the sign of the polarization components (i.e. the sign of P_i^I and P_i^M) at certain phases (e.g. $r - \tilde{r}$ phase) in order to retain a negative value of the term $\gamma_{44}(P_2^I P_3^I P_2^M P_3^M + P_1^I P_3^I P_1^M P_3^M + P_1^I P_2^I P_1^M P_2^M)$.

ii) The 3-D clamping condition and the mutual elastic interaction between the inclusion and the matrix renormalize the second- and fourth-order terms in the energy expression. The expressions of the renormalized second-order coefficients and the values of the renormalized fourth-order coefficients for $c_f = 0.5$ are given in Table 3.2. The paraelectric-ferroelectric phase transition temperature is determined by $b_1^{M*} = 0$ for the matrix and by $b_1^{I*} = 0$ for the inclusion. Comparing the expression of b_1^{M*} and b_1^{I*} , we find that in a regular range of ε^0 (e.g. $-0.1 \sim 0.1$), the transition temperature of the PTO matrix is always larger than the BTO inclusions. Therefore, P_i^M always appears first. Positive equiaxed strain ε^0 and low temperature accelerates the emergence of the ferroelectric phase. From $b_1^{M*} = 0$ a linear relation between the transition temperature T and the applied equiaxed strain ε^0 is obtained, with a gradient of 3.22×10^4 . This transition line is marked Line 1 in Fig. 3.1a.

Table 3.2 Renormalized second-order coefficients and fourth-order coefficients for $c_f = 0.5$

Coefficients	Values
$b_1^{I*} (\text{C}^{-2} \cdot \text{m}^2 \cdot \text{N})$	$1.65 \times 10^5 (T - 110 - 3.22 \times 10^4 \varepsilon^0)$
$b_{11}^{I*} (\text{C}^{-4} \cdot \text{m}^6 \cdot \text{N})$	$1.8 \times 10^6 (T - 56.6)$
$b_{12}^{I*} (\text{C}^{-4} \cdot \text{m}^6 \cdot \text{N})$	2.42×10^8
$b_1^{M*} (\text{C}^{-2} \cdot \text{m}^2 \cdot \text{N})$	$1.9 \times 10^5 (T - 479 - 3.26 \times 10^4 \varepsilon^0)$
$b_{11}^{M*} (\text{C}^{-4} \cdot \text{m}^6 \cdot \text{N})$	1.29×10^8
$b_{12}^{M*} (\text{C}^{-4} \cdot \text{m}^6 \cdot \text{N})$	4.50×10^8

iii) The transition line illustrating the appearance of inclusion polarization P_i^I is tagged Line 2 in Fig. 3.1a. One should notice that this line is composed of two sections, one denoting a transition from b_1 phase to $a_1 - b_2$ phase, and the other denoting a transition from b_1 phase to $a_{11} - b_3$ phase. Assuming a continuous change of P_i^M in the transition from b_1 phase to $a_1 - b_2$ phase, the energy difference between the two phases is given by

$$[\Delta\psi]_{b_1}^{a_1-b_2} = b_1^{I*} (P_1^I)^2 + b_{11}^{I*} (P_1^I)^4 + c_f a_{111}^I (P_1^I)^6 + \gamma_{12} (P_1^I)^2 (P_2^M)^2. \quad (3.5.6)$$

In the deduction of Eq. (3.5.6), we have assumed that P_1^M in the b_1 phase equals P_2^M in the $a_1 - b_2$ phase. Similarly, the energy difference between the b_1 phase and $a_{11} - b_3$ phase is given by

$$\begin{aligned} [\Delta\psi]_{b_1}^{a_{11}-b_3} = & 2b_1^{I*} (P_1^I)^2 + (2b_{11}^{I*} + b_{12}^{I*}) (P_1^I)^4 \\ & + 2c_f (a_{111}^I + a_{112}^I) (P_1^I)^6 + 2\gamma_{12} (P_1^I)^2 (P_3^M)^2. \end{aligned} \quad (3.5.7)$$

The governing equations for the two transitions to occur are the same, given by

$$b_1^{I*} + \gamma_{12} (P_2^M)^2 = 0. \quad (3.5.8)$$

This explains why the two sections of Line 2 are well connected. The second term in Eq. (3.5.8) is the source of the nonlinear character of Line 2 in Fig. 3.1a, since P_2^M in Eq. (3.5.8) is a quantity that varies with temperature and strains.

Now we try to find the critical point (Point A in Fig. 3.1a) where the transition from the b_1 phase to $a_1 - b_2$ phase switches to the $a_{11} - b_3$ phase. Next to transition points, we have

$$\delta = b_1^{I*} + \gamma_{12} (P_2^M)^2, \quad (3.5.9)$$

where δ is a small negative quantity. From Eq. (3.5.6) and Eq. (3.5.7) we can solve the polarizations that minimize the corresponding energy expression as

$$(P_1^I)^2[a_1 - b_2] = \frac{-b_{11}^* \pm \sqrt{(b_{11}^*)^2 - 3c_f a_{111} \delta}}{3c_f a_{111}} \approx \frac{-\delta}{2b_{11}^*}, \quad (3.5.10)$$

$$\begin{aligned} (P_1^I)^2[a_{11} - b_3] &= \frac{-(2b_{11}^{I*} + b_{12}^{I*}) \pm \sqrt{(2b_{11}^{I*} + b_{12}^{I*})^2 - 12c_f(a_{111}^I + a_{112}^I)\delta}}{6c_f(a_{111}^I + a_{112}^I)} \\ &\approx \frac{-\delta}{2b_{11}^{I*} + b_{12}^{I*}}. \end{aligned} \quad (3.5.11)$$

Substitution of Eqs. (3.5.10) and (3.5.11) into Eqs. (3.5.6) and (3.5.7) yields

$$[\Delta\psi]_{b_1}^{a_1 - b_2} = -\frac{\delta^2}{4b_{11}^{I*}} + O(\delta^3), \quad (3.5.12)$$

$$[\Delta\psi]_{b_1}^{a_{11} - b_3} = -\frac{\delta^2}{2b_{11}^{I*} + b_{12}^{I*}} + O(\delta^3), \quad (3.5.13)$$

where $O(\delta^3)$ means high order terms (≥ 3) of δ . Thus the critical point is reached when

$$2b_{11}^{I*} - b_{12}^{I*} = 0. \quad (3.5.14)$$

Substituting values listed in Table 3.2, we finally have $T_A = 123.8^\circ\text{C}$.

iv) According to the calculation, the phase transitions described by Lines 3, 4, 5 and 6 in Fig. 3.1a are all of first order. Appearance of the $r - \tilde{r}$ phase is promoted by the interaction term $\gamma_{44}(P_2^I P_3^I P_2^M P_3^M + P_1^I P_3^I P_1^M P_3^M + P_1^I P_2^I P_1^M P_2^M)$.

v) In order to examine the effect of the volume fraction c_f , we first calculate its influence on the elastic coupling stiffness tensors N^{ij} , which are related to the energy expression by Eq. (3.5.3). The expressions for these tensors are given in Eq. (3.4.18). For PTO-BTO composites, we plot the relation of the components of these tensors with the volume fraction c_f in Fig. 3.3. From Eq. (3.4.18) and related expressions, it is easy to discover that components of the elastic coupling stiffness tensors N^{ij} are high order functions of c_f . Yet, when the stiffness tensors of the matrix material $\mathbf{C}$ and the inclusion material $\mathbf{C}^*$ are close to each other (in our calculation, the PTO matrix and the BTO inclusions satisfy this condition well), the components of the elastic coupling stiffness tensors $\mathbf{N}^{11}$ and $\mathbf{N}^{12}$ can be approximated with a linear function of c_f , written as

$$N_{11}^{11} = d_1(1 - c_f), \quad (3.5.15a)$$

$$N_{12}^{11} = d_2(1 - c_f), \quad (3.5.15b)$$

$$N_{44}^{11} = d_3(1 - c_f), \quad (3.5.15c)$$

$$N_{11}^{12} = d_4 c_f, \quad (3.5.15d)$$

$$N_{12}^{12} = d_5 c_f, \quad (3.5.15e)$$

$$N_{44}^{12} = d_6 c_f, \quad (3.5.15f)$$

where d_1 as a function of c_f varies from -1.7458×10^{11} to -1.7442×10^{11} , d_2 varies from -8.0542×10^{10} to -7.9496×10^{10} , d_3 varies from -1.1104×10^{11} to -1.1047×10^{11} , d_4 varies from -1.7458×10^{11} to -1.7442×10^{11} , d_5 varies from -7.9237×10^{10} to -7.8212×10^{10} , d_6 varies from -1.1178×10^{11} to -1.1118×10^{11} . The components of the elastic coupling stiffness tensors $\mathbf{N}^{21}$, $\mathbf{N}^{22}$ and $\mathbf{N}^{Mf}$ can be approximated with a quadratic function of c_f , written as

$$N_{11}^{21} = (d_{11} c_f + d_{01})(1 - c_f), \quad (3.5.16a)$$

$$N_{12}^{21} = (d_{12} c_f + d_{02})(1 - c_f), \quad (3.5.16b)$$

$$N_{44}^{21} = (d_{13} c_f + d_{03})(1 - c_f), \quad (3.5.16c)$$

$$N_{11}^{22} = (d_{14} c_f + d_{04}) c_f, \quad (3.5.16d)$$

$$N_{12}^{22} = (d_{15} c_f + d_{05}) c_f, \quad (3.5.16e)$$

$$N_{44}^{22} = (d_{16} c_f + d_{06}) c_f, \quad (3.5.16f)$$

$$N_{11}^{Mf} = d_{17} c_f (1 - c_f), \quad (3.5.16g)$$

$$N_{12}^{Mf} = d_{18} c_f (1 - c_f), \quad (3.5.16h)$$

$$N_{44}^{Mf} = d_{19} c_f (1 - c_f), \quad (3.5.16i)$$

where $d_{01} = 8.730 \times 10^{10}$, $d_{02} = 3.9668 \times 10^{10}$, $d_{03} = 5.5561 \times 10^{10}$, $d_{04} = 2.4442 \times 10^{10}$, $d_{05} = 6.4041 \times 10^{10}$, $d_{06} = 1.2335 \times 10^{10}$, and d_{11} as a function of c_f varies from -6.2599×10^{10} to -6.2497×10^{10} , d_{12} varies from -3.1794×10^{10} to -3.0837×10^{10} , d_{13} varies from -4.4487×10^{10} to -4.4037×10^{10} , d_{14} varies from 6.2800×10^{10} to 6.2924×10^{10} , d_{15} varies from 3.2816×10^{10} to 3.3768×10^{10} , d_{16} varies from 4.2989×10^{10} to 4.3439×10^{10} , d_{17} varies from 1.2537×10^{11} to 1.255×10^{11} , d_{18} varies from 6.4075×10^{10} to 6.5167×10^{10} , d_{19} varies from 8.723×10^{10} to 8.7744×10^{10} .

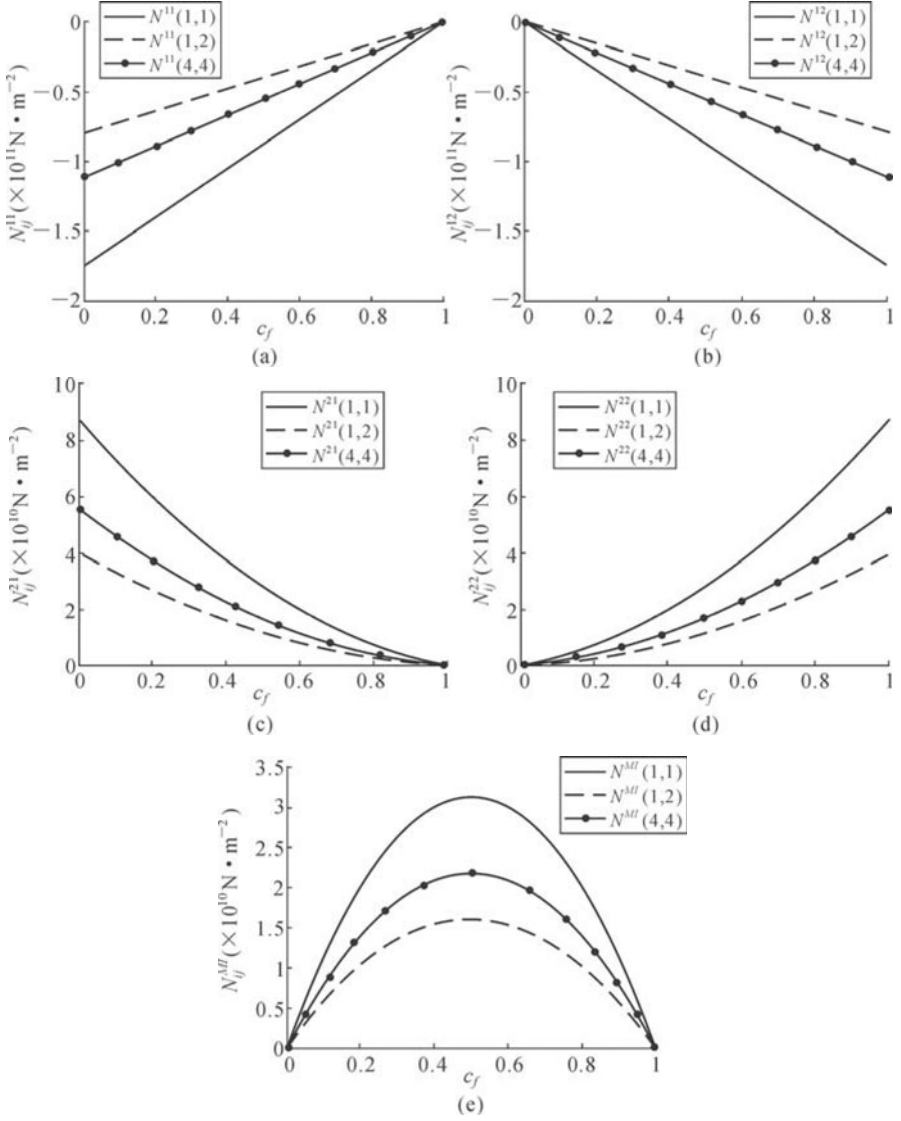


Fig. 3.3 Components of the elastic coupling stiffness tensors N_{ij}^{11} (a), N_{ij}^{12} (b), N_{ij}^{21} (c), N_{ij}^{22} (d), N_{ij}^{Ml} (e) versus volume fraction c_f

Substituting Eqs. (3.5.15) and (3.5.16) into Eqs. (3.5.2) and (3.5.3), we have

$$b_1^{I*} = c_f [a_1^I + (Q_{11}^I + 2Q_{12}^I)(d_4 + 2d_5)] + \frac{c_f(1-c_f)}{2[3c_f\alpha^I + 2(1-c_f)\alpha^M + (1-c_f)\alpha_d^I]}, \quad (3.5.17a)$$

$$b_1^{M*} = (1 - c_f)[a_1^M + (Q_{11}^M + 2Q_{12}^M)(d_1 + 2d_2)]. \quad (3.5.17b)$$

From Eq. (3.5.17b), we can see that for a second-order paraelectric-ferroelectric phase transition of PTO-BTO composite, the transition T - ε_0 relation is merely affected by the change in c_f . This phenomenon is clearly shown in Fig. 3.1b by the overlapping Line 1. On the other hand, coefficients such as b_{11}^{I*} , b_{11}^{M*} are quadratic functions of c_f . Thus, phase transitions determined by these coefficients would vary dramatically with an increase in c_f as demonstrated by the shift of Lines 4, 5 and 6 in Fig. 3.1b.

vi) On the basis of the above discussion, we try to analyze the effect of depolarization fields, which is calculated for PTO-BTO composites and plotted in Fig. 3.2. The depolarization fields have two effects on the energy expression Eq. (3.5.11): one is a renormalization of the coefficient b_1^{I*} by adding a positive

term $\frac{c_f(1 - c_f)}{2[3c_f\alpha^I + 2(1 - c_f)\alpha^M + (1 - c_f)\alpha_d^I]}$ and the other is the interaction term $\gamma_{01}(P_1^I P_1^M + P_2^I P_2^M + P_3^I P_3^M)$. From the discussion above, we learn that b_1^{I*} is closely related to the appearance of P_i^I . The positive term $\frac{c_f(1 - c_f)}{2[3c_f\alpha^I + 2(1 - c_f)\alpha^M + (1 - c_f)\alpha_d^I]}$

in b_1^{I*} depresses this appearance, and thus Line 2 is downward shifted in Fig. 3.2a compared with the same transition line in Fig. 3.1a. Since

$\gamma_{01} = -\frac{c_f(1 - c_f)\alpha^I}{2\alpha_i^M[3c_f\alpha^I + 2(1 - c_f)\alpha^M + (1 - c_f)\alpha_d^I]}$ is always negative, the term

$\gamma_{01}(P_1^I P_1^M + P_2^I P_2^M + P_3^I P_3^M)$ encourages coexistence of P_i^I and P_i^M ($i=1,2,3$). By comparing the phases in Fig. 3.1a and Fig. 3.2a, it appears that the b_1 phase is replaced by $a_1 - b_1$ phase, $a_1 - b_2$ phase is replaced by $a_{12} - b_1$ phase, $a_{11} - b_3$ phase is replaced by $r - b_1$ phase. These changes are all due to the introduction of the term $\gamma_{01}(P_1^I P_1^M + P_2^I P_2^M + P_3^I P_3^M)$. The energy change between the $a_{11} - b_3$ phase and $r - b_1$ phase can be written in the form of

$$[\Delta\psi]_{a_{11}-b_3}^{r-b_1} = r_{01}P_1^M P_1^I + O[(P_1^I)^2]. \quad (3.5.18)$$

In the deduction process of Eq. (3.5.18), we have assumed here that P_1^M in the $r - b_1$ phase equals P_3^M in the $a_{11} - b_3$ phase, and that P_2^I and P_3^I in the $r - b_1$ phase equal P_1^I and P_2^I in the $a_{11} - b_3$ phase, respectively. For the energy expression in Eq. (3.5.18), when $P_1^M \neq 0$ and $|P_1^I|$ is small enough, the first term on the right hand side dominates and it is always possible to find a P_1^I that makes

$[\Delta\psi]_{a_{11}-b_3}^{r-b_1}$ negative. Following exactly the same rule, the replacement of the b_1 phase by the $a_1 - b_1$ phase and the $a_1 - b_2$ phase by the $a_{12} - b_1$ phase can be explained. Another significant difference between Fig. 3.1a and Fig. 3.2a is the configuration of Line 4. To explain this difference, we start with analyzing the phase transition from the $a_{11} - b_3$ phase to $a_{11} - \tilde{r}$ phase in Fig. 3.1a. The energy change between the $a_{11} - b_3$ phase and $a_{11} - \tilde{r}$ phase gives

$$\begin{aligned} [\Delta\psi]_{a_{11}-\tilde{r}}^{a_{11}-b_3} = & [2b_1^{M*} + 2b_{12}^{M*}(P_3^M)^2 + (1-c_f)2a_{112}^M(P_3^M)^4 \\ & + (2\gamma_{11} + 2\gamma_{12} + \gamma_{44})(P_1^I)^2](P_1^M)^2 \\ & + [2b_{11}^{M*} + b_{12}^{M*} + (1-c_f)(2a_{112}^M + a_{123}^M)(P_3^M)^2](P_1^M)^4 \\ & + (1-c_f)2(a_{111}^M + a_{112}^M)(P_1^M)^6. \end{aligned} \quad (3.5.19)$$

According to Eq. (3.5.19), the transition from the $a_{11} - b_3$ phase to $a_{11} - \tilde{r}$ phase occurs when

$$2b_1^{M*} + 2b_{12}^{M*}(P_3^M)^2 + (1-c_f)2a_{112}^M(P_3^M)^4 + (2\gamma_{11} + 2\gamma_{12} + \gamma_{44})(P_1^I)^2 = 0. \quad (3.5.20)$$

In Eq. (3.5.19), it is assumed that in the $a_{11} - \tilde{r}$ phase, P_3^M and $P_1^I = P_2^I$ come from the $a_{11} - b_3$ phase. Yet, strictly speaking, in the $a_{11} - \tilde{r}$ phase the solution of P_3^M and $P_1^I = P_2^I$ is coupled with $P_1^M = P_2^M$. Thus, satisfaction of Eq. (3.5.20) is a sufficient but not necessary condition for the occurrence of the phase transition.

Now replacing the $a_{11} - b_3$ phase with the $r - b_1$ phase, since P_1^I is small in the $r - b_1$ phase, Eq. (3.5.19) is recast as

$$\begin{aligned} [\Delta\psi]_{a_{11}-\tilde{r}}^{r-b_1} = & [2b_1^{M*} + 2b_{12}^{M*}(P_1^M)^2 + (1-c_f)2a_{112}^M(P_1^M)^4 \\ & + (2\gamma_{11} + 2\gamma_{12} + \gamma_{44})(P_2^I)^2](P_2^M)^2 \\ & + [2b_{11}^{M*} + b_{12}^{M*} + (1-c_f)(2a_{112}^M + a_{123}^M)(P_1^M)^2](P_2^M)^4 \\ & + (1-c_f)2(a_{111}^M + a_{112}^M)(P_2^M)^6 + 2r_{01}P_2^M P_2^I - r_{01}P_1^M P_1^I. \end{aligned} \quad (3.5.21)$$

In Eq. (3.5.21), P_2^M is a small quantity compatible with P_1^I ; therefore $[\Delta\psi]_{a_{11}-\tilde{r}}^{r-b_1}$ is transformed as

$$[\Delta\psi]_{a_{11}-\tilde{r}}^{r-b_1} = 2r_{01}P_2^M P_2^I - r_{01}P_1^M P_1^I + O[(P_1^M)^2]. \quad (3.5.22)$$

Thus

$$2r_{01}P_2^M P_2^I - r_{01}P_1^M P_1^I = 0. \quad (3.5.23)$$

For $P_2^M \ll 1$, it provides an alternative mechanism for the transition from the $r - b_1$ phase to $a_{11} - \tilde{r}$ phase, and this explains the shape change in the configuration of Line 4 in Fig. 3.2a in the low temperature zone compared with Fig. 3.1a. By comparing Eq. (3.5.20) with Eq. (3.5.23), the competition of the two phase transition mechanism in one phase diagram is generally determined by the two coefficients r_{01} and b_1^{M*} .

Since the electric polarizations in each phase which minimize the total energy can be solved from Eq. (3.5.2), it is possible to discuss the variation in such physical quantities as average electric polarization, piezoelectric coefficients or electrostriction coefficients of the entire composite versus eigenstrains. The average electric polarizations are defined as

$$\bar{P}_i = c_f P_i^I + (1 - c_f) P_i^M. \quad (3.5.24)$$

The variation in the average electric polarization at room temperature (20 °C) for the PTO-BTO ferroelectric composite with $c_f = 0.5$ is plotted in Fig. 3.4 where the effect of depolarization fields is neglected and in Fig. 3.5 where the effect of depolarization fields is considered.

3.6 Remarks

Micromechanics models use the mean fields instead of the actual fields to study the overall physical properties, which avoids the solution of the actual fields within the complex microstructures of the composites. The most significant advantage of these sorts of models is their conciseness. The methods can be used when the volume fraction of each phase is known without knowledge of the distribution of the microstructures. By incorporating the basic solutions introduced in Chapter 2, the results obtained by using these methods are basically reasonable. The most important factor that influences the validity of these models is the interaction between different inclusions, whose effect is not included in the basic solutions. Thus, the various assumptions of different micromechanics models aim at incorporating this interaction effect.

For ferroelectric composites, the combined method of the micromechanics models and the Landau theory for phase transition is constructed here. The method is able to incorporate many effects such as the external mechanical loads, the shape, orientation and volume fraction of the inclusions, etc. in the phase diagram analysis of ferroelectric composites. The energy formulation is constructed for general external mechanical loads and for inclusions with a general ellipsoidal shape. For a specified load condition and inclusion shape, the result should be

simplified before actual application. The phase diagrams and corresponding analysis presented in Section 3.5 are constructed for spherical BaTiO_3 inclusions embedded in the PbTiO_3 matrix with equiaxed strain applied in all three directions. Concerning the many possible choices of load conditions and inclusion shapes, there is a lot of potential work to do. Besides, the formulation presented here did not incorporate the external electric fields as well as the effect of depolarization fields, which is another important yet complex task.

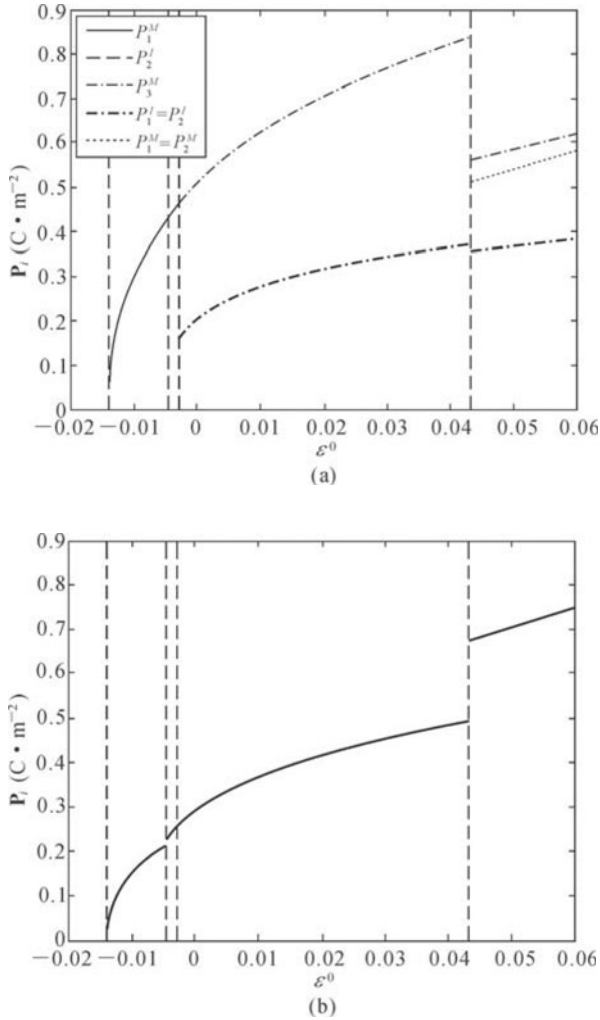


Fig. 3.4 ε_0 versus components of the electric polarization vector $\mathbf{P}_i$ (a) and average polarization of the ferroelectric composite $|\mathbf{P}|$ (b). Effect of depolarization fields is neglected

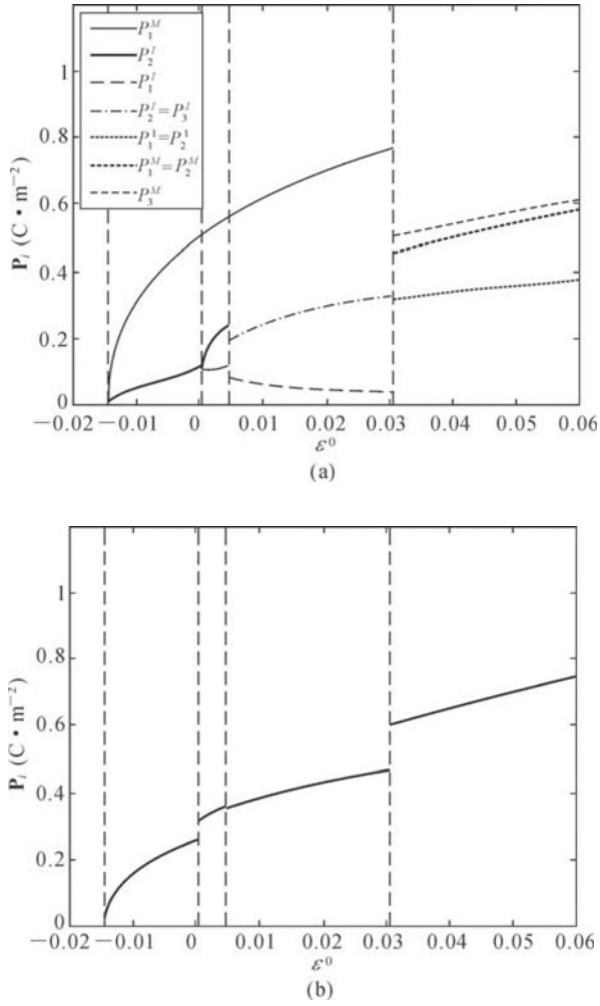


Fig. 3.5 ε_0 versus components of the electric polarization vector $\mathbf{P}_i$ (a) and average polarization of the ferroelectric composite $|\mathbf{P}|$ (b). Effect of depolarization fields is considered

Appendix A: Radon Transform

Considering function $f(\mathbf{x})$ defined in $\mathbb{R}^3$, the Radon transform of $f(\mathbf{x})$ is defined as

$$\hat{f}(s, \mathbf{n}) = R[f(\mathbf{x})] = \int f(\mathbf{x}) \delta(s - \mathbf{n} \cdot \mathbf{x}) d\mathbf{x}, \quad (\text{A1})$$

where $\mathbf{n}$ is a unit vector and δ is the three-dimensional Dirac delta. The Radon transform integrates $f(\mathbf{x})$ over all curves defined by $\mathbf{n} \cdot \mathbf{x} = s$.

From Eq. (A1), the following properties of the Radon transform hold.
Homogeneity:

$$\hat{f}(\alpha s, \alpha \mathbf{n}) = \frac{1}{|\alpha|} \hat{f}(s, \mathbf{n}). \quad (\text{A2})$$

Linearity:

$$R[c_1 f_1(\mathbf{x}) + c_2 f_2(\mathbf{x})] = c_1 \hat{f}_1(\alpha s, \alpha \mathbf{n}) + c_2 \hat{f}_2(\alpha s, \alpha \mathbf{n}). \quad (\text{A3})$$

Transformation of non-singular matrix:

$$R[f(\mathbf{A}^{-1} \mathbf{x})] = |\det \mathbf{A}| \hat{f}(s, \mathbf{A}^T \mathbf{n}), \quad (\text{A4})$$

where $\det \mathbf{A} \neq 0$.

Transformation of derivatives:

$$R\left[\frac{\partial f}{\partial x_i}\right] = n_i \frac{\partial \hat{f}(s, \mathbf{n})}{\partial s}; \quad R\left[\frac{\partial^2 f}{\partial x_i \partial x_j}\right] = n_i n_j \frac{\partial^2 \hat{f}(s, \mathbf{n})}{\partial s^2}. \quad (\text{A5})$$

The inverse Radon transform is defined as

$$f(\mathbf{x}) = -\frac{1}{8\pi^2} \int_S \hat{f}''(\mathbf{n} \cdot \mathbf{x}, \mathbf{n}) dS(\mathbf{n}), \quad (\text{A6})$$

where

$$\hat{f}''(\mathbf{n} \cdot \mathbf{x}, \mathbf{n}) = \left. \frac{\partial^2 \hat{f}(s, \mathbf{n})}{\partial s^2} \right|_{s=\mathbf{n} \cdot \mathbf{x}}. \quad (\text{A7})$$

$S(\mathbf{n})$ in Eq. (A.6) can be an arbitrary close surface containing the origin.

References

- [1] Avellaneda, M., Swart, P.J.: Calculating the performance of 1-3 piezoelectric composites for hydrophone applications: An effective medium approach. *J. Acoust. Soc. Am.* **103**, 1449-1467 (1998)
- [2] Bakhvalov, N., Panasenko, G.: Homogenization: Averaging Processes in Periodic Media: Mathematical Problems in the Mechanics of Composite Materials. Kluwer Academic Publishers, Dordrecht, Boston (1989)
- [3] Banno, H.: Recent developments of piezoelectric ceramic products and composites of synthetic rubber and piezoelectric ceramic particles. *Ferroelectrics* **50**, 3-12 (1983)
- [4] Benveniste, Y.: A new approach to the application of Mori-Tanaka's theory in composite materials. *Mech. Mater.* **6**, 147-157 (1987)
- [5] Benveniste, Y., Dvorak, G.J.: Uniform fields and universal relations in piezoelectric composites. *J. Mech. Phys. Solids* **40**, 1295-1312 (1992)
- [6] Benveniste, Y., Dvorak, G.J., Chen, T.: On diagonal and elastic symmetry of the approximate effective stiffness tensor of heterogeneous media. *J. Mech. Phys. Solids* **39**, 927-946 (1991)
- [7] Berger, H., Kari, S., Gabbert, U., Rodriguez-Ramos, R., Guinovart, R., Otero, J.A., Bravo-Castillero, J.: An analytical and numerical approach for calculating effective material coefficients of piezoelectric fiber composites. *Int. J. Solids Struct.* **42**, 5692-5714 (2005)
- [8] Bisegna, P., Luciano, R.: Variational bounds for the overall properties of piezoelectric composites. *J. Mech. Phys. Solids* **44**, 583-602 (1996)
- [9] Bisegna, P., Luciano, R.: On methods for bounding the overall properties of periodic piezoelectric fibrous composites. *J. Mech. Phys. Solids* **45**, 1329-1356 (1997)
- [10] Bohm, H.J.: Numerical investigation of microplasticity effects in unidirectional long fiber reinforced metal matrix composites. *Model. Simul. Mater. Sci. Eng.* **1**, 649-671 (1993)
- [11] Brockenbrough, J.R., Suresh, S.: Plastic deformation of continuous fiber-reinforced metal-matrix composites: Effects of fiber shape and distribution. *Scr. Metall. Mater.* **24**, 325-330 (1990)
- [12] Bruggeman, D.A.G.: Berechnung verschiedener physikalischer Konstanten von heterogen Substanzen. I. Dielektrizitätskonstanten und Leitfähigkeiten der Mischkörper aus isotropen Substanzen. *Ann. Phys.* **416**, 665-679 (1935)
- [13] Budiansky, B.: On the elastic moduli of some heterogeneous materials. *J. Mech. Phys. Solids* **13**, 223-227 (1965)
- [14] Chan, H.L.W., Unsworth, J.: Simple model for piezoelectric ceramic/polymer 1-3 composites used in ultrasonic transducer applications. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **36**, 434-441 (1989)
- [15] Chen, J.Q., Wang, B., Du, S.Y.: A micromechanics method to study the effect of domain switching on fracture behavior of polycrystalline ferroelectric ceramics. *Appl. Math. Mech.* **23**, 1250-1262 (2002)

- [16] Chen, T.Y.: Piezoelectric properties of multiphase fibrous composites: Some theoretical results. *J. Mech. Phys. Solids* **41**, 1781-1794 (1993)
- [17] Chen, T.Y.: Micromechanical estimates of the overall thermoelectroelastic moduli of multiphase fibrous composites. *Int. J. Solids Struct.* **31**, 3099-3111 (1994)
- [18] Chen, X., Fang, D.N., Hwang, K.C.: Micromechanics simulation of ferroelectric polarization switching. *Acta Mater.* **45**, 3181-3189 (1997)
- [19] Chiang, C.R.: Thermal mismatch stress of a spherical inclusion in a cubic crystal. *Int. J. Fract.* **139**, 313-317 (2006)
- [20] Cleveringa, H.H.M., van der giessen, E., Needleman, A.: Comparison of discrete dislocations and continuum plasticity predictions for a composite material. *Acta Mater.* **45**, 3163-3179 (1997)
- [21] Devonshire, A.F.: Theory of barium titanate-part I. *Phil. Mag.* **40**, 1040-1063 (1949)
- [22] Du, S.Y., Wang, B.: *Micromechanics of Composite Materials*. Science Press, Beijing, China (1998) (In Chinese)
- [23] Dunn, M.L.: Effects of grain shape anisotropy, porosity, and microcracks on the elastic and dielectric constants of polycrystalline piezoelectric ceramics. *J. Appl. Phys.* **78**, 1533-1541 (1995)
- [24] Dunn, M.L., Taya, M.: An analysis of piezoelectric composite materials containing ellipsoidal inhomogeneities. *Proc. R. Soc. Lond. A* **443**, 265-287 (1993a)
- [25] Dunn, M.L., Taya, M.: Electromechanical properties of porous piezoelectric ceramics. *J. Am. Ceram. Soc.* **76**, 1697-1706 (1993b)
- [26] Dunn, M.L., Taya, M.: Micromechanics predictions of the effective electroelastic moduli of piezoelectric composites. *Int. J. Solids Struct.* **30**, 161-175 (1993c)
- [27] Erhart, J., Cao, W.W.: Effective material properties in twinned ferroelectric crystals. *J. Appl. Phys.* **86**, 1073-1081 (1999)
- [28] Eshelby, J.D.: The determination of the elastic field of an ellipsoidal inclusion, and related problems. *Proc. R. Soc. Lond. A* **241**, 376-396 (1957)
- [29] Ferrari, M.: Asymmetry and the high concentration limit of the Mori-Tanaka effective medium theory. *Mech. Mater.* **11**, 251-256 (1991)
- [30] Fulton, C.C., Gao, H.: Microstructural modeling of ferroelectric fracture. *Acta Mater.* **49**, 2039-2054 (2001)
- [31] Grekov, A.A., Kramarov, S.O., Kuprienko, A.A.: Anomalous behavior of the two-phase lamellar piezoelectric texture. *Ferroelectrics* **76**, 43-48 (1987)
- [32] Grekov, A.A., Kramarov, S.O., Kuprienko, A.A.: Effective properties of a transversely isotropic piezocomposite with cylindrical inclusions. *Ferroelectrics* **99**, 115-126 (1989)
- [33] Gunawardena, S.R., Jansson, S., Leckie, F.A.: Modeling of anisotropic behavior of weakly bonded fiber reinforced MMCs. *Acta Metall. Mater.* **41**, 3147-3156 (1993)
- [34] Hershey, A.V.: The elasticity of an isotropic aggregate of anisotropic cubic crystals. *J. Appl. Mech.* **21**, 236-241 (1954)

- [35] Hill, R.: A self-consistent mechanics of composite materials. *J. Mech. Phys. Solids* **13**, 213-222 (1965)
- [36] Hu, Y.F., Wang, B.: Magnetolectric phase diagrams for ferromagnetic/ferroelectric composites. (Finished)
- [37] Huang, J.H., Kuo, W.S.: Micromechanics determination of the effective properties of piezoelectric composites containing spatially oriented short fibers. *Acta Mater.* **44**, 4889-4898 (1996)
- [38] Huber, J.E.: Micromechanical modeling of ferroelectrics. *Curr. Opin. Solid State Mater. Sci.* **9**, 100-106 (2005)
- [39] Huber, J.E., Fleck, N.A.: Ferroelectric switching: A micromechanics model versus measured behavior. *Eur. J. Mech. A/Solids* **23**, 203-217 (2004)
- [40] Jiang, C.P., Cheung, Y.K.: An exact solution for the three-phase piezoelectric cylinder model under antiplane shear and its applications to piezoelectric composites. *Int. J. Solids Struct.* **38**, 4777-4796 (2001)
- [41] Jiang, C.P., Tong, Z.H., Cheung, Y.K.: A generalized self-consistent method for piezoelectric fiber reinforced composites under antiplane shear. *Mech. Mater.* **33**, 295-308 (2001)
- [42] Kroner, E.: Berechnung der elastischen konstanten des vielkristalls aus den konstanten des einkristalls. *Z. Phys. A* **151**, 504-518 (1958) (In German)
- [43] Landau, L.D., Lifshitz, E.M.: *Electrodynamics of Continuous Media*. Pergamon Press, Oxford (1960)
- [44] Li, J., Weng, G.J.: A theory of domain switch for the nonlinear behavior of ferroelectrics. *Proc. R. Soc. Lond. A* **455**, 3493-3511 (1999)
- [45] Li, J., Weng, G.J.: A micromechanics-based hysteresis model for ferroelectric ceramics. *J. Intel. Mater. Syst. Struct.* **12**, 79-91 (2001)
- [46] Li, W.F., Weng, G.J.: A micromechanics-based thermodynamic model for the domain switch in ferroelectric crystals. *Acta Mater.* **52**, 2489-2496 (2004)
- [47] Li, Y.L., Cross, L.E., Chen, L.Q.: A phenomenological thermodynamic potential for BaTiO₃ single crystals. *J. Appl. Phys.* **98**, 064101 (2005)
- [48] Lu, W., Fang, D.N., Li, C.Q., Hwang, K.C.: Nonlinear electric-mechanical behavior and micromechanics modeling of ferroelectric domain evolution. *Acta Mater.* **47**, 2913-2926 (1999)
- [49] McLaughlin, R.: A study of the differential scheme for composite materials. *Int. J. Eng. Sci.* **15**, 237-244 (1977)
- [50] Mori, T., Tanaka, K.: Average stress in matrix and average elastic energy of materials with misfitting inclusions. *Acta Metall.* **21**, 571-574 (1973)
- [51] Muller, H.: Properties of Rochelle salt. *Phys. Rev.* **47**, 175-191 (1935)
- [52] Muller, H.: Properties of Rochelle salt. II. *Phys. Rev.* **57**, 829-839 (1940a)
- [53] Muller, H.: Properties of Rochelle salt. III. *Phys. Rev.* **58**, 565-573 (1940b)
- [54] Muller, H.: Properties of Rochelle salt. IV. *Phys. Rev.* **58**, 805-811 (1940c)
- [55] Mura, T.: *Micromechanics of Defects in Solids*, 2nd ed. Martinus Nijhoff Publishers, Hague, Netherlands (1987)
- [56] Nan, C.W., Clarke, D.R.: Effective properties of ferroelectric and/or ferromagnetic composites: A unified approach and its application. *J. Am. Ceram. Soc.* **80**, 1333-1340 (1997)

- [57] Newnham, R.E., Skinner, D.P., Cross, L.E.: Connectivity and piezoelectric-pyroelectric composites. *Mater. Res. Bull.* **13**, 525-536 (1978)
- [58] Pertsev, N.A., Zembilgotov, A.G., Tagantsev, A.K.: Effect of mechanical boundary conditions on phase diagrams of epitaxial ferroelectric thin films. *Phys. Rev. Lett.* **80**, 1988-1991 (1998)
- [59] Poizat, C., Sester, M.: Effective properties of composites with embedded piezoelectric fibers. *Comput. Mater. Sci.* **16**, 89-97 (1999)
- [60] Qiu, Y.P., Weng, G.J.: On the application of Mori-Tanaka's theory involving transversely isotropic spheroidal inclusions. *Int. J. Eng. Sci.* **28**, 1121-1137 (1990)
- [61] Reuss, A.: Berechnung der Fließgrenze von Mischkristallen auf Grund der Plastizitätsbedingung fuer Einkristalle. *Z. Angew. Math. Mech.* **9**, 49-58 (1929) (In German)
- [62] Sabina, F.J., Rodriguez-Ramos, R., Bravo-Castillero, J., Guinovart-Diaz, R.: Closed-form expressions for the effective coefficients of a fibre-reinforced composite with transversely isotropic constituents. II: Piezoelectric and hexagonal symmetry. *J. Mech. Phys. Solids* **49**, 1463-1479 (2001)
- [63] Smith, W.A., Auld, B.A.: Modeling 1-3 composite piezoelectrics: Thickness-mode oscillations. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **38**, 40-47 (1991)
- [64] Smit, R.J.M., Brekelmans, W.A.M., Meijer, H.E.H.: Prediction of the mechanical behavior of nonlinear heterogeneous systems by multi-level finite element modeling. *Comput. Methods Appl. Mech. Eng.* **155**, 181-192 (1998)
- [65] Su, Y., Weng, G.J.: A self-consistent polycrystal model for the spontaneous polarization of ferroelectric ceramics. *Proc. R. Soc. A* **462**, 1763-1789 (2006)
- [66] Suquet, P.: Elements of homogenization theory for inelastic solid mechanics. In: Sanchez-Palencia, E., Zaoui, A. (Eds.) *Homogenization Techniques for Composite Media*, pp. 194-275. Springer-Verlag, Berlin (1987)
- [67] Taya, M., Arsenault, R.J.: *Metal Matrix Composites: Thermomechanical Behavior*. Pergamon Press, Oxford (1989)
- [68] Teply, J.L., Dvorak, G.J.: Bounds on overall instantaneous properties of elastic-plastic composites. *J. Mech. Phys. Solids* **36**, 29-58 (1988)
- [69] Van Dommelen, J.A.W., Schrauwen, B.A.G., Van Breemen, L.C.A., Govaert, L.E.: Micromechanical modeling of the tensile behavior of oriented polyethylene. *J. Poly. Sci. B* **42**, 2983-2994 (2004)
- [70] Voigt, W.: Uber die Beziehung zwischen den beiden Elastizitätskonstanten isotroper Korper. *Weid. Ann.* **38**, 573-587 (1889) (In German)
- [71] Wang, B.: Three-dimensional analysis of an ellipsoidal inclusion in a piezoelectric material. *Int. J. Solids Struct.* **29**, 293-308 (1992)
- [72] Wang, B.: Effective behaviors of piezoelectric composites. *Appl. Mech. Rev.* **47**, S112-S121 (1994)
- [73] Wang, B., Woo, C.H.: Curie temperature and critical thickness of ferroelectric thin films. *J. Appl. Phys.* **97**, 084109 (2005)
- [74] Weng, G.J.: Some elastic properties of reinforced solids, with special

- reference to isotropic ones containing spherical inclusions. *Int. J. Eng. Sci.* **22**, 845-856 (1984)
- [75] Weng, G.J.: The theoretical connection between Mori-Tanaka's theory and the Hashin-Shtrikman-Walpole bounds. *Int. J. Eng. Sci.* **28**, 1111-1120 (1990)
- [76] Weng, G.J.: Explicit evaluation of Willis' bounds with ellipsoidal inclusions. *Int. J. Eng. Sci.* **30**, 83-92 (1992)
- [77] Wong, C.K., Poon, Y.M., Shin, F.G.: Explicit formulas for effective piezoelectric coefficients of ferroelectric 0-3 composites. *J. Appl. Phys.* **90**, 4690-4700 (2001)
- [78] Woo, C.H., Zheng, Y.: Depolarization in modeling nano-scale ferroelectrics using the Landau free energy functional. *Appl. Phys. A* **91**, 59-63 (2008)
- [79] Wu, T.L.: Micromechanics determination of electroelastic properties of piezoelectric materials containing voids. *Mater. Sci. Eng. A* **280**, 320-327 (2000)
- [80] Zhou, S.A.: A material multipole theory of elastic dielectric composites. *Int. J. Solids Struct.* **28**, 423-447 (1991)

Determination of the Smallest Sizes of Ferroelectric Nanodomains

4.1 Introduction

Different ferroelectric domains have different piezoelectric, pyroelectric, electro-optic and nonlinear optic properties, and many applications of ferroelectric materials originated from this characteristic. The unique ability of a ferroelectric film to maintain its polarization, even in the absence of an external voltage, makes it very useful as a nonvolatile memory device, etc. Ferroelectric random access memory (FRAM) devices made from ferroelectrics have attracted much attention recently, due to their ideal properties such as nonvolatility, high speed and low power consumption, etc. (Scott and Paz de Araujo, 1989; Song et al., 2002) (Fig. 4.1).

The development of high-density FRAMs requires the fabrication of switched domains of very small size under low voltage or short pulses. Recently, atomic-force microscopes (AFM) have been successfully used to imprint small ferroelectric domains onto the surface of ferroelectric materials (Luthi et al., 1993). Devices prepared in this way have been proposed for possible use as high-density nonvolatile memory because of the very small feature size achievable using this technique. Polarization in AFM writing is accomplished with the high electric field of the AFM tip (Fig. 4.2). There is no need to use a top electrode, like in the usual poling process. However, without electrodes, there is no bound charge to tie the polarization. Therefore, the depolarization energy may cause serious depoling when the AFM tip is removed (Fig. 4.3). Thus, it is very important for us to know the stability conditions for a domain in the absence of an external electric field.



Fig. 4.1 FRAM produced by Ramtron (picture from Wikipedia)

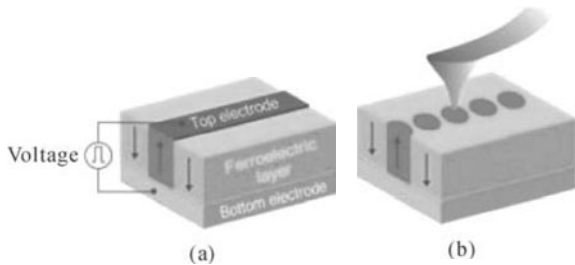


Fig. 4.2 Methods to switch the ferroelectric polarization. (a) In a traditional ferroelectric capacitor, a voltage is applied across the ferroelectric layer with metallic gate electrodes; (b) In the scanning-probe approach, the top electrode is replaced by a conducting, nanoscale scanning probe, to which a voltage is applied (Ahn et al., 1997). Copyright(1998) by Institute of Physics

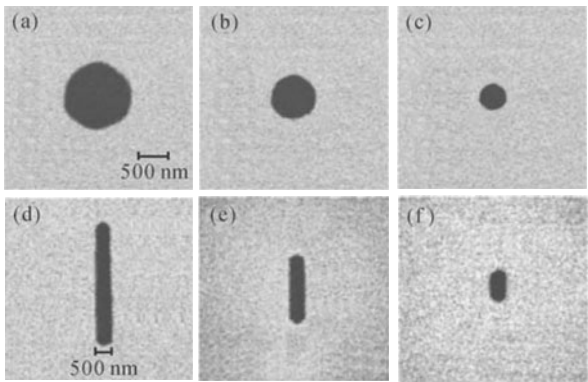


Fig. 4.3 Shrinkage of ferroelectric domain (Kalinin et al., 2010). Copyright(2010) by American Institute of Physics

Consider a ferroelectric material which is fully polarized. All of the dipole moments are aligned in the same direction. Then, by applying a reversed electric field over it, new domains with a reversed polarization will nucleate. This process

requires a certain amount of time, in the same manner as any nucleation process. When the nucleated nuclei reach a certain critical radius, they grow outwards, forming needle-like structures (Fig. 4.4). The growth of the nucleated ferroelectric domain is a non-equilibrium process in the consideration of thermodynamics. And generally speaking, a 180° domain evolves under the combined action of the externally applied electric field, the depolarization field, the surface tension of the domain wall, and an internal electric field that may originate from defects in the material. The basic discussion was made by Landauer (1957) with an emphasis on the electrostatic consideration.

In this chapter we focus on the problem of the stability of such a nucleated ferroelectric domain and related problems. Instead of using numerical simulation of the evolution process, we use the Liapunov criterion to determine the minimum stable ferroelectric domains.

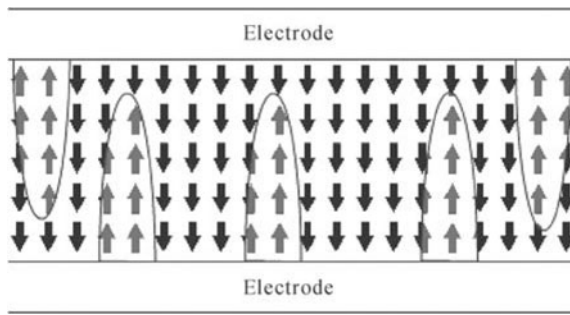


Fig. 4.4 Ferroelectric domains

4.2 Electric Fields in Ferroelectric Thin Film

4.2.1 General Expression of Electric Field of Ferroelectric Domain

Consider a ferroelectric crystal below its Curie temperature. The dependence of the polarization of the crystal on the applied electric field is sketched in Fig. 4.5. The crystal exhibits a permanent polarization $\mathbf{P}_s$ even in the absence of external electric fields. The parameter $\mathbf{P}_s$ is also called the spontaneous polarization. The constitutive equation for the ferroelectric crystal can be expressed in the following form:

$$\mathbf{D} = \boldsymbol{\varepsilon} \cdot \mathbf{E} + \mathbf{P}_s, \quad (4.2.1)$$

where $\boldsymbol{\varepsilon}$ is the dielectric permittivity tensor, $\mathbf{D}$ is the electric induction and $\mathbf{E}$ is the total electric field. The ferroelectric crystal can be made to be equivalent to an

ordinary dielectric material with a uniform distribution of spontaneous polarization. If we apply oppositely a high enough electric field, the spontaneous polarization in some regions will begin to reverse its direction. Such regions are called the 180° domains. Those 180° domains will grow until all the spontaneous dipole moments in the crystal switch off their direction. In the current investigation, the 180° growing domain is modeled as an ellipsoid inclusion. Then we are going to derive the electric field around it.

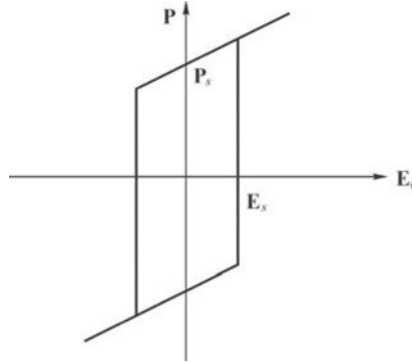


Fig. 4.5 Schematic of dependence $P(E)$ on the applied electric field E_0 for a ferroelectric crystal, where P_s is the spontaneous polarization and E_s is the coercive field

Since spontaneous polarization plays a role as the external field source, to obtain the induced electric fields by an ellipsoidal growing 180° domain, we can decompose the original problem shown in Fig. 4.6a into the two cases as shown in Fig. 4.6b and Fig. 4.6c. The superposition of the two solutions gives the solution of the original problem for a linear medium.

First, we consider the following case: Under the action of the electric field, a ferroelectric crystal with uniform spontaneous polarization P_s along the opposite direction of the electric field (Fig. 4.6b).

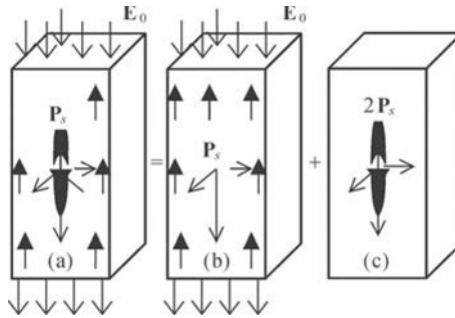


Fig. 4.6 A 180° domain growth in a parent phase driven by an electric field

Since the ferroelectric crystal can be treated as an ordinary dielectric material with spontaneous polarization P_s , the uniform polarization will generate a charge

distribution on the surface of the crystal, which creates a macroscopic electric field $-\mathbf{P}_s/\varepsilon$. ε is the dielectric constant. If the applied electric field in the dielectric material without the spontaneous polarization is $\mathbf{e}_0$, the total electric field in the crystal is given by

$$\mathbf{E}_0 = \mathbf{e}_0 - \frac{\mathbf{P}_s}{\varepsilon}, \quad (4.2.2)$$

where $\mathbf{E}_0$ is usually treated as the applied electric field in the ferroelectric crystal when we study the domain reverse problem.

Secondly we consider the following: In the dielectric medium there is a growing 180° domain with polarization $2\mathbf{P}_s$ along the direction of the applied electric field (Fig. 4.6c).

This is an electrodynamics problem. If the 180° domain is assumed to take an ellipsoidal shape with principle half axes a_1, a_2, a_3 , using the full Maxwell's equation, one can obtain the Hertz potential Π as follows:

$$\Pi = \frac{p_s \mathbf{k}}{2\pi\varepsilon} \iiint_{\Omega(t)} dv' \frac{1}{|\mathbf{r} - \mathbf{r}'|} = \frac{p_s \mathbf{k}}{4\pi\varepsilon} [I(\lambda) - x_n x_n I_N(\lambda)], \quad (4.2.3)$$

where $x_n x_n I_N(\lambda) = x_1^2 I_1(\lambda) + x_2^2 I_2(\lambda) + x_3^2 I_3(\lambda)$, and $\mathbf{k}$ is the unit vector along the polarization direction. The expressions of $\lambda, I, I_1, I_2, I_3$, together with the detailed derivation of Eq. (4.2.3), are given in the appendix.

By using the Hertz potential given in Eq. (4.2.3), the electric field can be derived as follows:

$$E_1 = \frac{\partial^2 \Pi_3}{\partial x_1 \partial x_3}, E_2 = \frac{\partial^2 \Pi_3}{\partial x_2 \partial x_3}, E_3 = \frac{\partial^2 \Pi_3}{\partial x_3^2} - \frac{1}{c^2} \frac{\partial^2 \Pi_3}{\partial t^2} \approx \frac{\partial^2 \Pi_3}{\partial x_3^2}. \quad (4.2.4)$$

• The electric field inside the domain

Consider the electric field inside the 180° domain (Fig. 4.6a). For interior points of an ellipsoid $\Omega(t)$, $\lambda = 0$, therefore, I, I_i become constants. Substituting Eq. (4.2.3) into Eq. (4.2.4), then taking the sum of Eq. (4.2.2) and Eq. (4.2.4), one obtains the electric field inside the domain as follows:

$$\begin{aligned} E_1^{\text{in}} &= 0, \\ E_2^{\text{in}} &= 0, \\ E_3^{\text{in}} &= E_0 - \frac{p_s}{2\pi\varepsilon} I_3. \end{aligned} \quad (4.2.5)$$

It is worth mentioning that for an ellipsoidal domain, the electric field remains uniform within the domain for uniform polarization. For a spherical domain with radius a ,

$$I = 4\pi a^2, \quad I_1 = I_2 = I_3 = \frac{4}{3}\pi. \quad (4.2.6)$$

Substituting Eq. (4.2.6) into Eq. (4.2.5) yields

$$E_3^{\text{in}} = E_0 - \frac{2p_s}{3\epsilon}. \quad (4.2.7)$$

The other components of the electric field within the domain are zero.

• *The electric field outside the domain*

Now we consider the electric field outside the 180° domain (Fig. 4.6a). By substituting Eq. (4.2.3) into Eq. (4.2.4) and adding Eq. (4.2.2), we obtain the electric field outside the domain as

$$\begin{aligned} E_1^{\text{out}} &= \frac{\partial^2 I_3}{\partial x_1 \partial x_3} = \frac{p_s}{4\pi\epsilon} \frac{\partial^2}{\partial x_1 \partial x_3} [I(\lambda) - x_n x_n I_N(\lambda)], \\ E_2^{\text{out}} &= \frac{\partial^2 I_3}{\partial x_2 \partial x_3} = \frac{p_s}{4\pi\epsilon} \frac{\partial^2}{\partial x_2 \partial x_3} [I(\lambda) - x_n x_n I_N(\lambda)], \\ E_3^{\text{out}} &= E_0 - \left(\frac{1}{c^2} \frac{\partial^2 I_3}{\partial t^2} - \frac{\partial^2 I_3}{\partial x_3^2} \right) \\ &= E_0 - \frac{p_s}{4\pi\epsilon} \left(\frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x_3^2} \right) [I(\lambda) - x_n x_n I_N(\lambda)]. \end{aligned} \quad (4.2.8)$$

Here, we have derived the general expression of the electric field inside and outside the ferroelectric domain (Wang and Xiao, 2000).

4.2.2 AFM-Induced Electric Field in Ferroelectric Thin Films

To obtain the explicit energy expression of the FE domain induced by the AFM tip in thin film, it is very important to know the electric field distribution induced by the conducting atomic force microscopy in the ferroelectric thin film. Here, we derive an explicit expression for the AFM-induced electric field in a ferroelectric thin film.

4.2.2.1 Electric Field Near the AFM Tip

In ferroelectric domain imaging and tailoring, polarization is produced using the super high electric field of the AFM tip, which can be modeled as a single charge Q trapped in a conducting sphere with radius a , corresponding to the tip radius. The electric potential created by the sphere in vacuum is

$$\Phi_0 = \begin{cases} -\frac{Q}{4\pi\epsilon_0 a}, & r \leq a \\ -\frac{Q}{4\pi\epsilon_0 r}, & r > a \end{cases} \quad (4.2.9)$$

where r is the distance from the charge at the origin to the field point, and ϵ_0 is the dielectric constant of the vacuum.

The AFM tip/ferroelectric thin film/dielectric substrate configuration is shown in Fig. 4.7. If the substrate serves as the bottom electrode, one can assume that it has an infinite dielectric constant $\epsilon_2 \rightarrow \infty$.

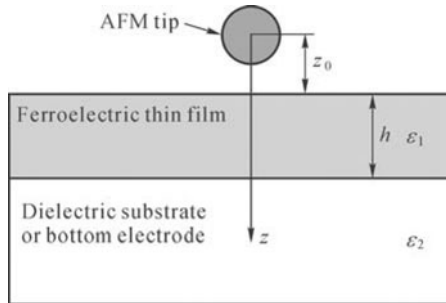


Fig. 4.7 AFM tip/ferroelectric thin film/dielectric substrate configuration

The electric field is symmetric with respect to the z axis for our problem, and it is convenient to use the cylindrical coordinate system (r, ϕ, z) , in which the Laplace equation for the electric potential is

$$\frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial \Phi}{\partial r} \right) + \frac{1}{r^2} \frac{\partial^2 \Phi}{\partial \phi^2} + \frac{\partial^2 \Phi}{\partial z^2} = 0. \quad (4.2.10)$$

By means of separating the variables

$$\Phi = \Theta(r)\Psi(\phi)\Omega(z), \quad (4.2.11)$$

one obtains

$$r \frac{d}{dr} \left(r \frac{d\Theta}{dr} \right) + (k^2 r^2 - n^2) \Theta = 0, \quad (4.2.12)$$

$$\frac{d^2 \Psi}{d\varphi^2} + n^2 \Psi = 0, \quad (4.2.13)$$

$$\frac{d^2 \Omega}{dz^2} - k^2 \Omega = 0, \quad (4.2.14)$$

where k and n are the separation parameters. Eq. (4.2.12) takes the form of the Bessel equation, with the Bessel functions being its solution. For azimuthal symmetry, $n = 0$. If the solution extends to infinity radially, so that there are no cylindrical boundaries that might impose restrictions on the radial function $\Theta(r)$, then there are correspondingly no restrictions on k , and the solution can be expressed as an integral over all k :

$$\Theta(r) = \int_0^\infty e^{\pm kz} f(k) J_0(kr) dk, \quad (4.2.15)$$

where $J_0(kr)$ is the Bessel function zeroth order of the first kind, and the function $f(k)$ is to be determined by the boundary conditions.

By using the properties of the Bessel functions and the electric potential outside the AFM tip, Eq. (4.2.9) can be written as

$$\Phi_0 = -\frac{Q}{4\pi\epsilon_0} \int_0^\infty e^{-k|z|} J_0(kr) dk. \quad (4.2.16)$$

The potential of Eq. (4.2.16) can be used with the induced potential in Eq. (4.2.15) to write the solution for the boundary-value problem. For plane boundaries normal to the z axis (Fig. 4.7),

$$\Phi_1 = -\frac{Q}{4\pi\epsilon_0} \left[\int_0^\infty e^{-k|z|} J_0(kr) dk + \int_0^\infty A(k) e^{kz} J_0(kr) dk \right], \quad z < z_0, \quad (4.2.17)$$

$$f_a > 0, \quad (4.2.18)$$

$$\Phi_3 = -\frac{Q}{4\pi\epsilon_0} \int_0^\infty D(k) e^{-k|z|} J_0(kr) dk, \quad z > z_0 + h. \quad (4.2.19)$$

The undetermined coefficients $A(k)$, $B(k)$, $C(k)$ and $D(k)$ have to be solved to

satisfy the boundary conditions for all values of r from 0 to ∞ . This requires that the integrands alone satisfy these conditions.

At $z = z_0$, we have $\Phi_1 = \Phi_2$, and $\varepsilon_0 \frac{d\Phi_1}{dz} = \varepsilon_1 \frac{d\Phi_2}{dz}$, giving, respectively,

$$e^{-kz_0} + A(k)e^{kz_0} = B(k)e^{-kz_0} + C(k)e^{kz_0}, \quad (4.2.20)$$

$$-\varepsilon_0 e^{-kz_0} + \varepsilon_0 A(k)e^{kz_0} = -\varepsilon_1 B(k)e^{-kz_0} + \varepsilon_1 C(k)e^{kz_0}. \quad (4.2.21)$$

At $z = z_0 + h$, $\Phi_2 = \Phi_3$ and $\varepsilon_1 \frac{d\Phi_2}{dz} = \varepsilon_2 \frac{d\Phi_3}{dz}$, giving, respectively,

$$D(k)e^{-k(z_0+h)} = B(k)e^{-k(z_0+h)} + C(k)e^{k(z_0+h)}, \quad (4.2.22)$$

$$-\varepsilon_2 D(k)e^{-k(z_0+h)} = -\varepsilon_1 B(k)e^{-k(z_0+h)} + \varepsilon_1 C(k)e^{k(z_0+h)}. \quad (4.2.23)$$

Solving Eqs. (4.2.20) – (4.2.23) of $A(k)$, $B(k)$, $C(k)$, $D(k)$, we have

$$\begin{aligned} B(k) &= \frac{2(1 + \varepsilon_1/\varepsilon_2)}{(1 + \varepsilon_1/\varepsilon_0)(1 + \varepsilon_1/\varepsilon_2) - (1 - \varepsilon_1/\varepsilon_0)(1 - \varepsilon_1/\varepsilon_2)e^{-2kh}}, \\ C(k) &= \frac{2(1 - \varepsilon_1/\varepsilon_2)e^{-2k(z_0+h)}}{(1 - \varepsilon_1/\varepsilon_0)(1 - \varepsilon_1/\varepsilon_2)e^{-2kh} - (1 + \varepsilon_1/\varepsilon_0)(1 + \varepsilon_1/\varepsilon_2)}, \\ D(k) &= \frac{4\varepsilon_1/\varepsilon_2}{(1 + \varepsilon_1/\varepsilon_0)(1 + \varepsilon_1/\varepsilon_2) - (1 - \varepsilon_1/\varepsilon_0)(1 - \varepsilon_1/\varepsilon_2)e^{-2kh}}, \\ A(k) &= [B(k) - 1]e^{-2kz_0}C(k). \end{aligned} \quad (4.2.24)$$

Substitution of Eq. (4.2.24) into Eqs. (4.2.7) and (4.2.8) gives the electric potential in the different regions. In the ferroelectric thin film, we have

$$\begin{aligned} \Phi_2 = -\frac{2Q}{4\pi\varepsilon_0} &\left(\frac{1}{1 + \varepsilon_1/\varepsilon_0} \int_0^\infty \frac{J_0(kr)e^{-kz}}{1 - \beta e^{-2kh}} dk \right. \\ &\left. - \frac{(1 - \varepsilon_1/\varepsilon_2)}{(1 + \varepsilon_1/\varepsilon_0)(1 + \varepsilon_1/\varepsilon_2)} \int_0^\infty \frac{J_0(kr)e^{k[z - 2(z_0+h)]}}{1 - \beta e^{-2kh}} dk \right), \end{aligned} \quad (4.2.25)$$

where $\beta = \frac{(1 - \varepsilon_1/\varepsilon_0)(1 - \varepsilon_1/\varepsilon_2)}{(1 + \varepsilon_1/\varepsilon_0)(1 + \varepsilon_1/\varepsilon_2)}$.

Despite the complicated appearance, the integrals in Eq. (4.2.25) can be readily evaluated numerically. Noting that $\beta < 1$, and expanding the denominator of the integrands with respect to β , we can use the following mathematical identity:

$$\frac{1}{R} = \frac{1}{\sqrt{z^2 + r^2}} = \int_0^\infty e^{-k|z|} J_0(kr) dk, \quad (4.2.26)$$

to obtain the solution in series form,

$$\begin{aligned} \Phi_2 = & -\frac{2Q}{4\pi\epsilon_0} \left[\frac{1}{1 + \epsilon_1/\epsilon_0} \sum_{n=0}^\infty \frac{\beta^n}{[(z + 2nh)^2 + r^2]^{1/2}} \right. \\ & \left. - \frac{(1 - \epsilon_1/\epsilon_2)}{(1 + \epsilon_1/\epsilon_0)(1 + \epsilon_1/\epsilon_2)} \sum_{n=0}^\infty \frac{\beta^n}{[\hat{z} + 2nh)^2 + r^2]^{1/2}} \right], \end{aligned} \quad (4.2.27)$$

where $\hat{z} = 2(z_0 + h) - z$.

If the substrate is covered by a bottom electrode and the film is absent, then $\epsilon_2 \rightarrow \infty$, $\epsilon_1 = \epsilon_0$ and Eq. (4.2.27) gives

$$\begin{aligned} \Phi_2 = & -\frac{Q}{4\pi\epsilon_0} \left(\int_0^\infty J_0(kr) e^{-kz} dk - \int_0^\infty J_0(kr) e^{k[z - 2(z_0 + h)]} dk \right) \\ = & -\frac{Q}{4\pi\epsilon_0} \left(\frac{1}{\sqrt{z^2 + r^2}} - \frac{1}{\sqrt{r^2 + [z - 2(z_0 + h)]^2}} \right). \end{aligned} \quad (4.2.28)$$

Eq. (4.2.28) can also be derived via the image method by placing the image charge in the substrate at a specific point $(z_0 + h)$ from the substrate surface.

In the following analysis, the ferroelectric material is treated as an ordinary dielectric with spontaneous polarization. The presence of spontaneous polarization charges in ferroelectric materials induces a depolarization field, which we shall now consider. The substrates of electronic devices may often contain a multilayer of different dielectric materials. We note that the present method can readily be extended to such cases. By taking the derivative of the electric potential, Eq. (4.2.25), with respect to z , the electric field component in the thickness direction is given by

$$E_z = -\frac{\partial \Phi_2}{\partial z} = -\frac{2Q}{4\pi\epsilon_0} \left(\frac{1}{1 + \epsilon_1/\epsilon_0} \int_0^\infty \frac{kJ_0(kr)e^{-kz}}{1 - \beta e^{-2kh}} dk + \frac{(1 - \epsilon_1/\epsilon_2)}{(1 + \epsilon_1/\epsilon_0)(1 + \epsilon_1/\epsilon_2)} \int_0^\infty \frac{kJ_0(kr)e^{k[z-2(z_0+h)]}}{1 - \beta e^{-2kh}} dk \right). \quad (4.2.29)$$

We consider the electric field in the film for two configurations: (1) the AFM-tip/PZT film/bottom electrode and (2) the AFM-tip/PZT film/MgO substrate. We suppose the dielectric constant of the electrode is infinite, whereas that of the MgO substrate is lower than that of the PZT thin film. Being near the ferroelectric film and the substrate, the electric charge inside the tip depends on the configuration, and can be determined using Eq. (4.2.17) by putting Φ_1 equal to the voltage applied at the AFM tip. The material constant is shown in Table 4.1, in which the AFM tip is on the thin film. In this case, the charge inside the tip is determined by putting $z = z_0$.

Table 4.1 Numerical parameters

Parameters	Values
Dielectric constant of PZT film, ϵ_1	$8.8419 \times 10^{-9} \text{ F} \cdot \text{m}^{-1}$
Dielectric constant of MgO substrate, ϵ_2	$7.1619 \times 10^{-11} \text{ F} \cdot \text{m}^{-1}$
Dielectric constant of vacuum, ϵ_0	$8.8419 \times 10^{-12} \text{ F} \cdot \text{m}^{-1}$
AFM-tip radius, a	20 nm
Thickness of thin film, h	270 nm
Distance from the center of tip to the surface of the film, z_0	20 nm
Voltage applied to the tip, Φ_1	20 V

For the two configurations, the electric potentials in the thin film are shown in Fig. 4.8 as a function of the coordinates z and r . The corresponding electric fields are shown in Fig. 4.9. It is clear that the substrate has a large effect on the electric field distribution in the thin film, which increases for thinner films as shown in Fig. 4.10.

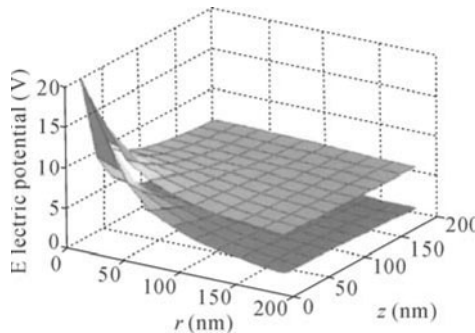


Fig. 4.8 Electric potential in the thin film as a function of coordinate systems (z, r)

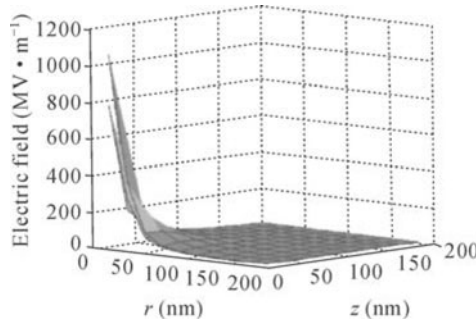


Fig. 4.9 Electric potential in the thin film as a function of coordinate systems (z, r)

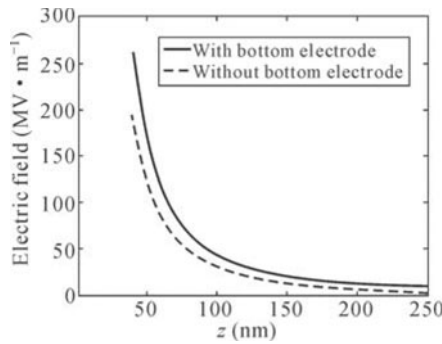


Fig. 4.10 Electric field in the thin film vs. z at $r=0$

4.2.2.2 Depolarization Field of a Ferroelectric Domain in Thin Film

The depolarization field in ferroelectric films is created by spontaneous polarization charges ρ_p ,

$$\rho_p = -\nabla \cdot \mathbf{P} = -P_{i,i}, \quad (4.2.30)$$

where P_i is the i -th component of spontaneous polarization. The depolarization field induced by the domain can be obtained by integrating, on the boundary of the domain, Green's function in the thin film for the vacuum/thin film/substrate configuration, which we will derive in the following.

- *Green's function of the vacuum/film/substrate configuration*

The solution follows a procedure similar to that adopted above, except that a unit positive point charge is in the ferroelectric thin film (Fig. 4.11) in this case. Using the cylindrical function to reflect the azimuthal symmetry, for a positive charge at

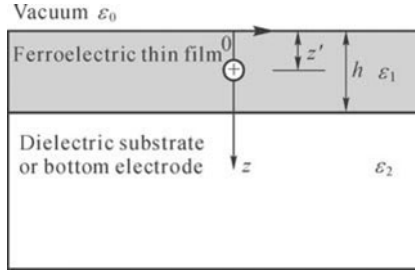


Fig. 4.11 Schematic of a unit positive charge in the ferroelectric thin film

$z = z', r = 0$, the electric potentials in the vacuum, thin film and dielectric substrate can be expressed, respectively, as

$$G_1 = \frac{1}{4\pi\epsilon_0} \int_0^\infty M(k) e^{k(Z-z')} J_0(kr) dk, \quad Z < 0; \quad (4.2.31)$$

$$G_2 = \frac{1}{4\pi\epsilon_0} \left(\frac{\epsilon_0}{\epsilon_1} \int_0^\infty e^{-k|Z-z'|} J_0(kr) dk + \int_0^\infty F(k) e^{k(Z-z')} J_0(kr) dk + \int_0^\infty N(k) e^{-k(Z-z')} J_0(kr) dk \right), \quad 0 \leq Z \leq h; \quad (4.2.32)$$

$$G_3 = \frac{1}{4\pi\epsilon_0} \int_0^\infty L(k) e^{-k(Z-z')} J_0(kr) dk, \quad Z > h; \quad (4.2.33)$$

where $z' \geq 0$. $M(k)$, $N(k)$, $F(k)$ and $L(k)$ are coefficients to be determined by the boundary conditions for all values of r from 0 to ∞ . This requires that the integrands alone satisfy these conditions (Panofsky and Phillips, 1962).

At $Z = 0$, $G_1 = G_2$ and $\epsilon_0 \frac{dG_1}{dz} = \epsilon_1 \frac{dG_2}{dz}$, leading, respectively, to

$$M(k) e^{-kz'} = \frac{\epsilon_0}{\epsilon_1} e^{-kz'} + N(k) e^{kz'} + F(k) e^{-kz'}, \quad (4.2.34)$$

$$\epsilon_0 M(k) e^{-kz'} = \epsilon_0 e^{-kz'} - \epsilon_1 N(k) e^{kz'} + \epsilon_1 F(k) e^{-kz'}. \quad (4.2.35)$$

At $Z = h$, $G_2 = G_3$ and $\epsilon_1 \frac{dG_2}{dz} = \epsilon_2 \frac{dG_3}{dz}$, leading, respectively, to

$$L(k) e^{-k(h-z')} = \frac{\epsilon_0}{\epsilon_1} e^{-k(h-z')} + N(k) e^{-k(h-z')} + F(k) e^{k(h-z')}, \quad (4.2.36)$$

$$-\varepsilon_2 L(k) e^{-k(h-z')} = -\varepsilon_0 e^{-k(h-z')} - \varepsilon_1 N(k) e^{-k(h-z')} + \varepsilon_1 F(k) e^{k(h-z')}. \quad (4.2.37)$$

Solving Eqs. (4.2.34) – (4.2.37) for $M(k), N(k), F(k)$ and $L(k)$, we have

$$\begin{aligned} N(k) &= \frac{(1 - \varepsilon_1/\varepsilon_0)(\varepsilon_0/\varepsilon_1 - \varepsilon_0/\varepsilon_2) e^{-kh} + (1 - \varepsilon_0/\varepsilon_1)(1 + \varepsilon_1/\varepsilon_2) e^{k(h-2z')}}{(1 + \varepsilon_1/\varepsilon_0)(1 + \varepsilon_1/\varepsilon_2) e^{kh} - (1 - \varepsilon_1/\varepsilon_0)(1 - \varepsilon_1/\varepsilon_2) e^{-kh}}, \\ F(k) &= \frac{(1 + \varepsilon_1/\varepsilon_0)(\varepsilon_0/\varepsilon_1 - \varepsilon_0/\varepsilon_2) e^{-k(h-2z')} + (1 - \varepsilon_1/\varepsilon_2)(1 - \varepsilon_0/\varepsilon_1) e^{-kh}}{(1 + \varepsilon_1/\varepsilon_0)(1 + \varepsilon_1/\varepsilon_2) e^{kh} - (1 - \varepsilon_1/\varepsilon_0)(1 - \varepsilon_1/\varepsilon_2) e^{-kh}}. \end{aligned} \quad (4.2.38)$$

$M(k)$ and $F(k)$ are then obtained by substituting Eq. (4.2.38) into Eq. (4.2.34) and Eq. (4.2.36). By substituting Eq. (4.2.38) into Eq. (4.2.32) one obtains Green's function in the thin film.

$$\begin{aligned} G_2 &= \frac{1}{4\pi\varepsilon_0} \left(\frac{\varepsilon_0}{\varepsilon_1} \int_0^\infty e^{-k|Z-z'|} J_0(kr) dk + \int_0^\infty F(k) e^{k(Z-z')} J_0(kr) dk \right. \\ &\quad \left. + \int_0^\infty N(k) e^{-k(Z-z')} J_0(kr) dk \right) \\ &= \frac{1}{4\pi\varepsilon_1} \frac{1}{\left[(Z-z')^2 + r^2 \right]^{1/2}} \\ &\quad + \frac{1}{4\pi\varepsilon_0} \sum_{n=0}^\infty \left(\frac{\alpha_1}{\left[(z_1 + 2nh)^2 + r^2 \right]^{1/2}} + \frac{\alpha_2}{\left[(z_2 + 2nh)^2 + r^2 \right]^{1/2}} \right. \\ &\quad \left. + \frac{\alpha_3}{\left[(z_3 + 2nh)^2 + r^2 \right]^{1/2}} + \frac{\alpha_4}{\left[(z_4 + 2nh)^2 + r^2 \right]^{1/2}} \right) \beta^n, \end{aligned} \quad (4.2.39)$$

where

$$\begin{aligned} \alpha_1 &= \frac{(1 - \varepsilon_1/\varepsilon_0)(\varepsilon_0/\varepsilon_1 - \varepsilon_0/\varepsilon_2)}{(1 + \varepsilon_1/\varepsilon_0)(1 + \varepsilon_1/\varepsilon_2)}, \quad \alpha_2 = \frac{1 - \varepsilon_0/\varepsilon_1}{1 + \varepsilon_1/\varepsilon_0}, \\ \alpha_3 &= -\frac{(1 - \varepsilon_0/\varepsilon_1)(1 - \varepsilon_1/\varepsilon_2)}{(1 + \varepsilon_1/\varepsilon_0)(1 + \varepsilon_1/\varepsilon_2)}, \quad \alpha_4 = \frac{\varepsilon_0/\varepsilon_2 - \varepsilon_0/\varepsilon_1}{1 + \varepsilon_1/\varepsilon_2}, \end{aligned} \quad (4.2.40)$$

and

$$z_1 = 2h + Z - z', \quad z_2 = Z + z', \quad z_3 = 2h - Z + z', \quad z_4 = 2h - (Z + z'). \quad (4.2.41)$$

Eq. (4.2.39) gives the electric potential at (Z, r, θ) in the thin film due to a unit positive charge at $Z=z', r=0$. If the unit positive charge is located at (z', r', θ') , the electric potential at (Z, R, θ) is given by Eq. (4.2.39) with replacement as follows:

$$r = \left[R^2 + (r')^2 - 2Rr' \cos(\theta - \theta') \right]^{1/2}. \quad (4.2.42)$$

• *Depolarization field of a ferroelectric domain*

For a ferroelectric domain with uniform spontaneous polarization, it is clear from Eq. (4.2.30) that the charges ρ_p induced are concentrated on the domain boundary. Since the vacuum and substrate are assumed to extend to infinity in our analysis, one can express, in terms of Green function, the solution of the electric potential created by ρ_p as follows:

$$\begin{aligned} \Phi(\mathbf{x}) &= \iiint_V G_2(\mathbf{x}, \mathbf{x}') \rho_p d\mathbf{v} = - \iiint_V G_2(\mathbf{x}, \mathbf{x}') P_{i,i'} d\mathbf{v} \\ &= \iiint_V G_{2,i'}(\mathbf{x}, \mathbf{x}') P_i d\mathbf{v} = P_i \iint_S G_2(\mathbf{x}, \mathbf{x}') n_i ds, \end{aligned} \quad (4.2.43)$$

where S is the surface of the domain, and Eq. (4.2.32) gives the Green function $G_2(\mathbf{x}, \mathbf{x}')$. In deriving Eq. (4.2.43), we have assumed that P_i is restricted in a finite region and the boundary term vanishes. Eq. (4.2.43) can be used to calculate the depolarization electric potential created by any shape domain by substituting Eq. (4.2.39) into it. Under action of the charged AFM tip, we can assume that a cylindrical domain forms normal to the surface from $z=0$ to a , and that the spontaneous polarization direction is also normal to the surface (z direction). The integration of Eq. (4.2.43) reduces to integration on end sections $z=0$ and $z=a$ only, except that a delta function multiplied by a constant has to be added to obtain the electric potential inside the upper and lower circular surfaces. It is experimentally found that the surface polarization charges are substantially neutralized due to charge transfer. In such a case, instead of P_i , one has to substitute the remaining charges on the surface for integration on the end section, $z=0$. The depolarization energy of the ferroelectric domain created by the polarization charges is given by

$$U_{dp} = \frac{1}{2} \left[\iint_{z=0} q_1 \Phi(\mathbf{x}) ds + \iint_{z=a} q_2 \Phi(\mathbf{x}) ds \right], \quad (4.2.44)$$

where q_1 and q_2 are polarization charges at $z=0$ and a , respectively.

In this part, we have derived explicit expressions for the AFM-generated electric fields in ferroelectric thin films. The depolarization field created by the

polarization charges in ferroelectrics was also obtained. The effects of the thin film thickness and of the dielectric properties of the film and the substrate on switching properties of the near surface domain can be analyzed accordingly (Wang and Woo, 2003b).

4.3 Energy Expressions

4.3.1 Energy Expression for 180° Domain in a Ferroelectric Film Covered with Top and Bottom Electrodes

In this section, consider the energy expression for a 180° domain in a ferroelectric thin film covered with top and bottom electrodes. The ferroelectric domain under consideration is assumed to have the shape of a prolate semi-spheroid that intersects the upper electrode in a circle of radius a , and have its tip at a distance c from the electrode (Fig. 4.12).

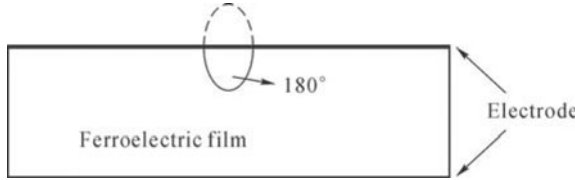


Fig. 4.12 Schematic of the half of a prolate spheroidal domain

The other electrode is assumed to be far away, so that its effects on the depolarization energy can be neglected. Since the electrode on which the domain originates is an equipotential surface, the depolarization field created by the prolate semi-spheroidal domain, based on the image approach, is the same as that created by a complete prolate spheroidal domain in an infinite dielectric matrix. Such a depolarization field inside the domain has been obtained above:

$$E_3 = -\frac{P_s}{2\pi\epsilon} I_3, \quad E_1 = E_2 = 0, \quad (4.3.1)$$

where for the ellipsoidal domain, I_3 is given by

$$I_3 = \begin{cases} 4\pi ca^2 \left\{ \log \left[c/a + \sqrt{(c/a)^2 - 1} \right] - \sqrt{c^2 - a^2}/c \right\} / (c^2 - a^2)^{3/2}, & c > a \\ 4\pi/3, & c = a \\ 4\pi a^2 c \left[\sqrt{(a/c)^2 - 1} - \arccos(c/a) \right] / (a^2 - c^2)^{3/2}, & c < a \end{cases} \quad (4.3.2)$$

where the energy expression holds for all three cases, even if the initial shape of the domain is assumed to be a prolate semi-spheroid. When a sufficiently large electrical field is applied, a 180° domain will nucleate and expand. The evolution of the domain is driven by the free energy reduction in the process. Since no strain is involved in the 180° domain, only the electrostatic energy and domain-wall energy enter into our analysis. If the coercive field E^C of the ferroelectric material plays the role of the external electric field, the Gibbs free energy change can be written as (Wang and Woo, 2003a)

$$\Delta G = G - G_0 = -\frac{1}{2} \iiint_{\Omega} E_i P_i^* dv - \iiint_{\Omega} E_i^C P_i^* dv = \frac{4\pi}{3} ca^2 \left(\frac{P_s^2}{2\pi\epsilon} I_3 - 2E^C P_s \right), \quad (4.3.3)$$

where $\mathbf{P}^* = (0, 2, 2P_s)$.

Eq. (4.3.3) only gives the electrostatic energy change due to the introduction of a 180° domain. The surface energy of the wall has to be added:

$$\Delta T = \frac{4\pi}{3} ca^2 \left[\frac{P_s^2}{2\pi\epsilon} I_3 - 2E^C P_s \right] + \Gamma(a, c) \sigma, \quad (4.3.4)$$

where σ designates the surface energy per unit area of the wall and Γ is the surface area of an ellipsoidal domain, given by

$$\Gamma(a, c) = \begin{cases} 2\pi a \left[a + \frac{c}{\sqrt{1-a^2/c^2}} \arcsin\left(\sqrt{1-a^2/c^2}\right) \right], & a < c \\ 4\pi a^2, & a = c \\ 2\pi a \left[a + c/\sqrt{(a/c)^2 - 1} \log\left(a/c + \sqrt{(a/c)^2 - 1}\right) \right], & a > c \end{cases} \quad (4.3.5)$$

In reality, the initial polarization charges prior to the induction of the 180° domain can be considered as discharged; thus, one should take $\mathbf{P}^* = P_s \mathbf{k}$ in the above energy expressions.

4.3.2 Energy Expression for 180° Domain in Ferroelectric Film Induced by an AFM Tip Without the Top Electrode

For a cylindrical domain on the bottom electrode, the depolarization energy U_{dp} created by the polarization charges has been calculated earlier. It is given by the

following integral, which we derived above:

$$\begin{aligned}
 U_{dp} = & \frac{P_s^2}{2\pi\epsilon_1} \int_0^{2\pi} \int_0^R \int_0^{2\pi} \int_0^R g_0(r, \theta; r', \theta') rr' dr' d\theta' dr d\theta \\
 & + \frac{P_s^2}{2\pi\epsilon_0} \int_0^{2\pi} \int_0^R \int_0^{2\pi} \int_0^R [\alpha_1 g_1(r, \theta; r', \theta') + \alpha_2 g_2(r, \theta; r', \theta') \\
 & + \alpha_3 g_3(r, \theta; r', \theta') + \alpha_4 g_4(r, \theta; r', \theta')] rr' dr' d\theta' dr d\theta,
 \end{aligned} \quad (4.3.6)$$

where

$$\begin{aligned}
 g_0(r, \theta; r', \theta') &= 2 \left[\frac{1}{\rho} - \frac{1}{(Z^2 + \rho^2)^{1/2}} \right], \\
 g_1(r, \theta; r', \theta') &= \sum_0^\infty \left\{ \frac{2}{[(2h+2nh)^2 + \rho^2]^{1/2}} - \frac{1}{[(2h+2nh+Z)^2 + \rho^2]^{1/2}} - \frac{1}{[(2h+2nh-Z)^2 + \rho^2]^{1/2}} \right\} \beta^n, \\
 g_2(r, \theta; r', \theta') &= \sum_0^\infty \left\{ \frac{1}{[(2Z+2nh)^2 + \rho^2]^{1/2}} - \frac{2}{[(2nh+Z)^2 + \rho^2]^{1/2}} + \frac{1}{[(2nh)^2 + \rho^2]^{1/2}} \right\} \beta^n, \\
 g_3(r, \theta; r', \theta') &= \sum_0^\infty \left\{ \frac{2}{[(2h+2nh)^2 + \rho^2]^{1/2}} - \frac{1}{[(2h+2nh+Z)^2 + \rho^2]^{1/2}} - \frac{1}{[(2h+2nh-Z)^2 + \rho^2]^{1/2}} \right\} \beta^n, \\
 g_4(r, \theta; r', \theta') &= \sum_0^\infty \left\{ \frac{1}{[(2h+2nh-2Z)^2 + \rho^2]^{1/2}} - \frac{2}{[(2h+2nh-Z)^2 + \rho^2]^{1/2}} + \frac{1}{[(2h+2nh)^2 + \rho^2]^{1/2}} \right\} \beta^n, \\
 \alpha_1 &= \frac{(1-\epsilon_1/\epsilon_0)(\epsilon_0/\epsilon_1)}{(1+\epsilon_1/\epsilon_0)}, \quad \alpha_2 = \frac{(1-\epsilon_0/\epsilon_1)}{(1+\epsilon_1/\epsilon_0)}, \quad \alpha_3 = -\frac{(1-\epsilon_0/\epsilon_1)}{(1+\epsilon_1/\epsilon_0)}, \quad \alpha_4 = -\epsilon_0/\epsilon_1, \\
 \beta &= \frac{1-\epsilon_1/\epsilon_0}{1+\epsilon_1/\epsilon_0}, \\
 \rho &= [r^2 + r'^2 - 2rr' \cos(\theta - \theta')]^{1/2},
 \end{aligned} \quad (4.3.7)$$

and ϵ_1 , ϵ_0 are the dielectric constants of the ferroelectric thin film and vacuum, respectively. It is easy to see that the depolarization energy of the ferroelectric

domain depends on the materials dielectric constant, spontaneous polarization, the domain size parameters R and Z and the film thickness h . In this investigation, we use expression Eq. (4.3.6) for the depolarization energy, which increases with increasing size (Z, R) of the domain. For the cylindrical domain shown in Fig. 4.13, the total domain-wall energy is,

$$U_{dw} = 2\pi RZ\gamma_1 + \pi R^2\gamma_2, \quad (4.3.8)$$

where γ_1, γ_2 are the domain-wall energy per unit area on the cylindrical surface and the bottom surface, respectively. The energy of the upper surface does not change when the polarization reverses. Therefore, it is not included in the sum in Eq. (4.3.8).

The interaction between the AFM tip field and the spontaneous polarization field provides the main force driving the fresh creation of the written domain.

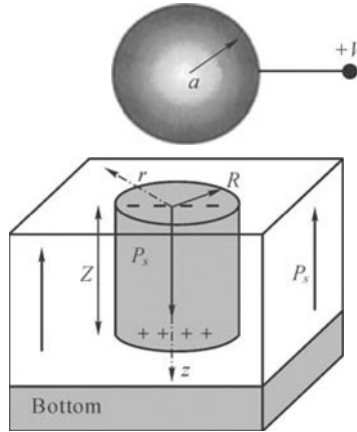


Fig. 4.13 Schematic of a 180° domain created by an AFM tip

When the tip is removed, there appears to be no driving force to stabilize the domain. However, the reversal of the polarization field has to be accompanied with the motion of the domain wall, which is a dissipative process characterized by a coercive field E_c . In such a case, one can treat E_c as an effective electric field in favor of the new domains, and the interaction energy between E_c and the spontaneous polarization can be written as

$$U_{int} = -2\pi R^2 Z E_c P_s, \quad (4.3.9)$$

where the negative sign means that it is against the shrinkage of the domain. The total energy is given by taking the summation of Eqs. (4.3.6), (4.3.8) and (4.3.9):

$$\Delta T = U_{dp} + U_{dw} + U_{\text{int}}. \quad (4.3.10)$$

Under the assumption of a cylindrical domain, the two parameters R and Z completely describe its geometry. If the initial polarization charges are considered to be discharged before introducing the 180° domain, the spontaneous polarization should take $1/2$ of its value.

4.4 Driving Force and Evolution Equations of Domain Growth

In this section, consider a general shape of the 180° domain wall in an electric field. With the movement of the surface, the free energy will be reduced and the released energy serves as the driving force to compensate for the energy dissipation due to the surface energy increase and the finite velocity of the movement. In this part, a general form of the energy release rate will be used to establish the evolution equation.

As shown in Fig. 4.14, the body containing a surface of discontinuity in an electromagnetic field can be divided into two regions, i.e., inside and outside the surface. The internal energy of an electromagnetic field is given by

$$W = \frac{1}{2} \iiint_v (\mathbf{E} \cdot \mathbf{D} + \mathbf{H} \cdot \mathbf{B}) dv = \iiint_{\Omega(t)} \omega dv + \iiint_{v-\Omega(t)} \omega dv = W_1 + W_2, \quad (4.4.1)$$

where $\Omega(t)$, v are the volumes inside the moving surface and the whole body, respectively. $\omega = 1/2 (\mathbf{E} \cdot \mathbf{D} + \mathbf{H} \cdot \mathbf{B})$ is the energy density. For the first integral in Eq. (4.4.1), the region of integration and the integrand are the function of time t . Using Reynolds' transport theorem, its change rate is obtained as

$$\frac{dW_1}{dt} = \frac{d}{dt} \iiint_{\Omega(t)} \omega dv = \iiint_{\Omega(t)} [\dot{\omega} + \text{div}(\omega \mathbf{u})] dv = \iiint_{\Omega(t)} \dot{\omega} dv + \iint_s \omega \mathbf{u} \cdot \mathbf{n} ds, \quad (4.4.2)$$

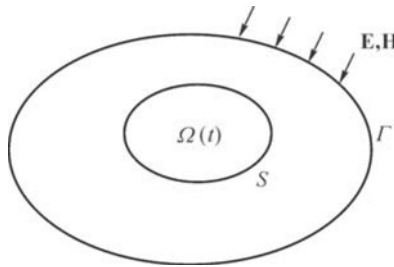


Fig. 4.14 A surface of discontinuity in an electromagnetic field

where $\mathbf{u}$ is the velocity of the surface movement and S^+ is the internal side of the surface. Eq. (4.4.2) asserts that the changing rate of the integral ω over $\mathcal{Q}(t)$ is equal to the rate computed as if $\mathcal{Q}(t)$ were fixed in its current position plus the rate at which ω is carried out of this region across its boundary. The rate of the energy density is given by

$$\begin{aligned}\dot{\omega} &= \frac{1}{2} \frac{d}{dt} (\mathbf{E} \cdot \mathbf{D} + \mathbf{H} \cdot \mathbf{B}) = \dot{\mathbf{E}} \cdot \dot{\mathbf{D}} + \dot{\mathbf{H}} \cdot \dot{\mathbf{B}} = \mathbf{E} \cdot (\nabla \times \mathbf{H} - \mathbf{J}) + \mathbf{H} \cdot (-\nabla \times \mathbf{E}) \\ &= -\mathbf{E} \cdot \mathbf{J} - [\mathbf{H} \cdot (\nabla \times \mathbf{E}) - \mathbf{E} \cdot (\nabla \times \mathbf{H})] = -\mathbf{E} \cdot \mathbf{J} - \text{div}(\mathbf{E} \times \mathbf{H}).\end{aligned}\quad (4.4.3)$$

In deriving Eq. (4.4.3), we have used Maxwell's equation to replace the time-derivative factors in the integrand, $\mathbf{\Sigma} = \mathbf{E} \times \mathbf{H}$, where $\mathbf{\Sigma}$ is called the Poynting vector, which is identified as the energy flux, or rate of energy flow per unit area. Substituting Eq. (4.4.3) into Eq. (4.4.2) and considering the current $\mathbf{J} = 0$ inside the region, we have

$$\frac{dW_1}{dt} = \iint_{S^+} (\omega \mathbf{u} - \mathbf{\Sigma}) \cdot \mathbf{n} ds. \quad (4.4.4)$$

In a similar way, one can derive the rate of the second integral in Eq. (4.4.1) as follows:

$$\frac{dW_2}{dt} = - \iint_{\Gamma} \mathbf{\Sigma} \cdot \mathbf{n} ds + \iint_{S^-} (\mathbf{\Sigma} - \omega \mathbf{u}) \cdot \mathbf{n} ds, \quad (4.4.5)$$

where S^- is the external side of the surface and Γ is the outside boundary of the body. The changing rate of the energy provided by the environment across the outside boundary of the body is given by

$$\dot{U} = - \iint_{\Gamma} \mathbf{\Sigma} \cdot \mathbf{n} ds. \quad (4.4.6)$$

Therefore, the energy release rate in the process of the surface movement can be obtained as

$$\dot{G} = \dot{U} - \dot{W} = \iint_S \{ [\omega] u_i - [\Sigma_i] \} n_i ds, \quad (4.4.7)$$

where $[\omega] = \omega^O - \omega^I$, $[\Sigma_i] = \Sigma_i^O - \Sigma_i^I$ are the jumps of corresponding quantities across the surface. Eshelby (1970) proposed the well-known energy-momentum tensor as the driving force for a defect to propagate in materials. In fact, Eq. (4.4.7)

represents its counterpart in an electromagnetic field. Similarly, we can define the energy-momentum vector as

$$P_i = \omega u_i - \Sigma_i. \quad (4.4.8)$$

The energy release rate $\dot{G}$ represents the driving force for the growth of the 180° domain, which we will adopt in the following domain stability problem (Wang and Xiao, 2000).

The energy release rate expression derived above can be used to establish the evolution equations for any shape of ferroelectric domains. For a spheroidal domain or a cylindrical domain, two parameters, a and c (Z and R) completely describe the geometry of the domain. For such a shape of a domain, the explicit energy expression has been obtained in Section 4.3. Therefore, the rate of evolution of the domain then depends on the driving forces which are the free energy reduction rates accompanying the change in a and c , i.e.

$$\begin{aligned} f_a &= -\frac{\partial(\Delta T)}{\partial a}, \\ f_c &= -\frac{\partial(\Delta T)}{\partial c}, \end{aligned} \quad (4.4.9)$$

where ΔT is the free energy.

According to near-equilibrium thermodynamics, the evolution rate can be reasonably assumed to be proportional to the corresponding driving force,

$$\begin{aligned} \frac{\partial c}{\partial t} &= M f_c = -M \frac{\partial(\Delta T)}{\partial c}, \\ \frac{\partial a}{\partial t} &= M f_a = -M \frac{\partial(\Delta T)}{\partial a}, \end{aligned} \quad (4.4.10)$$

where M is a material constant. And Eq. (4.4.10) is the evolution equation of the ferroelectric domain evolution.

4.5 Stability Analysis

Eq. (4.4.10) is a system of nonlinear differential equations, which we will not attempt to solve. Our interest is the conditions of stability of the solution. Knowing that $(a, c) = (0, 0)$ is a stationary point, as we determine the condition under which the back-switching can be avoided, we find its corresponding area of attraction. All solutions of Eq. (4.4.10) starting in this area must end up at this

point.

It is easy to understand that if $f_a > 0, f_c > 0$ at some point (a, c) , the back-switching will not happen, and if $f_a < 0, f_c < 0$, the domain will definitely shrink. And the expansion and shrinkage boundary always sits in the range $f_a > 0, f_c < 0$ or $f_a < 0, f_c > 0$. That is why it is much more difficult to determine the domain of attractions for a system of differential equations with more than one variable. The condition $f_a > 0$ and $f_c > 0$ will give the upper limit of the stable domain, whereas $f_a < 0$ and $f_c < 0$ will give the lower limit of the size of the domain. In this section, for a spheroidal 180° domain in an infinite ferroelectric material, we can establish a Liapunov function $V(a, c) = a^2 + c^2 > 0$, for $a > 0, c > 0$ to determine the domain of attraction around $a = 0, c = 0$. The domain of attraction, or the strict stability domain, in which $\frac{dV}{dt} < 0$, can be determined by the solution of the following equation

$$\frac{dV}{dt} = \frac{\partial V}{\partial a} \cdot \frac{\partial a}{\partial t} + \frac{\partial V}{\partial c} \cdot \frac{\partial c}{\partial t} = 0. \quad (4.5.1)$$

The physical meaning of the domain of attraction is that for any given initial state in this domain, the final state of the system will be $a = 0, c = 0$, which determines the smallest sizes of the ferroelectric domain to avoid the back-switching after the external electric field is removed.

In the following, determine the smallest sizes of the 180° domain in the ferroelectric film covered and not covered by electrodes, respectively. Taking PZT material as an example (Scott and Paz de Araujo, 1989), assume a coercive field of $60 \text{ kV} \cdot \text{cm}^{-1}$, spontaneous polarization of $P_s = 30 \text{ } \mu\text{C} \cdot \text{cm}^{-2}$, a relative dielectric constant of $\epsilon_r = 1000$, and the domain wall energy of $\sigma = 4 \text{ mJ} \cdot \text{m}^{-2}$. And assuming that the initial polarization charges were discharged before introducing the 180° domain, one can carry out the following calculation.

Case 1: A 180° domain in a ferroelectric film covered with top and bottom electrodes.

In such a case, the evolution equation of a and c can be established explicitly by substituting Eq. (4.3.4) into Eq. (4.4.10). Then substituting into Eq. (4.5.1) yields the domain of attraction (Fig. 4.15). In the same figure, the areas in which $f_a > 0, f_c > 0$ and $f_a < 0, f_c < 0$ are also shown. It can be seen that, in this case, the minimum stable size of the 180° domain is about $a \approx 2 \text{ nm}$, $c \approx 10 \text{ nm}$. Any solution that starts from this region will eventually end up at the origin, namely shrink to zero.

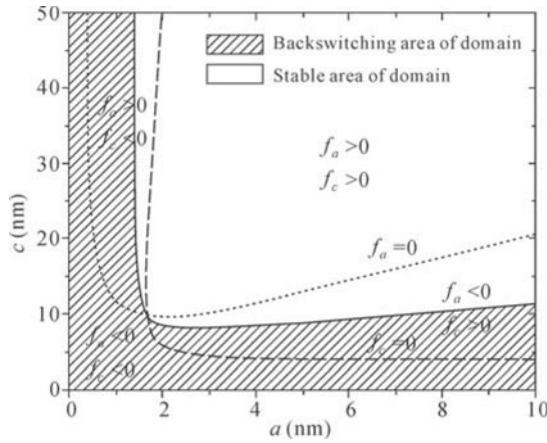


Fig. 4.15 Unstable area of 180° domain in ferroelectric thin film covered by electrodes

Case 2: A 180° domain in ferroelectric film induced by an AFM tip without the top electrode.

In such a case, the evolution equation of a and c can be established by substituting Eq. (4.3.10) into Eq. (4.4.10). Then substituting into Eq. (4.4.11) yields the domain of attraction (Fig. 4.16) under such conditions. The areas in which $f_a > 0$, $f_c > 0$ and $f_a < 0$, $f_c < 0$ are also shown. In this case, the minimum stable size of the 180° domain is about $a \approx 10$ nm, $c \approx 20$ nm. The stable domain should be larger than the one with the electrode, since there is no bound charge to tie the polarization charges without the top electrode. One should also bear in mind that a is the radius of the 180° domain, and it is also the minimum stable size of the domain.

In summary, a general approach was established to determine the minimum sizes of the stable domain to avoid back switching in ferroelectric thin film by stability analysis on the evolution equations (Wang et al., 2006).

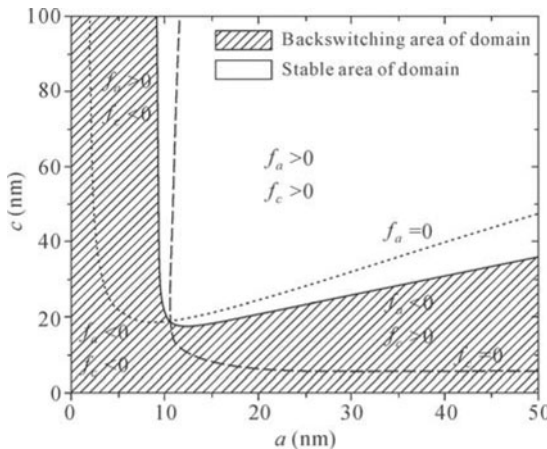


Fig. 4.16 Unstable area of 180° domain in ferroelectric thin film without top electrode

4.6 Remarks

Based on evolution equations, we present a general formulation for the investigation of the stability of microstructures. If only one parameter x is needed to describe the kinetic equation, $\frac{dx}{dt} = f(x)$.

One can determine its stable point and the associated attraction region by considering its right-hand side. And if more than two parameters are needed to describe the system evolution, the stability properties of the system can then be analyzed by using the Liapunov function based on the evolution equations. As an important application, the method is used to determine the critical dimensions in which switched domains in ferroelectric materials will switch back upon removal of the external voltage. Besides the obvious scientific interest, such a theory can be used to provide the necessary guidance for the production of stable switched domains. This is especially important for the fabrication of ferroelectric memory devices. Although the analysis has focused on an ellipsoidal domain in a bulk ferroelectric material, in which an explicit expression for the depolarization energy can be derived, it can be extended to cases of more complicated geometry, in which the depolarization field and energy may be evaluated numerically. The effect of the temperature can also be taken into account through the temperature-dependent spontaneous polarization. In the example addressed above, we assumed that the driving force for shrinkage of the domain is the depolarization energy and the domain-wall energy, and that resistance is provided by the internal electric field. However, the reversal of polarization would necessitate domain-wall motion, which is a dissipative process commonly characterized by a coercive field. Thus one can also replace or add the resistant terms by the coercive field in the energy expressions. Furthermore, the analysis is sufficiently general to be extended to many similar phenomena, such as phase transformation, dislocation loop nucleation, expansion, etc.

Appendix B: Derivation of the Electric and Magnetic Field for a Growing 180° Domain

In any time-invariant reference system of coordinates, the differential form of Maxwell's equations reads as follows:

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \quad (\text{B1a})$$

$$\nabla \times \mathbf{H} = \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t}, \quad (\text{B1b})$$

$$\nabla \cdot \mathbf{D} = \rho, \quad (\text{B1c})$$

$$\nabla \cdot \mathbf{B} = 0, \quad (\text{B1d})$$

where $\mathbf{E}$ is electric field ($\text{V} \cdot \text{m}^{-1}$); $\mathbf{H}$ is magnetic field ($\text{A} \cdot \text{m}^{-1}$); $\mathbf{D}$ is the electric conduction ($\text{C} \cdot \text{m}^{-2}$), $\mathbf{B}$ is the magnetic induction (T), $\mathbf{J}$ is the electric current density ($\text{A} \cdot \text{m}^{-2}$); and ρ is the electric charge density ($\text{C} \cdot \text{m}^{-2}$), and the prescribed source current $\mathbf{J}$ and charge density ρ are related by the continuity equation

$$\nabla \cdot \mathbf{J} + \frac{\partial \rho}{\partial t} = 0. \quad (\text{B2})$$

In the present work, we impose a further simplification by ignoring the anisotropic effect of the dielectric constant. For linear isotropic materials, the constitutive relation can be written as follows:

$$\begin{aligned} \mathbf{D} &= \varepsilon \mathbf{E}, \\ \mathbf{B} &= \mu \mathbf{H}, \end{aligned} \quad (\text{B3})$$

where ε, μ are the permittivity and permeability of the isotropic medium, respectively.

The electric polarization gives rise to the charge distribution with volume density

$$\rho = -\nabla \cdot (2\mathbf{P}_s). \quad (\text{B4a})$$

If the polarization changes with time, similarly, it corresponds to the current density as follows:

$$\rho = \mathbf{J} = \frac{\partial (2\mathbf{P}_s)}{\partial t}. \quad (\text{B4b})$$

By taking the curl of Eq. (B1a) and substituting $\nabla \times \mathbf{H}$ from Eq. (B1b), we obtain

$$\nabla \times \nabla \times \mathbf{E} = -\mu \varepsilon \frac{\partial^2 \mathbf{E}}{\partial t^2} - \mu \frac{\partial \mathbf{J}}{\partial t}. \quad (\text{B5})$$

A combination of Eqs. (B1c) and (B5) yields

$$\nabla^2 \mathbf{E} - \frac{1}{c^2} \frac{\partial^2 \mathbf{E}}{\partial t^2} = \mu \frac{\partial \mathbf{J}}{\partial t} + \frac{1}{\varepsilon} \nabla \rho, \quad c^2 = 1/(\varepsilon \mu). \quad (\text{B6a})$$

Substitution of Eq. (B4) into Eq. (B6a) leads to

$$\nabla^2 \mathbf{E} - \frac{1}{c^2} \frac{\partial^2 \mathbf{E}}{\partial t^2} = \left(\frac{1}{c^2} \frac{\partial^2}{\partial t^2} \mathbf{I} - \nabla \nabla \right) \cdot \frac{2\mathbf{P}_s}{\varepsilon}, \quad (\text{B6b})$$

where $\mathbf{I}$ is the diagonal unit matrix. The above equation relates the unknown electric field with the source field $2\mathbf{P}_s/\varepsilon$. By introducing the Hertz potential $\mathbf{\Pi}$ as follows:

$$\begin{bmatrix} E_1 \\ E_2 \\ E_3 \end{bmatrix} = - \begin{bmatrix} \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2} & -\frac{\partial^2}{\partial x \partial y} & -\frac{\partial^2}{\partial x \partial z} \\ -\frac{\partial^2}{\partial x \partial y} & \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial y^2} & -\frac{\partial^2}{\partial y \partial z} \\ -\frac{\partial^2}{\partial x \partial z} & -\frac{\partial^2}{\partial y \partial z} & \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial z^2} \end{bmatrix} \begin{bmatrix} \Pi_1 \\ \Pi_2 \\ \Pi_3 \end{bmatrix}, \quad (\text{B7})$$

we have

$$\nabla^2 \mathbf{\Pi} - \frac{1}{c^2} \frac{\partial^2 \mathbf{\Pi}}{\partial t^2} = -\frac{2\mathbf{P}_s}{\varepsilon}. \quad (\text{B8})$$

Once the Hertz potential $\mathbf{\Pi}$ is determined, the electric field can be found through Eq. (B7), while the magnetic field is computed from Eq. (B1a):

$$\xi \mathbf{H} = \nabla \times \frac{1}{c} \frac{\partial \mathbf{\Pi}}{\partial t}, \quad \xi = \sqrt{\mu/\varepsilon}. \quad (\text{B9})$$

For such a problem, we first need to derive the solution for the following equation subjected to the initial condition $G(\mathbf{r}, t < 0) = 0$ and to the radiation condition at infinity as the Green function

$$\nabla^2 G(\mathbf{r}, t) - \frac{1}{c^2} \frac{\partial^2 G(\mathbf{r}, t)}{\partial t^2} = -\delta(\mathbf{r})\delta(t). \quad (\text{B10})$$

The solution of Eq. (B10) in a spherical system of coordinates is

$$G(\mathbf{r}, t) = \frac{\delta(t - r/c)}{4\pi r} = \frac{\delta(t^*)}{4\pi r}, \quad (\text{B11})$$

where $t^* = t - r/c$ is the retarded time, $\mathbf{r}$ is the position vector and r is its magnitude.

Green function represents the response of the system described by Eq. (B10) to a localized unit source of space-time density $\delta(\mathbf{r})\delta(t)$. It can be used to derive the final expression for the field radiated by an arbitrary source distribution. For the present case, the applied source is the dipole density $2\mathbf{P}(\mathbf{r},t)/\varepsilon$. The total Hertz potential can be obtained by superposition as shown in Eqs. (B8) and (B10):

$$\begin{aligned}\Pi &= 2 \iiint d\mathbf{v}' \int dt' \frac{\mathbf{P}_s(\mathbf{r}',t')}{\varepsilon} G(\mathbf{r}-\mathbf{r}',t-t') \\ &= \frac{1}{2\pi\varepsilon} \iiint d\mathbf{v}' \int dt' \mathbf{P}_s(\mathbf{r}',t') \times \frac{\delta(t-t'-|\mathbf{r}-\mathbf{r}'|/c)}{|\mathbf{r}-\mathbf{r}'|} = \frac{1}{2\pi\varepsilon} \iiint d\mathbf{v}' \frac{\mathbf{P}_s(\mathbf{r}',t^*)}{|\mathbf{r}-\mathbf{r}'|},\end{aligned}\quad (\text{B12})$$

where $t^* = t - |\mathbf{r}-\mathbf{r}'|/c$. For the current problem, the polarization source is restricted to the finite region. Generally speaking, since c is the light speed, $ct \gg |\mathbf{r}-\mathbf{r}'|$. Therefore the polarization can be expressed using the Heaviside function as follows:

$$\mathbf{P}_s(\mathbf{r},t^*) \approx \mathbf{P}_s(\mathbf{r},t) = p_s \mathbf{k} H[\mathcal{Q}(t)], \quad (\text{B13})$$

where

$$H[\mathcal{Q}(t)] = \begin{cases} 1, & \mathbf{r} \in \mathcal{Q}(t) \\ 0, & \text{others} \end{cases}, \quad (\text{B14})$$

in which $\mathcal{Q}(t)$ is the region of the 180° domain. Since the domain is expanding, it is a function of time. In what follows, we assume that the 180° domain takes an ellipsoidal shape, and derive the electric and magnetic field both inside and outside the domain.

Substitution of Eq. (B13) into Eq. (B12) yields

$$\Pi = \frac{p_s \mathbf{k}}{2\pi\varepsilon} \iiint_{\mathcal{Q}(t)} d\mathbf{v}' \frac{1}{|\mathbf{r}-\mathbf{r}'|}. \quad (\text{B15})$$

If the 180° domain is assumed to take an ellipsoidal shape with principle half axes a_1, a_2, a_3 , the integral in Eq. (B15) can be expressed in terms of the following simple formula, as shown in Mura:

$$\Pi = \frac{p_s \mathbf{k}}{2\pi\varepsilon} \iiint_{\mathcal{Q}(t)} d\mathbf{v}' \frac{1}{|\mathbf{r}-\mathbf{r}'|} = \frac{p_s \mathbf{k}}{4\pi\varepsilon} [I(\lambda) - x_n x_n I_n(\lambda)], \quad (\text{B16})$$

where $x_n x_n I_N(\lambda) = x_1^2 I_1(\lambda) + x_2^2 I_2(\lambda) + x_3^2 I_3(\lambda)$, and the expressions of λ , I , I_1 , I_2 , I_3 are given as follows:

$$\begin{aligned} I(\lambda) &= 2\pi a_1 a_2 a_3 \int_{\lambda}^{\infty} \frac{ds}{\Delta(s)}, \\ I_i(\lambda) &= 2\pi a_1 a_2 a_3 \int_{\lambda}^{\infty} \frac{ds}{(a_i^2 + s)\Delta(s)}, \end{aligned} \quad (\text{B17})$$

where $\Delta(s) = [(a_1^2 + s)(a_2^2 + s)(a_3^2 + s)]^{1/2}$, and λ is the largest position root of the equation

$$\frac{x_1^2}{a_1^2 + \lambda} + \frac{x_2^2}{a_2^2 + \lambda} + \frac{x_3^2}{a_3^2 + \lambda} = 1. \quad (\text{B18})$$

Thus

$$\begin{aligned} I(\lambda) &= 4\pi a_1 a_2 a_3 (a_1^2 - a_3^2)^{-1/2} F(\theta, k), \\ I_1(\lambda) &= 4\pi a_1 a_2 a_3 (a_1^2 - a_2^2)^{-1} (a_1^2 - a_3^2)^{-1/2} [F(\theta, k) - E(\theta, k)], \\ I_2(\lambda) &= 4\pi a_1 a_2 a_3 \left[(a_1^2 - a_3^2)^{1/2} (a_1^2 - a_2^2)^{-1} (a_2^2 - a_3^2)^{-1} E(\theta, k) \right. \\ &\quad \left. - (a_1^2 - a_2^2)^{-1} (a_1^2 - a_3^2)^{-1/2} F(\theta, k) \right. \\ &\quad \left. - (a_2^2 - a_3^2)^{-1} (a_3^2 + \lambda)^{1/2} (a_1^2 + \lambda)^{-1/2} (a_2^2 + \lambda)^{-1/2} \right], \\ I_3(\lambda) &= 4\pi a_1 a_2 a_3 (a_2^2 - a_3^2)^{-1} (a_1^2 - a_3^2)^{-1/2} \\ &\quad \times \left[(a_2^2 + \lambda)^{1/2} (a_1^2 - a_3^2)^{1/2} (a_1^2 + \lambda)^{-1/2} (a_3^2 + \lambda)^{-1/2} - E(\theta, k) \right], \end{aligned} \quad (\text{B19})$$

where the first and second elliptic integrals are given by

$$\begin{aligned} F(\theta, k) &= \int_0^{\theta} \frac{dw}{(1 - k^2 \sin^2 w)^{1/2}}, \\ E(\theta, k) &= \int_0^{\theta} (1 - k^2 \sin^2 w)^{1/2} dw, \\ \theta &= \arcsin(1 - a_3^2/a_1^2)^{1/2}, \quad k = [(a_1^2 - a_2^2)(a_1^2 - a_3^2)]^{1/2}. \end{aligned} \quad (\text{B20})$$

For some special shapes of the domain, Eq. (B19) becomes an elementary function as shown below.

1) Sphere ($a_1 = a_2 = a_3 = a_4$)

$$I(\lambda) = \frac{4\pi a^3}{(a^2 + \lambda)^{1/2}},$$

$$I_1(\lambda) = I_2(\lambda) = I_3(\lambda) = \frac{4\pi a^3}{3(a^2 + \lambda)^{3/2}}.$$
(B21)

2) Elliptical cylinder ($a_3 \rightarrow \infty$)

$$I(\lambda) = 0, \quad I_1(\lambda) = \frac{4\pi a_1 a_2}{a_2^2 - a_1^2} \left[\frac{(a_2^2 + \lambda)^{1/2}}{(a_1^2 + \lambda)^{1/2}} - 1 \right],$$

$$I_2(\lambda) = \frac{4\pi a_1 a_2}{a_1^2 - a_2^2} \left[\frac{(a_1^2 + \lambda)^{1/2}}{(a_2^2 + \lambda)^{1/2}} - 1 \right], \quad I_3(\lambda) = 0.$$
(B22)

3) Oblate spheroid ($a_1 = a_2 > a_3$)

$$I(\lambda) = \frac{4\pi a_1^2 a_3}{\sqrt{a_1^2 - a_3^2}} \arccos b,$$

$$I_1(\lambda) = I_2(\lambda) = 2\pi a_1^2 a_3 (\arccos b - bd) / (a_1^2 - a_3^2)^{3/2},$$

$$I_3(\lambda) = 4\pi a_1^2 a_3 \left(\frac{d}{b} - \arccos b \right) / (a_1^2 - a_3^2)^{3/2},$$
(B23)

where $b = \sqrt{(a_3^2 + \lambda)/(a_1^2 + \lambda)}$ and $d = \sqrt{(a_1^2 - a_3^2)/(a_1^2 + \lambda)}$.

4) Prolate spheroid ($a_1 > a_2 = a_3$)

$$I(\lambda) = \frac{4\pi a_1 a_3^2}{\sqrt{a_1^2 - a_3^2}} \arccos(h\bar{b}),$$

$$I_1(\lambda) = 4\pi a_1 a_3^2 (\arccos(h\bar{b}) - \bar{b}/\bar{d}) / (a_1^2 - a_3^2)^{3/2},$$

$$I_2(\lambda) = I_3(\lambda) = 2\pi a_1 a_3^2 (\bar{b}\bar{d} - \arccos(h\bar{b})) / (a_1^2 - a_3^2)^{3/2},$$
(B24)

where $\bar{b} = \sqrt{(a_1^2 + \lambda)/(a_3^2 + \lambda)}$ and $\bar{d} = \sqrt{(a_1^2 - a_3^2)/(a_3^2 + \lambda)}$.

If the point is inside the domain, $\lambda = 0$.

References

- [1] Ahn, C.H., Tybell, T., Antognazza, L., Char, K., Hammond, R.H., Beasley, M.R., Fischer, O., Triscone, J.M.: Local, nonvolatile electronic writing of epitaxial $\text{Pb}(\text{Zr}_{0.52}\text{Ti}_{0.48})\text{O}_3/\text{SrRuO}_3$ heterostructures. *Science* **276**, 1100-1103 (1997)
- [2] Kalinin, S.V., Morozovska, A.N., Chen, L.Q., Rodriguez, B.J.: Local polarization dynamics in ferroelectric materials. *Rep. Prog. Phys.* **73**, 056502 (2010)
- [3] Landauer, R.: Electrostatic considerations in BaTiO_3 domain formation during polarization reversal. *J. Appl. Phys.* **28**, 227-234 (1957)
- [4] Luthi, R., Haefke, H., Grütter, P., Güntherodt, H.J., Szczesniak, L., Meyer, K.P.: Surface and domain structures of ferroelectric GASH crystals studied by scanning force microscopy. *Surf. Sci.* **285**, L498-L502 (1993)
- [5] Panofsky, W.K.H., Phillips, M.: *Classical Electricity and Magnetism*. Addison-Wesley, Reading, MA (1962)
- [6] Scott, J.F., Paz de Araujo, C.A.: Ferroelectric memories. *Science* **246**, 1400-1405 (1989)
- [7] Song, Y.J., Koo, B.J., Lee, J.K., Kim, C.J., Jang, N.W., Kim, H.H., Jung, D.J., Lee, S.Y., Kim, K.: Electrical properties of highly reliable plug buffer layer for high-density ferroelectric memory. *Appl. Phys. Lett.* **80**, 2377-2379 (2002)
- [8] Wang, B., Woo, C.H.: Stability of 180° domain in ferroelectric thin films. *J. Appl. Phys.* **94**, 610-617 (2003a)
- [9] Wang, B., Woo, C.H.: Atomic force microscopy-induced electric field in ferroelectric thin films. *J. Appl. Phys.* **94**, 4053-4059 (2003b)
- [10] Wang, B., Xiao, Z.M.: On the dynamic growth of a 180° domain in a ferroelectric material. *J. Appl. Phys.* **88**, 1464-1472 (2000)
- [11] Wang, B., Woo, C.H., Zheng, Y.: Stability analysis of 180° domains in ferroelectric thin films. *Solid Mechanics and Its Applications* **127**, 57-67 (2006)

Size and Surface Effects of Phase Transition on Nanoferroelectric Materials

5.1 Introduction and Overview of Ferroelectrics in Nanoscale Dimensions

Ferroelectricity is a collective phenomenon, the characteristics of which depend on the combined effects of many factors, such as the ambient temperature, boundary conditions, sample dimensions, misfit epitaxial stresses, external mechanical loads, defects, interface and surface effects, etc. This is particularly true for the state of ferroelectricity itself. Therefore, studies of the effects of nanoscale dimensions on the polarization, dielectric constants, Curie temperature, critical sizes and other critical properties of ferroelectrics have received much attention over the last few decades. Recently, research in this area has found new relevance due to the surge of nanotechnological interest in ultra-miniaturized electronics, nanoscale piezoelectric transducers, ultra-high density memory devices and electro-optic devices, etc. More importantly, many experimental and theoretical works on nanoscale ferroelectrics, such as ferroelectric thin films (FTF), nanoparticles (FNP), nanowires (FNW), nanorods (FNR), nanodisks (FND), nanotubes (FNT), nanoislands (FNI) and nanograins (FNG), indicate strong size and surface effects on their critical properties.

5.1.1 Ferroelectric Thin Films in Nanoscale Dimensions

For ferroelectric thin film, its properties, such as its spontaneous polarization and Curie-Weiss relation, are known to depend on the following factors, including temperature, film thickness, misfit epitaxial strain, electromechanical boundary

conditions, nature of the electrodes and interfacial properties. For example, the Curie temperature will generally decrease with a decrease in the thickness of the thin film, and if the size is reduced to some critical size, the phase transition temperature may reduce to 0 K. Thus the intrinsic ferroelectric material will lose its ferroelectricity totally. This limiting thickness or size for different shapes of material is called the critical size of ferroelectricity. Early studies of the size and surface effects on critical properties of ferroelectric thin films gave the impression that the critical size of the ferroelectric thin films was a few tens of nanometers. More recently, many important experiments and theoretical calculations showed that the critical size was much smaller.

Experimentally, Tybell et al. (1999) demonstrated the presence of a stable polarization in a 4-nm-thick epitaxial film of the perovskite ferroelectric $\text{PbZr}_{0.2}\text{Ti}_{0.8}\text{O}_3$ at room temperature by combining electric-force microscopy and piezoelectric scanning-probe microscopy. Streiffer et al. (2002) adopted X-ray scattering and observed ferroelectric properties of epitaxial PbTiO_3 films down to 1.2 nm, showing 180° stripe domains on SrTiO_3 insulator substrate, but a monodomain structure on the conducting SrRuO_3 electrode (Fong et al., 2004).

Theoretically, characteristics governing the para/ferro-electric transitions in ferroelectrics on a nanoscale have been investigated within the framework of the thermodynamic theory, molecular dynamic simulations and first-principle calculations, and so on. However, a clear picture has yet to be completed, and there are various discrepancies between experimental measurements and theoretical calculations. Some theoretical calculations themselves still remain unresolved. Although it is fairly plausible that the critical dimensions have their origin in the size, surface or interface effects on the para/ferroelectric transition, the details of the competing mechanisms, their relative importance and their relation to various electro-mechanical surface conditions, have yet to be clarified.

5.1.2 Ferroelectric Tunneling Junctions and Capacitors in Nanoscale Dimensions

Ferroelectric tunneling junctions or capacitors (Fig. 5.1) using ferroelectric thin films as barriers on the nanoscale show many interesting electronic transport properties, such as giant electro-resistance (GER), giant piezo-resistance (GPR) and non-destructive read out (NDRO) (Zhuravlev et al., 2005; Velev et al., 2007; Pertsev and Kohlstedt, 2007; Contreras et al., 2003; Duan et al., 2006a; Tsymbal and Kohlstedt, 2006; Kohlstedt et al., 2005; Zheng and Woo, 2009b; Luo et al., 2010; Choi et al., 2009). Many of them present attractive prospects for applications in the functional design of nano-transducers, ultra high-density nonvolatile random-access memories, ferroelectric nano-diodes and nano-photovoltaic devices, etc.

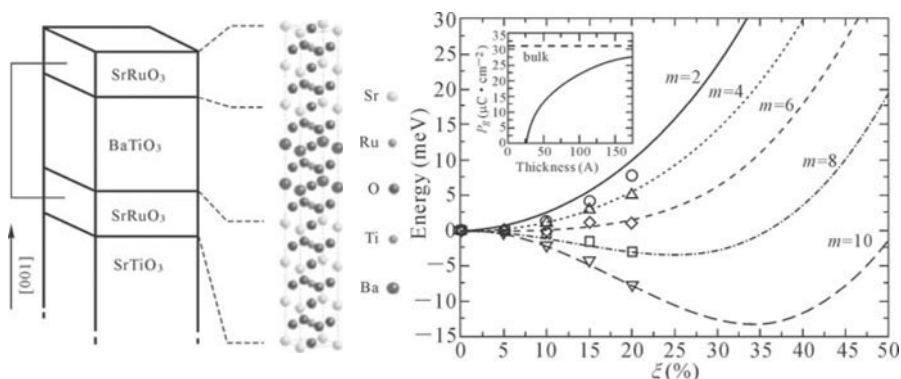


Fig. 5.1 Structure of a typical ferroelectric capacitor or junctions. (a) Schematic view of the system considered here: a BaTiO₃ film between SrRuO₃ electrodes that are short-circuited. The whole structure is grown on top of a SrTiO₃ substrate, atomic representation of the related supercell that is simulated; (b) First principles results (symbols) and electrostatic model results (lines) are shown for different thicknesses of the ferroelectric thin film. Inset is evolution of the spontaneous polarization with thickness (Junquera and Ghosez, 2003). Copyright(2003) by the Nature Publishing Group

Generally speaking, the most important and useful characteristic of a ferroelectric tunneling junction or capacitor is the order of magnitude change in the tunneling conductance it may display in response to moderate changes of the polarization in the tunnel barrier. Naturally, the stability of the ferroelectric state in the barrier is crucial to all these wonderful possibilities. It is well known that the polarization in a ferroelectric tunnel junction or capacitor can be retained only above some critical thickness of about 1~10 nm (Kim et al., 2005). A good understanding of the factors governing the critical thickness might thus help greatly in the full utilization of ferroelectric tunnel junctions or capacitors in nano-devices.

For example, the first-principle calculations of Junquera and Ghosez (2003) predicted a critical size of about 2.4 nm for ferroelectric thin films with symmetrical electrodes such as SrRuO₃/BaTiO₃/SrRuO₃. Umeno et al. (2006) investigated the stability of the ferroelectric state of ultrathin Pt/PbTiO₃/Pt based on *ab initio* DFT-LDA calculation, and found that the critical thickness was about 1.6 nm and 2.4 nm for PbO and TiO₂-terminated films, respectively. Note that ferroelectrics models based on the first principle calculations with as little as a thousand atoms are already very difficult to perform due to the large demand on computer resources. Studies of anything other than ultra-thin perfect samples are next to impossible. It is well known that no reliable first principle calculation scheme has yet been developed which is directly applicable in the thermodynamic regime near phase transition, because of the heavy involvement of many-body excited states such as phonons, magnons, polarons and excitons, which lie far beyond what a density function theory description can offer. Nevertheless, the thermodynamic approach may then serve as a conceptual bridge between microscopic and macroscopic description in such cases.

Indeed, Duan et al. (2006b) suggested that information on some parameters in the thermodynamic model, such as the extrapolation lengths, can be obtained from the first principle calculations (Fig. 5.2). Similarly, Tagantsev et al. investigated the size, surface effect and interface problems in metal-ferroelectric-metal heterostructures by using a multi-scale scheme, in which models based on thermodynamics and first-principle calculations were combined, such as coefficients about surfaces and interfaces obtained using the first principle calculations (Gerra et al., 2007; Tagantsev et al., 2008). They also tested the scheme by evaluating the transition temperature as the temperature at which the paraelectric phase became unstable, along the lines of the stability analysis (Wang and Woo, 2005).

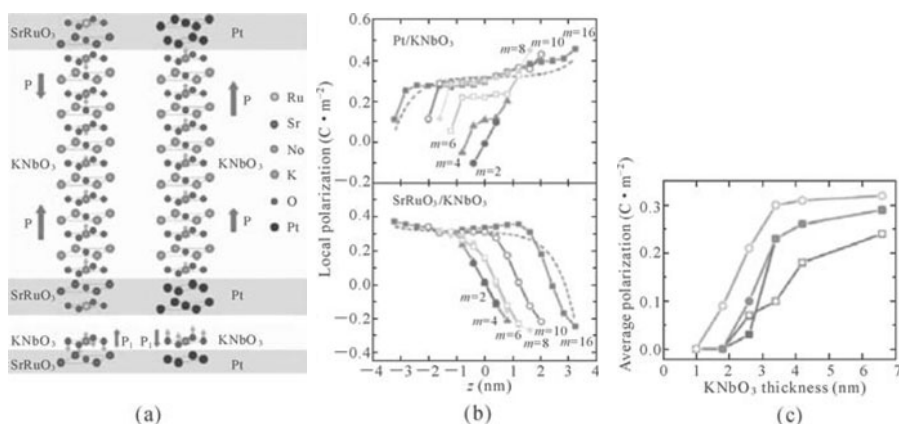


Fig. 5.2 (a) Atomic structures of SrRuO₃/KNbO₃ (left) and Pt/KNbO₃ (right) superlattices; (b) Local polarization in a ferroelectric KNbO₃ layer; (c) Average polarization as a function of KNbO₃ film thickness for Pt electrodes (circles) and SrRuO₃ electrodes (squares). Solid symbols represent results of the calculation based on soft mode displacements; open symbols are results obtained from fully relaxed structures using local polarizations (Duan et al., 2006b). Copyright(2006) by the American Chemical Society

5.1.3 Ferroelectric Multilayers in Nanoscale

Artificially structured ferroelectrics are experiencing a rapid surge of research interest, driven by the demand of new functional materials that provide new opportunities for novel applications. Studies of ferroelectric bilayers, multilayers and superlattices, such as KNbO₃/KTaO₃, PbTiO₃/SrTiO₃, BaTiO₃/SrTiO₃, PbTiO₃/PbZrO₃, PbTiO₃/BaTiO₃, etc., have revealed that the properties of ferroelectric multilayers can be controlled by adjusting structural periodicity, misfit strain, thickness ratio, surface boundary conditions, electrostatic interactions etc.

For example, using ultraviolet Raman spectroscopy, Tenne et al. (2006) found

that the transition temperature of ferroelectric ultrathin $\text{BaTiO}_3/\text{SrTiO}_3$ superlattices grown on SrTiO_3 could change by up to 500 K, depending on the thickness ratio of BaTiO_3 and SrTiO_3 layers (Fig. 5.3). This illustrates the important role of electrical and mechanical boundary conditions on critical properties of a ferroelectric superlattice. Dawber et al. (2007) studied $\text{PbTiO}_3/\text{SrTiO}_3$ superlattices grown on SrTiO_3 by X-ray diffraction and piezoelectric atomic force microscopy and found that the polarization and phase transition temperature were functions of the thicknesses of PbTiO_3 and SrTiO_3 . Using both experiments and first-principle calculations, they also found that the ferroelectricity of a superlattice would disappear when the thickness or ratio of the PbTiO_3 layer was below some critical values. In addition, first-principle calculations, atomistic simulations and thermodynamic models have contributed significantly to understanding the effects of artificial structures on the properties of ferroelectrics (Qu et al., 1997; Shen and Ma, 2000; Li et al., 2007a; 2007b; Tenne et al., 2006; Ahn et al., 2004; Dawber et al., 2005a; 2007).

5.1.4 Ferroelectric Nanowires and Nanotubes

The proximity of the surface is also known to affect the polarization distribution of a ferroelectric sample. Indeed, since the polarization depends on the eigenstrain (transformation strain) associated with the structural instability of the crystal lattice, it naturally varies according to the atomistic environment of the location under consideration. The large disruption in crystalline uniformity near the surface is expected to affect the eigenstrain and hence the polarization. This effect of surface relaxation on the spontaneous polarization has been modeled using an extrapolating length in many theoretical investigations of the ferroelectric properties of the ferroelectric thin films, nanoparticles, nanowires and nanotubes, etc. (Ma et al., 1998; Uchino et al., 1989; Morozovska et al., 2006; Zheng et al., 2007; 2008; Liu et al., 2006; Alexe et al., 2006; Luo et al., 2003; Zhou et al., 2007; Spanier et al., 2006; Jang et al., 2005; Scott, 2007) (Fig. 5.4 and Fig. 5.5).

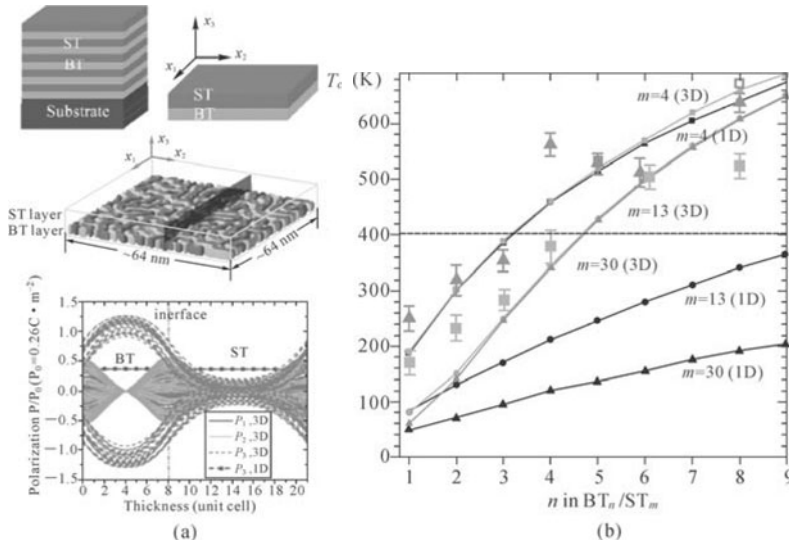


Fig. 5.3 (a) Schematic structures of an epitaxial BT_n/ST_m superlattice on a substrate and the simulation cell, where BT, ST are BaTiO₃ and SrTiO₃, respectively; Domain morphology of a BT/ST superlattice and projections of the polarization components. (b) Dependence of the phase transition temperature T_c on superlattices BT_n/ST_m , where open triangles are from temperature-dependent XRD measurements, circles with lines are from the 3D phase-field model calculations (Li et al., 2007b). Copyright(2007) by the American Institute of Physics

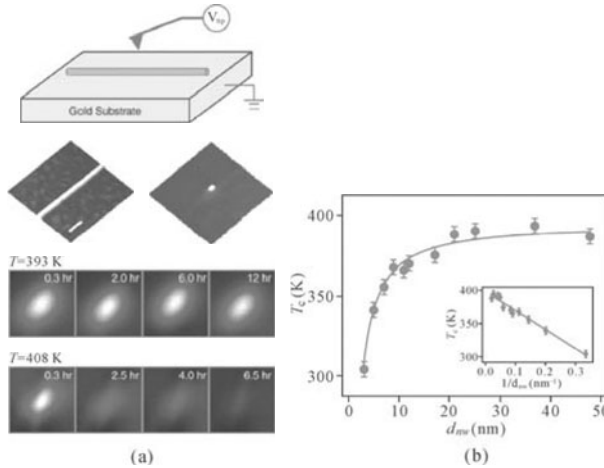


Fig. 5.4 (a) Schematic diagram illustrating the experimental geometry. Here V_{tip} represents the voltage applied to a conductive tip; a topographic image (left) and a spatial map of the electric polarization (right) of an 11-nm-diameter nanowire; time series of EFM images for a 25-nm-diameter nanowire following the polarization writing below and above the phase transition temperature. (b) Ferroelectric phase transition temperature T_c as a function of d_{nw} , where the solid circles are the experimentally determined T_c , and the error bars are uncertainties, the magenta solid line is the result of the fit to the data using the $1/d_{nw}$ scaling relation, the inset plots T_c as a function of $1/d_{nw}$ and illustrates the inverse diameter dependence (Spanier et al., 2006). Copyright(2006) by the American Chemical Society

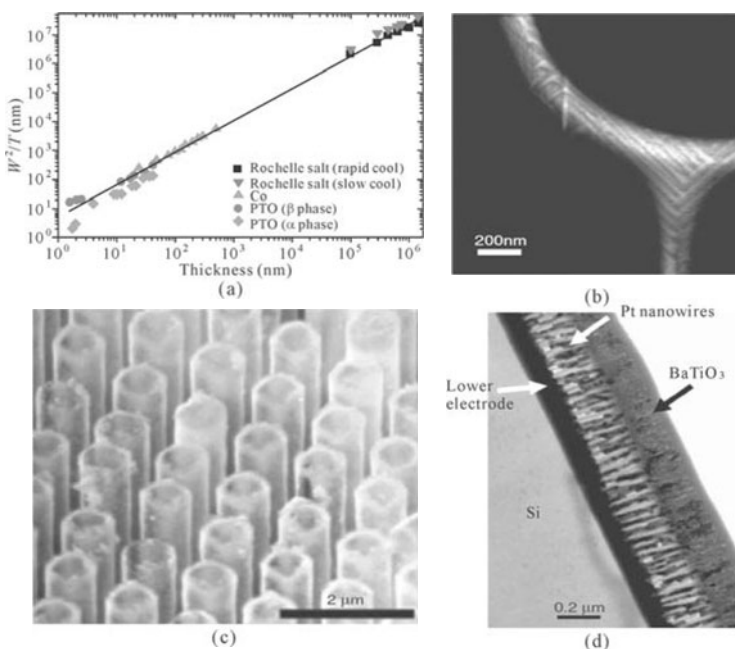


Fig. 5.5 (a) Stripe width squared divided by domain wall thickness versus film thickness; (b) Plan view micrograph of 3D nanodomains in BaTiO_3 ; (c) Scanning electron microscopy (SEM) micrograph of $\text{SrBi}_2\text{Ta}_2\text{O}_9$ nanotubes at oblique incidence. (d) SEM cross-sectional view of 0.5 Tb/in^2 nanoarray of Pt nanowires and PZT (Scott, 2007). Copyright (2007) by the Nature Publishing Group

Based on the thermodynamic theory, size and surface effects on the Curie temperature and polarization of a spherical nanoparticle have been investigated (Zhong et al., 1994). They also found a diminishing polarization and Curie temperature as the extrapolation length decreased. Using first-principle calculations, the polarization and Curie temperature of one-dimensional ferroelectric nanowires have been investigated, the results of which showed that the polarization and Curie temperature was size-dependent, and that ferroelectricity disappeared completely when the radius was reduced below a critical value (Naumov et al., 2005).

More importantly, for a nanowire or nanotube, the large surface to volume ratio naturally leads to the other strong surface effect, i.e., surface tension. Indeed, the tendency to minimize the surface energy introduces a large intrinsic surface stress in samples of nano-dimension, and is one of the most important surface effects considered. The role of surface tension may be seen from the significant polarization enhancement of $\text{Pb}_{0.25}\text{Ba}_{0.15}\text{Sr}_{0.6}\text{TiO}_3$ in nanotube form compared with that in thin-film form (Liu et al., 2006). In another case, a giant reversible piezoelectric strain in a 70 nm diameter $\text{PbZr}_{0.2}\text{Ti}_{0.8}\text{O}_3$ nanowire was observed in the presence of a large compressive stress due to the surface tension (Zhou et al., 2007). In ultra-small $\text{PbZr}_{0.52}\text{Ti}_{0.48}\text{O}_3$ nanowires and nanotubes with radii R

between 20–30 nm and length 50 m, a rectangular shaped piezoelectric hysteresis loop with an effective remnant piezoelectric coefficient comparable with that typically found in PZT films. By using X-ray diffraction (XRD), Uchino et al. (1989) investigated the effect of the surface tension effect on the Curie temperature of BaTiO₃ particle. The abrupt change in the Curie temperature with the observed particle size was found to corroborate the large compressive pressure induced by the surface-tension.

Both effects, i.e., the surface extrapolation length and surface tension, alter the free energy of a bulk ferroelectric, causing the phase transition temperature to shift and polarization to change. In recent works, using a thermodynamic model and linear stability analysis of the associated dynamic Ginzburg-Landau equation, analytic expressions of the transition temperature and Curie-Weiss relation of ferroelectric nanowires and nanotubes were derived and analyzed, paying particular attention to the effects of the surface tension and surface relaxation (Glinchuk et al., 2003; Morozovska et al., 2006). The transition temperature, polarization and Curie-Weiss relation were considered as a function of wire and tube dimensions through the actions of the surface tension, polarization gradient coefficient, extrapolation length and electrostriction coefficient (Zheng et al., 2008a; 2008b).

5.1.5 *Ferroelectric Nanograins or Nanoislands on Substrates*

Although the origin of the polarization instability was not clearly addressed therein, it was pointed out that the ferroelectric size effect should be three-dimensional in nature. The divergence in the theoretical predictions arises from the different electrical and mechanical boundary conditions considered in the computations, for example, screening length of the electrodes, strain conditions, and degree of surface relaxation. The disagreement between experimental results and theoretical predictions can be associated with the simplification of the electromechanical boundary conditions involved in calculations by neglecting defects frequently formed at the interfaces. To bridge these experimental and theoretical gaps it is essential to probe the role of interfacial defects in nanoscale ferroelectrics with a tetragonal perovskite structure.

Roelofs et al. (2002) fabricated ferroelectric lead titanate nanoislands down to 15 nm on Si/SiO₂/TiO₂/Pt substrate, and they found that the limiting size for the ferroelectric phase was about 20 nm. In a more recent paper, Chu et al. (2004) investigated the effect of lattice dislocations at the interface between the ferroelectric nanometer-sized island and electrodes on the size limit, and they concluded that controlling such defects to keep the epitaxial stress would be crucial to pushing the storage density of ferroelectric memories to such a limit (Fig. 5.6). In fact, the misfit dislocations play the role of releasing the misfit stresses. Using quantitative high resolution electron microscopy, they visualized the strain fields of edge type misfit dislocations, extending predominantly into a PZT region with a height of ~4 nm and width of ~8 nm. The lattice within this region deviates from the regular crystal structure. Piezoresponse force microscopy

indicates that such PZT nanoislands do not show ferroelectricity (Scott, 2007). Wang and Woo (2005) suggested that misfit engineering was indispensable for obtaining nanostructured ferroelectrics with stable polarization. From the stability analysis, an analytic expression of critical properties of ferroelectric nanograins or nanoislands can be established considering size and surface effects (Wang et al., 2010).

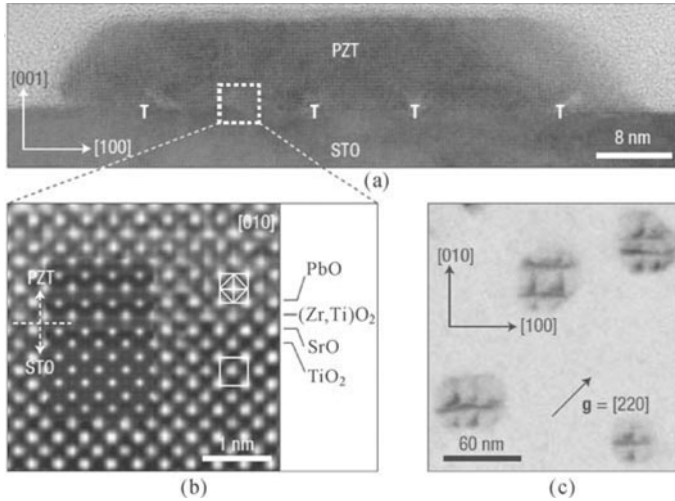


Fig. 5.6 Cross-sectional and plan-view micrographs and calculated contrasts of the PZT nanoislands. (a) Cross-sectional HREM image of a (001)-oriented PZT nanoisland in [010] projection, with misfit dislocations indicated by *T*. (b) An enlarged interface zone of (a). The inserted small white rectangle and square indicate respective PZT and STO lattices; inset (dark area on the left), multislice contrast simulation at $t = 4$ nm and $\Delta f = -60$ nm with PbO-(Zr,Ti)O₂-SrO-TiO₂ stacking across the interface. At this imaging condition, bright and dark contrasts represent the cation and anion columns, respectively. (c) Bright-field plan view image recorded under $g = [220]$ on specimens annealed at 950 °C for 1 h, the dark contrasts showing the network of misfit dislocations

5.2 Thermodynamic Modeling and Stability Analysis of Ferroelectric Systems

5.2.1 Background of the Thermodynamic Modeling for Ferroelectrics

Theoretically, the thermodynamic models, first principle calculations and atomistic simulations have contributed significantly to understanding properties of nanoscale ferroelectric systems. By taking into account symmetric conditions, Landau's theory can provide a reliable and reasonable description of a system's equilibrium behavior near a phase transition. Thermodynamic models of ferroelectrics formulated based on the Landau free-energy expansion are popular and suitable for studying the properties of the ferroelectrics (Landau et al., 1960;

Lines and Glass, 1979). This phenomenological approach is also a conceptual bridge between microscopic models and macroscopic phenomena of materials. The simplicity and easy experimental compatibility of the model are among the attractions. Indeed, it is the general validity of thermodynamics from which the strength of the model was derived. For example, Landau's symmetry-based treatment of phase transitions was applied to the case of ferroelectrics by Devonshire and ferroelectrics in finite size by Ginzburg (Gonzalo and Jimenez, 2005; Rabe et al., 2007; Haun et al., 1987; Li et al., 1997; Hu et al., 2003). Then many theoretical models based on the Landau theory (i.e., Landau-Devonshire, Ginzburg-Landau, Landau-Ginzburg-Devonshire, the time dependent Ginzburg-Landau equations, phase field model, etc.) have been widely used for investigating properties of ferroelectrics in different dimension scales. Despite inaccuracies due to the neglect of stochastic fluctuations of a thermodynamic system, these approaches have been proven powerful and effective, and widely used for investigating properties of ferroelectrics.

Recent modeling efforts of nanoscale ferroelectrics using the Landau approach have also faced challenges in the formulation when contributions of the surface effects, inhomogeneous polarization distribution and depolarization field have to be incorporated in the derivation of the free-energy functional. In this regard, it is of paramount importance that the thermodynamics, mechanics and electromagnetics involved are considered with full self-consistency.

When surface effects from polarization gradients, near-surface effect and the depolarization field E_d are added, the accuracy of the Landau approach suffers, particularly when E_d is significant. It is known that the depolarization field is mainly determined by the surface polarization charge of ferroelectric with the open-circuit boundary conditions. For the short-circuit boundary conditions, in general, the depolarization field has two components, one coming from the incomplete charge compensation in the electrodes and the other from the inhomogeneity of the polarization field near the interface. Here the first component can be formulated in terms of the screening length in the electrodes, and the second component is more complex and is determined by several factors, including space and surface charges, the small size effect, interface effect or Schottky barriers, etc. Indeed, misunderstandings in this regard have led to controversies concerning the contribution from the depolarization field, which results in large discrepancies in the outcome of different calculations.

Under open-circuit boundary conditions, very often the depolarization field calculation based on the usual expression $E_d = -P^T/\epsilon_0$ completely overwhelms the total free energy of a typical ferroelectric thin film of PbTiO_3 or BaTiO_3 . Here P^T and ϵ_0 are the total polarization and permittivity of the vacuum, respectively. Indeed, the depolarization field in such cases reaches strengths of $\sim 10^5 \text{ kV}\cdot\text{cm}^{-1}$, which are two orders of magnitude larger than the sum of all other contributions to the total free energy. The consequences include a downward shift in the phase transition temperatures of an unphysical amount of $\sim 10^4 \text{ K}$, or an 180° domain shrinkage of over 100% at very low temperature. Even under short-circuit boundary conditions (i.e., $E_d = -(P^T - q)/\epsilon_0$, here q is the induced surface charge

density in the electrodes), this depolarization contribution to the total free energy is also about two orders of magnitude larger than the sum of all the other contributions (Mehta et al., 1973; Kretschmer and Binder, 1979; Woo and Zheng, 2008; Zheng and Woo, 2009a). This is true even when the deviation from uniformity $P^I(z) - \langle P^I \rangle$ is only 1% of the polarization for bulk BaTiO₃ at room temperature ($\sim 0.25 \text{ C}\cdot\text{m}^{-2}$). Here $\langle P^I \rangle$ is the average total polarization. Due to the dominance of this term, a minimized total free energy cannot afford a relative deviation of P^I stronger than 10^{-4} from a uniform distribution, even in the presence of lattice defects, such as dislocation, cracks and surfaces. In order to minimize the non-uniformity of the polarization distribution under boundary conditions imposed by the presence of defects with significant surface and interface effects, the action of the Ginzberg term (i.e., gradient term) would force the polarization field to deviate significantly from zero. In this way, the effect of the depolarization would be so overpowering that the permanent polarization that determines its ferroelectric character would be practically eliminated, buried under its dielectric properties. The dominance of the depolarization field could lead to the quenching of domain patterns and other properties in the presence of strain fields created by interfacial dislocations, point defects, composition gradient, bilayers, multilayers or superlattices, etc. As a result, no successful application of the Landau approach to the study of multilayered structures has been carried out so far. Indeed, without a good understanding of the source of the problem a satisfactory method of resolution, any study in this direction cannot be very meaningful. Importantly, Tagantsev (2008) also discussed “The Landau expansion for ferroelectrics: which variable to use?”, and generally the problem of the choice of the order parameter in the Landau expansion of proper ferroelectrics.

Based on the premise that phase transition from a stable state of a dynamic system occurred only when the state becomes unstable, irrespective of the order of the transition, a thermodynamic model has been developed to describe critical parameters in ferroelectric transformation in thin films with approximately considering effects of the background dielectric material (Zheng et al., 2009). Instead of adopting a numerical approach, the critical values of the parameters, such as the Curie temperature and critical thickness, were calculated via a linear stability analysis of the evolution equation of the system. Subsequently, recent investigations comprehensively showed that the effect of the depolarization field determined by the spontaneous polarization and background dielectric constant was very important to critical properties of ferroelectric thin film. These works have investigated the source of the foregoing difficulties and suggested that the problem was the result of a serious underestimation of the dielectric screening of the electrostatic interaction between the spontaneous polarization and the depolarization field E_d , and that the simple solution was to use the spontaneous polarization as the order parameter. This suggestion has also been adopted in several recent works.

Indeed, although the order parameter P^I used in the Landau free-energy expansion is supposed to represent the total polarization, the approach may be

called the Total-Polarization-Order-Parameter (TPOP) approach. In reality, the order parameter at this stage is only an unknown of the Euler-Lagrange (E-L) equation derived from the free-energy functional through the variational principle. The dynamic characteristic of P^T , such as its instability that marks the phase transition, is totally determined by the form and the parameters of the E-L equation. Thus, in order for the polarization effect of the depolarization field to enter into the dynamic character of the E-L equation, it must be included explicitly in the free-energy functional, and cannot be done implicitly through the order parameter. In the TPOP approach, the polarization induced by an applied field is only implicitly included in the free-energy functional. Its contribution is thus not recognized and, independent of what the assigned meaning of P^T is supposed to be, its effect will not be present in the dynamical behavior of its solution P^T . As a result, there is a lack of dielectric screening in the effect of the depolarization field.

Using spontaneous polarization as the order parameter in the Landau free-energy expansion and including the contribution from the induced polarization explicitly in the free-energy functional, recent works have shown that their effects enter the associated E-L equation properly, and the difficulties encountered in the TPOP approach can be avoided. For an easy description, and to highlight the substantial difference between the results of the two approaches, particularly in regard to the critical properties, this approach is named the Spontaneous-Polarization-Order-Parameter (SPOP) approach.

5.2.2 Electrostatics for Ferroelectrics

In the SPOP approach, the spontaneous polarization field $\mathbf{P}$ (simply referred to as the polarization in the following) is used as the order parameter. As mentioned, $\mathbf{P}$ is the polarization from the permanent electric moment formed from spontaneous atomic displacements generated in a dielectric material when going through a ferroelectric phase transformation. There is also another component of the polarization $\mathbf{P}^E$, which is the result of mechanisms such as the electronic polarization, and other *non-permanent* displacement of the ionic charge distribution, induced by an applied electric field $\mathbf{E}$, and which accounts for the dielectric screening of the effect of $\mathbf{E}$.

This description is consistent with Lines and Glass (1979) who observed that the total dielectric susceptibility χ^T of ferroelectrics could be expressed in general as the sum of linear and non-linear parts:

$$\begin{aligned} \chi^T &= \chi^0 + \left(\frac{\partial \chi^T}{\partial \mathbf{E}} \right)_{\sigma, T} \mathbf{E} + \left(\frac{\partial \chi^T}{\partial \boldsymbol{\sigma}} \right)_{\mathbf{E}, T} \boldsymbol{\sigma} + \left(\frac{\partial \chi^T}{\partial T} \right)_{\mathbf{E}, \boldsymbol{\sigma}} T + \left\{ \begin{array}{l} \text{higher-order} \\ \text{non-linearities} \end{array} \right\} \\ &\equiv \chi^0 + \chi(T, \mathbf{E}, \boldsymbol{\sigma}), \end{aligned} \quad (5.2.1)$$

according to which the dependence of the susceptibility, Curie-Weiss relation, etc. on the temperature T , stress σ and electric field $\mathbf{E}$ were described by the non-linear component χ . The total polarization $\mathbf{P}^T$ of ferroelectrics in this description is a sum of two components, one due to a permanent electric moment, which can be identified with $\mathbf{P}$ and whose relation to T , σ and $\mathbf{E}$ can be expressed via χ . This is embedded in a background material with dielectric response to $\mathbf{E}$, which is explicitly describable by a susceptibility χ^0 independent of $\mathbf{P}$, and which creates an induced polarization $\mathbf{P}^E$ proportional to $\mathbf{E}$. The sum of $\mathbf{P}^E$ and $\mathbf{P}$ constitutes the total polarization $\mathbf{P}^T$. To a good approximation, one may consider χ^0 as the dielectric susceptibility of the ferroelectrics in the paraelectric state far away from the phase transition (e.g., $\frac{\partial \chi^T}{\partial T} = \frac{\partial \chi}{\partial T} = 0$ at $T \gg T_c$ for $\mathbf{E} = \sigma = \mathbf{0}$, and $\chi^T = \chi^0$ for $T \gg T_c$). It should thus be accurately measured at $T \gg T_c$ for $\mathbf{E} = \sigma = \mathbf{0}$. Here T_c is the phase transition temperature. To be consistent with the notations, χ^0 with the background susceptibility χ_b or χ_d was used in previous works (Li et al., 2002; Wang and Woo, 2005; 2006; Lupascu, 2004; Zheng and Woo, 2009a). Note that the response of $\mathbf{P}$ to an applied electric field $\mathbf{E}$, and for that matter the non-linear component of the total susceptibility, is implicit in this formulation, and is realizable only through the thermodynamic relation between $\mathbf{P}^T$ and $\mathbf{E}$ via the total free energy functional, after the associated E-L equation is solved.

At constant applied stress and temperature, the electric displacement field $\mathbf{D}$ can thus be expressed in terms of the linear-part induced polarization and the nonlinear-part spontaneous polarization as

$$\mathbf{D} = \varepsilon_0 \mathbf{E} + \mathbf{P}^T = \varepsilon_0 \mathbf{E} + \mathbf{P}^E + \mathbf{P} = \varepsilon_0 \mathbf{E} + \chi_b \mathbf{E} + \mathbf{P} = \varepsilon_b \mathbf{E} + \mathbf{P}, \quad (5.2.2)$$

where ε_b is the background dielectric constant. Indeed, in the absence of an external electric field, the total electric field has only one contribution that is the depolarization field $\mathbf{E}_d$, i.e., $\mathbf{E} = \mathbf{E}_d$. The electric displacement is given by

$$\mathbf{D} = \varepsilon_0 \mathbf{E}_d + \mathbf{P}^T = \varepsilon_0 \mathbf{E}_d + \mathbf{P}^E + \mathbf{P} = \varepsilon_0 \mathbf{E}_d + \chi_b \mathbf{E}_d + \mathbf{P} = \varepsilon_b \mathbf{E}_d + \mathbf{P}. \quad (5.2.3)$$

5.2.3 Thermodynamics of Ferroelectrics

In Eqs. (5.2.2) and (5.2.3), the induced component of the total polarization is explicitly taken into account. A crystal is considered as the thermodynamic reference, which is of infinite extent (surfaceless) under the action of any applied field, i.e., electric, magnetic, or mechanical (internal or external). The general free-energy functional of ferroelectrics can be written as

$$g(\mathbf{P}, \mathbf{E}, \boldsymbol{\sigma}) = g(\mathbf{P}, \mathbf{E}, \boldsymbol{\sigma}) \Big|_{\mathbf{E}=0, \boldsymbol{\sigma}=0} - \int_0^{\boldsymbol{\sigma}} \mathbf{u} d\boldsymbol{\sigma} - \int_0^{\mathbf{E}} \mathbf{D} d\mathbf{E} = g_0 - \int_0^{\boldsymbol{\sigma}} \mathbf{u} d\boldsymbol{\sigma} - \int_0^{\mathbf{E}} \mathbf{D} d\mathbf{E}, \quad (5.2.4)$$

where $\mathbf{u}$ and $\boldsymbol{\sigma}$ are the strain and stress tensor fields, respectively. Here one should bear in mind that g_0 is the free-energy functional of the field-free, uniform and infinite crystal (surfaceless and thus with no depolarization charges) used as reference.

In the following discussions, we concentrate only on considering the Gibbs free energy density (including that of the applied fields) in an unstressed sample at constant temperature. If the electric field in the ferroelectrics is not zero, the thermodynamic potential can be obtained by starting from the relation $\frac{dg}{d\mathbf{E}} = -\mathbf{D}$, where $\mathbf{D}$ is given by Eq. (5.2.2). In this case, it follows from Eq. (5.2.4) and the fact that $\mathbf{P}$ has no explicit functional dependence on $\mathbf{E}$ that,

$$g(\mathbf{P}, \mathbf{E}) = g(\mathbf{P}, \mathbf{E}) \Big|_{\mathbf{E}=0} - \mathbf{P} \cdot \mathbf{E} - \int_0^{\mathbf{E}} \boldsymbol{\varepsilon}_b \cdot \mathbf{E} d\mathbf{E} = g_0 - \mathbf{P} \cdot \mathbf{E} - \frac{1}{2} \boldsymbol{\varepsilon}_b \mathbf{E}^2. \quad (5.2.5)$$

Here Eq. (5.2.5) is similar to that derived by Landau et al. (1960) with consideration of the linear part of the polarization.

As the free-energy functional of the field-free (i.e., $\mathbf{E} = 0$), uniform and infinite (surfaceless) crystal, g_0 can be expanded near the phase-transition instability in terms of $\mathbf{P}$ in the form of the Landau free-energy functional. Then, for the case in which $\mathbf{E}$ and $\mathbf{P}$ are parallel, and are both spatially uniform, substituting $E = E_{\text{ext}} + E_d$ in Eq. (5.2.5) leads to

$$g(T, E) = \frac{A}{2}(T - T_{c0})P^2 + \frac{B}{4}P^4 + \frac{C}{6}P^6 - P(E_{\text{ext}} + E_d) - \frac{1}{2}\boldsymbol{\varepsilon}_b(E_{\text{ext}} + E_d)^2, \quad (5.2.6)$$

where $A = 1/(\varepsilon_0 C_0)$, C_0 being the Curie-Weiss constant and ε_0 the permittivity of the vacuum. T_{c0} is the Curie temperature and B and C the expansion coefficients of the free energy of the reference crystal. It can be seen that the corresponding E-L equation will give the correct phase transition characteristics for $E = 0$.

Instead of Eq. (5.2.6), a number of authors prefer to express the free energy functional and depolarization field in terms of the total polarization and adopt the TPOP approach, in which the free energy functional is written as

$$g(T, E) = \frac{A}{2}(T - T_{c0})P^{\text{T}2} + \frac{B}{4}P^{\text{T}4} + \frac{C}{6}P^{\text{T}6} - P^{\text{T}}E_{\text{ext}} - \frac{1}{2}P^{\text{T}}E_d, \quad (5.2.7)$$

where the effect of the extrapolation length has been neglected. It can be seen immediately that both the free energies of the reference state and the electrostatic contributions in Eqs. (5.2.6) and (5.2.7) are different. Although the free energy

expressed by Eq. (5.2.7) has the correct form in the limit of $E \rightarrow 0$, yet Eq. (5.2.7) no longer satisfies $\frac{\partial g}{\partial E} = -D$, as it should. As mentioned, in order for the effect of the applied field E to be fully reflected, it should be explicitly excluded from the order parameter when deriving the E-L equation from the free-energy functional via the variational principle. Indeed, it can be clearly seen that the dielectric screening effect due to P^E is totally missing from the E-L equation derived from the free-energy expression in Eq. (5.2.7). Furthermore, ignoring the interrelation between E_d and P^T could also produce errors in evaluating contributions from the internal energy of the field $\int D(E)dE$ in Eq. (5.2.7), and from missing the interaction term with E_{ext} in the last term of Eq. (5.2.7).

More importantly, according to Eq. (5.2.6), the application of E produces an indirect effect on P when the variational minimum of the free-energy functional is sought. The total susceptibility χ^T is then a sum of the linear part χ_b and non-linear part χ , which measures this indirect effect of E on P , and can be obtained by solving the corresponding E-L equation. Accordingly, the dielectric properties of ferroelectrics are governed by a Curie-Weiss-type relation as

$$\chi^T = \chi + \chi_b = \begin{cases} \frac{\varepsilon_0 C_0}{2|T - T_c|} + \chi_b, & \text{for } T < T_c, \\ \frac{\varepsilon_0 C_0}{|T - T_c|} + \chi_b, & \text{for } T > T_c. \end{cases} \quad (5.2.8)$$

It also gives the correct limit of the total susceptibility far away from its phase transition temperature T_c , i.e., $\chi^T \Big|_{T \rightarrow \infty} = \left[\frac{\varepsilon_0 C_0}{T - T_c} + \chi_b \right] \Big|_{T \rightarrow \infty} = \chi_b$, which should be experimentally determined (Pintilie et al., 2007; 2008; Morozovska et al., 2008; Eliseev et al., 2009; Eliseev and Morozovska, 2009). For example, the background susceptibility χ_b or dielectric constants ε_b of PbTiO_3 , BaTiO_3 , SrTiO_3 , CaTiO_3 are in the order of $50\varepsilon_0$ in the limit of infinite temperature (Charles, 2005; Rupprecht and Bell, 1964; Li et al., 2002; Woo and Zheng, 2008; Zheng and Woo, 2009a; Zheng et al., 2009). The corresponding relation obtained using the TPOP approach of Eq. (5.2.7), on the other hand, means the limit of the total susceptibility tends to zero, i.e., $\chi^T \Big|_{T \rightarrow \infty} = \frac{\varepsilon_0 C_0}{T - T_c} \Big|_{T \rightarrow \infty} = 0$.

5.2.4 *Stability Analysis of Critical Properties of Ferroelectric Systems*

It is self-evident that phase transitions from a stable state of a dynamical system cannot occur without the state becoming unstable, independent of the order and the direction of the transition. The evolution equation of the state of polarization in a ferroelectric nanostructure can be derived from its free energy, in terms of control parameters such as the ambient temperature, boundary conditions, surface characteristics, sample dimensions, and misfit epitaxial stresses. Then important characteristics of ferroelectrics in nanoscale, such as the transition temperatures, critical thicknesses, domain morphology, substrate nature, etc., are related through the instabilities of the initial state, which is accessible via a stability analysis, the linear nature of which opens the system to many established analytic techniques. This is the approach that had been adopted by Wang and Woo (2005). An analytic approach to establish conditions of the stability of the paraelectric versus ferroelectric phases in ferroelectrics nanostructures was obtained. The polarizing/depolarizing effects due to the presence of the ferroelectric surface, the restraint of the transformation strain, the electromechanical surface conditions involving the epitaxial stress, external mechanical loads and the induced surface charges, etc., are taken into account. Expressions for the para/ferroelectric transition temperatures as a function of size, and the corresponding critical dimensions, were derived, and the complex relation of the critical dimensions with the various parameters was summarized.

Furthermore, straight adherence to thermodynamic principles is emphasized, and the relationship between the polarization and depolarization field in the free-energy expansion is carefully considered. Using the SPOP in the Landau free-energy functional, the difficulties encountered in the TPOP approach can be avoided. Based on the SPOP approach and stability analysis, the phase transition temperature, polarization and susceptibility of ferroelectrics in nanoscale, with size and surface effects, are calculated and discussed in the following sections.

A thermodynamic model of ferroelectric film is derived in Section 5.3. Moreover, size and surface effects on critical properties of the ferroelectric ultrathin films (in Section 5.3), tunneling junctions and capacitors (in Section 5.4), superlattices (in Section 5.5), nanowires and nanotubes (in Section 5.6), nanograins and nanoislands (in Section 5.7) are discussed using the evolution equation and stability analysis of the stationary states.

5.3 Ferroelectric Thin Films in Nanoscale

5.3.1 *Thermodynamic Model for a Thick Ferroelectric Film*

Firstly, details of the reason why the TPOP approaches have to face challenges in

the formulation are given, when contributions of the surface effects, inhomogeneous polarization distribution and depolarization field have been considered. Let us consider the simple case of a thick ferroelectric film with parallel surfaces perpendicular to the z -direction, free standing in a vacuum in the absence of an electric field (i.e., $\mathbf{E}_{\text{ext}} = 0$). Due to translational symmetry in the x - and y -directions, all variables here depend only on z , with constant x - and y -components. Accordingly, only the z -component has to be solved and only in terms of a single variable z .

Thus, inside the plate in the dielectric, there are no free charges. Outside the film in the field-free vacuum, $D = 0$ and continuity of the normal component of the displacement field gives

$$D = P^T + \epsilon_0 E_d = 0, \text{ and } E_d = -P^T / \epsilon_0. \quad (5.3.1)$$

Using Eq. (5.2.3), the electric displacement may also be written as

$$D = P + \epsilon_b E_d = 0, \text{ and } E_d = -P / \epsilon_b. \quad (5.3.2)$$

If Eq. (5.2.7) is used, the free-energy density expression with TPOP is given by

$$g_{\text{TPOP}} = \frac{A}{2}(T - T_{c0})P^2 + \frac{B}{4}P^4 + \frac{C}{6}P^6 + \frac{P^2}{2\epsilon_0}. \quad (5.3.3)$$

Firstly, a single domain in ferroelectric film is considered, and the condition of dynamic instability gives the equation of the transition temperature $T_c = T_{c0} - 1/(\epsilon_0 A)$. More importantly, with a value of $8.85 \times 10^{-12} \text{ F} \cdot \text{m}^{-1}$ for ϵ_0 , typical systems like KH_2PO_4 , BaTiO_3 and PbTiO_3 , etc. have values of $1/(\epsilon_0 A)$ from $\sim 10^3 \text{ K}$ to $\sim 10^5 \text{ K}$. Using Eq. (5.2.7), the Curie temperature of these films would be far less than 0 K, and could never have existed, in particular if we consider the 180° domain appears due to effect of the large depolarization field in the film. This depolarization contribution in Eq. (5.2.7) is about three orders of magnitude larger than the rest of the contributions to the total free energy. When people minimize the total free energy, this large depolarization field will reduce the domain width to about three orders of magnitude smaller than the thickness of film at low temperature (i.e., $T \ll T_{c0}$). This result means that a very thick BaTiO_3 film must be in a paraelectric state, because the domain will shrink and disappear due to the effect of the depolarization field.

If the derived free-energy density with the SPOP approach in which Eq. (5.2.6) is used instead, then

$$g_{\text{SPOP}} = \frac{A}{2}(T - T_{c0})P^2 + \frac{B}{4}P^4 + \frac{C}{6}P^6 + \frac{P^2}{2\epsilon_b}. \quad (5.3.4)$$

The corresponding condition of dynamic instability gives, for the transition

temperature under the influence of the depolarization field, $T_c = T_{c0} - 1/(\epsilon_b A)$, yielding a downward transition temperature shift of $\Delta T_c = -1/(\epsilon_b A)$. Compared with the one calculated using the traditional expression Eq. (5.2.11), the temperature shift is reduced by a factor of ϵ_b/ϵ_0 for KH_2PO_4 , BaTiO_3 , PbTiO_3 , etc., and falls to much more reasonable values. Of course, the depolarization field can induce the 180° domains in the film with a reasonable width when the temperature is below the Curie temperature.

Another commonly studied system is the ferroelectric capacitor in which the ferroelectric film of thickness h is sandwiched between perfect metallic electrodes, which means that the incomplete charge compensation in the electrodes is neglected. In this case, inside the conducting electrode plates, $D = 0$, so that $\mathbf{D} \cdot \mathbf{n} = D = q_c$ on the ferroelectric boundary and inside. Here q_c is the induced surface charge density in the electrode plates. Using Eq. (5.2.2) and the SPOP approach, q_c can be written as $q_c/\epsilon_0 = D/\epsilon_0 = E(z) + P(z)/\epsilon_b$, from which the applied electric field is given by $E(z) = -[P(z) - q_c]/\epsilon_b$. When a potential gradient $\Delta\Phi/h$ is applied across the upper and lower surfaces, a relation between $q_c \neq 0$ and $\Delta\Phi/h$ can be found by integrating Eq. (5.3.2). Thus, for the short-circuit electrodes, $\Delta\Phi = 0$, and $E_d(z) = -[P(z) - \langle P \rangle]/\epsilon_b$ are obtained. Here $\langle P \rangle$ is the average spontaneous polarization.

The TPOP approach, such as the model derived by Kretschmer and Binder (1979) as cited in many works, has the expression of the free energy as

$$g_{\text{TPOP}} = \frac{A}{2}(T - T_{c0})P^{\text{T}2} + \frac{B}{4}P^{\text{T}4} + \frac{C}{6}P^{\text{T}6} + \frac{1}{2\epsilon_0}[P^{\text{T}2} - \langle P^{\text{T}} \rangle^2]. \quad (5.3.5)$$

This can be compared with the corresponding expression derived from Eq. (5.2.5),

$$g_{\text{SPOP}} = \frac{A}{2}(T - T_{c0})P^2 + \frac{B}{4}P^4 + \frac{C}{6}P^6 + \frac{1}{2\epsilon_b}[P^2 - \langle P \rangle^2]. \quad (5.3.6)$$

It can be seen that the depolarization contribution in Eq. (5.3.5) is about two orders of magnitude larger than the rest of the contributions to the total free energy, even when the deviation from uniformity $P^{\text{T}}(z) - \langle P^{\text{T}} \rangle$ is only 1% of the polarization for bulk BaTiO_3 at room temperature ($\sim 0.25 \text{ C} \cdot \text{m}^{-2}$). As a result of the dominance of this term, a minimized total free energy cannot afford a relative deviation of P^{T} stronger than 10^{-4} from a uniform distribution, even in the presence of lattice defects, such as dislocation, cracks and surfaces. Moreover, in order to minimize the non-uniformity of the polarization distribution under boundary conditions imposed by the presence of defects with significant surface effects, the action of the Ginzburg term would dictate that the polarization field cannot significantly deviate from zero. For this reason, the effect of the depolarization would be so much overestimated that the ferroelectric nature would

be largely eliminated. The dominance of the depolarization field would lead to quenching domain patterns or other properties induced by interface dislocations, point defects, compositionally graded bilayers, multilayers or superlattices, etc.

5.3.2 Size and Surface Effects on Ferroelectric Thin Films

A ferroelectric thin film of dimensions $\infty \times \infty \times h$, with perfect metal electrodes and short-circuit boundary conditions, is considered. Here h is the film thickness and below 100 nm. The origin of the coordinate system is at the center of the cell. The spontaneous polarization P is supposed to be orthogonal to the surface of the film. The polarizing effects due to the presence of the film surface, the restraint of the transformation strain, the electromechanical surface conditions involving the epitaxial stress and the induced surface charges, are taken into account.

Another important effect for ferroelectric thin film is the depolarization field that comes from the polarization-induced surface charges, which is governed by the electric boundary conditions. In this section, two cases are considered: (a) the ferroelectric film is sandwiched between two perfect metallic electrodes with short-circuit conditions, and (b) the ferroelectric film is sandwiched between two dielectric substrates. Although ideal screening in perfect electrodes is achieved under the short-circuit condition, partial compensation by the electrodes still reduces the depolarization field due to the surface effect of the ferroelectric thin film on the variation of the spontaneous polarization in the z direction. In this case, the relation between the depolarization field E_d and the spontaneous polarization P is given by

$$E_d = -\frac{1}{\epsilon_b} \left(P - \frac{1}{h} \int_{-h/2}^{h/2} P dz \right), \quad (5.3.7)$$

where the second term on the right hand side is due to the compensating charges on the upper and lower metallic electrodes. For dielectric substrates there are no compensation charges and the corresponding relation is given by

$$E_d = -\frac{P}{\epsilon_b}. \quad (5.3.8)$$

The depolarization field can also be written down as a general expression that encompasses both contact conditions,

$$E_d = -\frac{1}{\epsilon_b} \left(P - \frac{\phi}{h} \int_{-h/2}^{h/2} P dz \right), \quad (5.3.9)$$

where ϕ takes on the value of 0 or 1 depending on the contact condition as discussed in the foregoing.

For rigid substrates, the transformation strain $\varepsilon^T = \varepsilon_{xx}^T = \varepsilon_{yy}^T = QP^2$ is fully constrained, causing an elastic contribution to the total free energy density $F_e = GQ^2P^4$, where $G = (C_{11} + C_{12} - 2C_{12}^2/C_{11})$, C_{11} and C_{12} being components of the elastic modulus of the film. Due to the free upper surface of the films, the component ε_{zz}^T of the transformation does not contribute to the elastic energy. Including the effects of the depolarization field E_d , the epitaxial stresses $\sigma_{xx} = \sigma_{yy} = \sigma_r$ and upper ($z = h/2$) and lower ($z = -h/2$) surfaces, the free energy (per unit area) of the ferroelectric thin film can be written as

$$F = \int_{-h/2}^{h/2} \left\{ \frac{A}{2} (T - T_{c0}) P^2 + \frac{1}{2\varepsilon_b} P^2 + \frac{B + 4GQ^2}{4} P^4 + \frac{C}{6} P^6 + \frac{D}{2} \left(\frac{\partial P}{\partial z} \right)^2 - 2\sigma_r \varepsilon^T - \frac{1}{2\varepsilon_b} \langle P \rangle^2 \right\} dz + \frac{D}{2\delta} P_{z=h/2}^2 + \frac{D}{2\delta} P_{z=-h/2}^2, \quad (5.3.10)$$

where T_{c0} is defined as the cooling phase-transition temperature of the bulk crystal, δ is the so-called extrapolation length that measures the effect of lattice relaxation on the surface, and $\langle P \rangle$ is the average spontaneous polarization along the z direction, i.e., $\langle P \rangle = \frac{1}{h} \int_{-h/2}^{h/2} P dz$. Two separate cases may be considered: the more common $\delta > 0$ case, corresponding to a reduction of self-polarization on the surface, and the other $\delta < 0$ case, corresponding to an enhancement. It should be noted that in general δ is not an absolute constant.

5.3.3 The Evolution Equation and Stability of the Stationary States

In this section, an analytic approach is used to establish conditions of the stability of the paraelectric versus ferroelectric phases in a thin film grown on a compressive substrate. Using Eq. (5.3.10), the time evolution of the system is governed by the time-dependent Ginzburg-Landau (TDGL) equation,

$$\begin{aligned}
\frac{\partial P}{\partial t} &= -M \frac{\delta F}{\delta P} \\
&= M \left\{ -A(T - T_{c0})P + 4Q\sigma_r P - \frac{P}{\varepsilon_b} - (B + 4GQ^2)P^3 \right. \\
&\quad \left. - CP^5 + D \frac{\partial^2 P}{\partial z^2} + \frac{\phi}{\varepsilon_b h} \int_{-h/2}^{h/2} P dz \right\},
\end{aligned} \tag{5.3.11}$$

where M is the kinetic coefficient related to the domain wall mobility. The surface term in Eq. (5.3.10) yields the boundary conditions,

$$\frac{\partial P}{\partial z} = \mp \frac{P}{\delta}, \text{ for } z = \pm \frac{h}{2}. \tag{5.3.12}$$

Eq. (5.3.11) has a trivial solution of $P = 0$, representing the stationary paraelectric state. To be able to transform between the paraelectric ($P = 0$) and the ferroelectric ($P \neq 0$) states, the initial state must become unstable. The stability of the initial state can be probed by applying a small perturbation Δ to the corresponding stationary solution of Eq. (5.3.11). The dynamics of Δ follows from Eq. (5.3.11) by retaining only terms linear in Δ , which is given by,

$$\begin{aligned}
\frac{\partial \Delta}{\partial t} &= M \left\{ -A(T - T_{c0})\Delta + 4Q\sigma_r \Delta - \frac{\Delta}{\varepsilon_b} - 3(B + 4GQ^2)P^2 \Delta \right. \\
&\quad \left. - 5CP^4 \Delta + D \frac{\partial^2 \Delta}{\partial z^2} + \frac{\phi}{\varepsilon_b h} \int_{-h/2}^{h/2} \Delta dz \right\}.
\end{aligned} \tag{5.3.13}$$

Here the boundary condition is the same as Eq. (5.3.12), only with P replaced by Δ . For a symmetric configuration, the condition $\frac{d\Delta}{dz} = 0$, at $z = 0$ must also hold (Wang and Woo, 2005).

This equation governs the stability of the stationary solution of Eq. (5.3.11). It is a linear equation of Δ . Its analysis in the context of phase stabilities, without having to resort to a solution of the nonlinear Eqs (5.3.10) or (5.3.11), is the main advantage. It is obvious from Eq. (5.3.10) that the equation governing Δ in the paraelectric and ferroelectric states is different in general, from which it follows that the stability conditions for the two branches are also different. However, as we shall see, for second-order transitions, P is continuous at the transition point. The ferroelectric and paraelectric branches are both unstable when crossing this point from opposite directions, and transform into one another. The Curie temperature and the critical thickness are then independent of the initial state, and are thus well defined.

For first-order transitions, on the other hand, P is discontinuous across the

transition temperature. The paraelectric ($P = 0$) and the ferroelectric ($P \neq 0$) branches obey different stability conditions, due to the difference in the governing equations of Δ in Eq. (5.3.13). Suppose the paraelectric branch is stable/metastable for temperatures down to T_c , and the ferroelectric branch ($P \neq 0$) is stable/metastable for temperatures up to T_{ch} . Then the two transition temperatures are likely to be unequal. Indeed, two transition temperatures do exist, depending on the direction of the transformation between the two states. T_c is called the supercooling transition temperature and T_{ch} the Curie temperature (superheating transition temperature).

5.3.3.1 Stability of the Paraelectric State and Critical Characteristics

In view of its simplicity, the case in which the initial state is paraelectric, i.e., $P = 0$, is considered first. Transition from the ferroelectric state ($P \neq 0$) will be considered in Section 5.2.2.2, where it will be shown that the two transition temperatures are related by a constant shift. This section does not differentiate between the first- and second-order transitions, to which this analysis applies equally. In the following, the cases $\delta > 0$ and $\delta < 0$ are separately considered.

(1) Reduced surface polarization with the $\delta > 0$ case

In most ferroelectric materials, lattice relaxation weakens the polarization on the surface and $\delta > 0$.

By using the method of separation of variables, and taking into account $\frac{d\Delta}{dz} = 0$, at $z = 0$, Eq. (5.3.13) can be separated into time-dependent and time-independent parts, connected by the eigenvalue. For $P = 0$, the solution can be written as

$$\Delta_c(z, t) = e^{\omega_c t} \varphi_{\omega_c}(z) = \Delta_0 e^{\omega_c t} [\cos(kz) + R], \quad (5.3.14)$$

where ω_c is the eigenvalue, and $\varphi_{\omega_c}(z)$ the corresponding eigenfunction of Eq. (5.3.13),

$$\omega_c = M[-A(T - T_{c0}) + 4Q\sigma_r - \varepsilon_b^{-1} - Dk^2], \quad (5.3.15)$$

and R a constant given by

$$R = -\frac{2M\phi}{\varepsilon_b h k \{M[4Q\sigma_r - A(T - T_{c0})] - \omega_c\}} \sin\left(\frac{kh}{2}\right). \quad (5.3.16)$$

The $P = 0$ solution is unstable when $\omega_c > 0$ because in this case Δ increases exponentially with time. It can be seen from Eq. (5.3.15) that when the

temperature T is sufficiently high, $\omega_c < 0$ and the paraelectric state is stable. When T is sufficiently low, ω_c turns positive and the paraelectric state is no longer stable, since any small perturbation Δ will grow exponentially beyond all bounds. The critical condition, $\omega = 0$, yields, in this case, the supercooling transition temperature T_c of the film

$$T_c = T_{c0} + \frac{4Q\sigma_r}{A} - \frac{1}{A\varepsilon_b} - \frac{D}{A}k_c^2, \quad (5.3.17)$$

where the first term on the RHS is related to the bulk ferroelectric property, the second term to the misfit of the substrate, the third term to the induced surface charges, and the fourth term to the lattice relaxation on the surface. While the last two terms act against the transformation by lowering the transition temperature, the second term can act either way, depending on the direction of the misfit, i.e., the sign of σ_r . Here, k_c depends on the film thickness h , the extrapolating length δ , and the electric boundary condition through ϕ , as the smallest non-zero root of the equation

$$\cot\left(\frac{k_c h}{2}\right) = k_c \delta + \frac{2\phi}{\varepsilon_b k_c h [4Q\sigma_r - A(T_c - T_{c0})]}. \quad (5.3.18)$$

The general relation between T_c and the film thickness h can be obtained by solving Eqs. (5.3.17) and (5.3.18) simultaneously. Furthermore, as it will be shown in the following, simple analytic approximations can also be derived in many important cases. The cooling critical thickness (CCT) of the film h_{cc} , below which ferroelectric transition by cooling is not possible, can be derived by putting $T_c = 0$ in Eq. (5.3.17), yielding the following expression for the corresponding k_c^2 , which we denote by k_{cc}^2 in terms of the material properties of the film,

$$k_{cc}^2 = \frac{1}{D}(AT_{c0} + 4Q\sigma_r - \frac{1}{\varepsilon_b}). \quad (5.3.19)$$

h_{cc} can then be obtained by solving for h as the smallest positive root of Eq. (5.3.18) corresponding to $T_c = 0$, with k_{cc} given by Eq. (5.3.19). Since h_{cc} varies inversely with k_{cc} according to Eq. (5.3.18) and Eq. (5.3.19), the CCT for these materials will increase with decreasing bulk transition temperature and substrate constraint, and with increasing depolarization field. It is further noted that, for k_{cc} to be real, $\varepsilon_b(AT_{c0} + 4Q\sigma_r) > 1$ must be satisfied, putting a necessary condition on the possibility of the cooling transition from the paraelectric state.

a) *Asymptotic solutions for thin films (i.e., $h \ll \delta$)*

When $hk_c \ll 1$, Eq. (5.3.18) can be solved via a series expansion of the left hand side, yielding

$$k_c^2 \approx \frac{2 - \phi G_c}{h(\delta + h/6)}, \quad (5.3.20)$$

from which T_c is given by

$$T_c \approx T_{c0} + \frac{4Q\sigma_r}{A} - \frac{1}{A\epsilon_b} - \frac{D(2 - \phi G_c)}{Ah(\delta + h/6)}, \quad (5.3.21)$$

where G_c is a function of T_c , and $G_c \equiv 2\epsilon_b^{-1}[4Q\sigma_r - A(T_c - T_{c0})]^{-1}$.

When $h \ll \delta$, the condition $k_c h \ll 1$ is satisfied from Eq. (5.3.20). Thus, T_c in Eq. (5.3.21) can be solved explicitly to give

$$T_c \approx T_{c0} + \frac{4Q\sigma_r}{A} - \frac{1}{A\epsilon_b} - \frac{2D}{Ah(\delta + h/6)} \quad \text{for } \phi=0; \quad (5.3.22a)$$

$$T_c \approx T_{c0} + \frac{4Q\sigma_r}{A} - \frac{2D}{Ah(\delta + h/6)} \quad \text{for } \phi=1. \quad (5.3.22b)$$

Thus, independent of the contact condition, the critical temperature decreases monotonically as the film thickness h decreases. The corresponding CCT can be obtained by putting $T_c = 0$ in Eq. (5.3.22), and solving for h :

$$h_{cc} \approx \frac{2D}{\delta[(4Q\sigma_r + AT_{c0}) - \epsilon_b^{-1}]} \quad \text{for } \phi=0 \text{ (dielectric contacts)}, \quad (5.3.23a)$$

$$h_{cc} \approx \frac{2D}{\delta(4Q\sigma_r + AT_{c0})} \quad \text{for } \phi=1 \text{ (short-circuit electrodes)}. \quad (5.3.23b)$$

Eqs. (5.3.23a) and (5.3.23b) show that the CCT decreases with decreasing surface relaxation, i.e., increasing extrapolation length δ for both contact conditions. In addition, the short-circuit boundaries, with a smaller depolarization field, give a lower CCT than the dielectric boundaries. Furthermore, h_{cc} may become negative if σ_r is negative and large. Both results can be predicted from physical considerations.

b) *Asymptotic solutions for thick films (i.e., $h \gg \delta$) with $\phi G_c \ll 1$*

In the neighborhood of $k_c h \approx \pi$, the left hand side of Eq. (5.3.18) can be approximated by $\cot(x) \approx \frac{\pi}{2} - x$. This approximation yields

$$k_c \approx \frac{\pi}{h + 2\delta} - \frac{2\phi G_c}{\pi h}. \quad (5.3.24)$$

Since the condition $k_c h \approx \pi$ is satisfied when $h \gg \delta$ and $G_c \ll 1$, k_c in Eq. (5.3.24) is a solution of Eq. (5.3.18) for thick films with a small depolarization field. In this case, an explicit relation of T_c can be obtained from Eq. (5.3.17), which, taking into account $\delta \ll h$, and $\phi G_c \ll 1$, can be solved for T_c to give

$$T_c \approx T_{c0} + \frac{4Q\sigma_r}{A} - \frac{D}{A} \left(\frac{\pi}{h} \right)^2. \quad (5.3.25)$$

The behavior of T_c in Eqs. (5.3.22) and (5.3.25) is similar. As the film thickness h is reduced, its critical temperature decreases monotonically.

The corresponding CCT is given by

$$h_{cc} \approx \sqrt{\frac{\pi^2 D}{4Q\sigma_r + AT_{c0}}}. \quad (5.3.26)$$

Note that, even for thick films, the constraint of the substrate and the surface relaxation cannot be neglected, as long as there is a surface in the sample. Furthermore, h_{cc} may become imaginary if σ_r is negative and sufficiently large.

(2) Enhanced surface polarization with the $\delta < 0$ case

In rare cases, polarization may be enhanced, instead of reduced, on the surface of some materials. In this case, $\delta < 0$ and, similar to the $\delta > 0$ case discussed in the foregoing, the solution of Eq. (5.3.13) can be written as

$$A = e^{\omega_c z} \varphi_{\omega_c}(z). \quad (5.3.27)$$

The eigenvalue ω_c is given by

$$\omega_c = M \left[-A(T - T_{c0}) + 4Q\sigma_r - \frac{1}{\varepsilon_b} + Dk^2 \right]. \quad (5.3.28)$$

The eigenfunction is given by $\varphi_{\omega_c}(z) = \Delta_0 [\cosh(kz) + R]$, with

$$R = -\frac{2M\phi}{\varepsilon_b h k \{M[4Q\sigma_r - A(T - T_{c0})] - \omega_c\}} \sinh\left(\frac{kh}{2}\right). \quad (5.3.29)$$

The cooling critical temperature T_c is obtained by putting $\omega_c = 0$ in Eq. (5.3.27):

$$T_c = T_{c0} + \frac{4Q\sigma_r}{A} - \frac{1}{A\varepsilon} + \frac{D}{A} k_c^2, \quad (5.3.30)$$

where k_c is related to the film thickness through boundary conditions, as the largest root, corresponding to the highest T_c , of the equation

$$k_c h \coth\left(\frac{k_c h}{2}\right) = k_c^2 h \gamma + \frac{2\phi}{\varepsilon_b [4Q\sigma_r - A(T_c - T_{c0})]}, \quad (5.3.31)$$

where γ is defined as $\gamma \approx \delta > 0$. Similar to the $\delta > 0$ case, h and k_c can be shown to vary inversely with each other in general. This behavior in Eq. (5.3.30) leads to a value of T_c that decreases as the film thickness h increases, i.e., opposite to the $\delta > 0$ case. Physically, the reason for this behavior is clear, as caused by the enhanced polarization on the surface.

The major difference between the $\delta > 0$ and $\delta < 0$ cases lies in the sign for the polarization gradient in the two cases. In the $\delta < 0$ case, the enhanced surface polarization, as represented by the positive gradient term, together with the misfit interaction of the transformation strain (if $\sigma_r > 0$), favors the ferroelectric state over the paraelectric state. From Eq. (5.3.30), it is clear that, unless the depolarization field caused by the surface induced charges is sufficiently large, T_c cannot be zero as long as a solution of k_c in Eq. (5.3.31) exists, independent of the film thickness.

The foregoing results show that, if the polarization is enhanced on the surface, i.e., $\delta < 0$, a positive value of the CCT in ferroelectric materials need not exist and, in such a case, the transition to the ferroelectric state by cooling is always possible. In this case, the existence of the CCT depends on whether the depolarization field due to the surface induced charges is sufficiently large. For the more common $\delta > 0$ case, on the other hand, this condition does not have to hold for the CCT to exist.

a) Asymptotic solutions for thin films, i.e., $h \ll \gamma$

Similar to the case of $\delta > 0$, analytic approximations of the cooling transition temperature and the CCT can also be derived.

When $h k_c \ll 1$, expanding the left hand side of Eq. (5.3.31) to second order in $h k_c$, the solution can be written as

$$k_c^2 \approx \frac{2 - \phi G_c}{h(\gamma - h/6)}, \quad (5.3.32)$$

so that

$$T_c \approx T_{c0} + \frac{4Q\sigma_r}{A} - \frac{1}{A\varepsilon} + \frac{D(2 - \phi G_c)}{Ah(\gamma - h/6)}. \quad (5.3.33)$$

The validity of solution, i.e., Eq. (5.3.32), is satisfied when $h \ll \gamma$. An explicit solution of T_c from Eq. (5.3.33) is given by

$$T_c \approx T_{c0} + \frac{4Q\sigma_r}{A} - \frac{1}{A\varepsilon_b} + \frac{2D}{Ah(\gamma - h/6)} \quad \text{for } \phi = 0; \quad (5.3.34a)$$

$$T_c \approx T_{c0} + \frac{4Q\sigma_r}{A} + \frac{2D}{Ah(\gamma - h/6)} \quad \text{for } \phi = 1. \quad (5.3.34b)$$

These expressions show explicitly that, independent of the contact condition, the critical temperature increases monotonically as the film thickness h decreases. This is opposite to the $\delta > 0$ case. At the same time, T_c for $\phi = 1$ is positive definite for any film thickness $h < \gamma/6$. In such a case, the CCT does not exist.

From Eq. (5.3.34a), when the depolarization field is sufficiently large, i.e., $(AT_{c0} + 4Q\sigma_r) < \varepsilon_b^{-1}$, a solution of h_{cc} for $T_c = 0$ exists when $\phi = 0$. This is given by

$$h_{cc} \approx \frac{2D}{\gamma[\varepsilon_b^{-1} - (4Q\sigma_r + AT_{c0})]} \quad \text{for } \phi = 0 \text{ (dielectric boundaries)}, \quad (5.3.35)$$

if $\gamma/h_{cc} \gg 1/6$.

b) Asymptotic solutions for thick films ($h \gg \gamma$) with $\phi G_c \ll 1$

When the solution lies in the neighborhood of $k_c h > \pi$, the left hand side of Eq. (5.3.31) can be approximated by $\coth(x) \approx 1$. This approximation yields

$$k_c^2 \approx \left(\frac{1}{\gamma} - \frac{\phi G_c}{h} \right)^2, \quad (5.3.36)$$

that satisfies $k_c h \approx \pi$ when $h \gg \gamma$ and $\phi G_c \ll 1$. An explicit expression of T_c can be derived by substituting into Eq. (5.3.30):

$$T_c = T_{c0} + \frac{4Q\sigma_r}{A} - \frac{1}{A\varepsilon_b} + \frac{D}{A\gamma^2}, \quad \text{for } \phi = 0 \text{ and } \phi = 1. \quad (5.3.37)$$

In both cases, T_c is independent of h , and the critical thickness does not exist.

At this point, it is obvious that the analytic results for second-order transitions,

obtained using the static Ginzburg-Landau equation, are consistent with the present results. Results are also applicable to first-order transitions from the paraelectric branch to the ferroelectric branch yielding, instead of the Curie temperature, the stability limit of the paraelectric phase during cooling, i.e., the supercooling transition temperature. To obtain the Curie temperature and the critical thickness, the next section will analyze the stability of the ferroelectric branch and investigate the stability limit of the ferroelectric phase on heating, i.e., transitions from the ferroelectric branch to the paraelectric branch.

5.3.3.2 Stability of the Ferroelectric State and Transition Characteristics

In the ferroelectric state, the stationary polarization field $P(z)$ is non-zero and the corresponding Eq. (5.3.13) is no longer the same as that in the paraelectric state. Treating the P dependent terms as a perturbation and expanding in terms of the complete orthonormal set of eigenfunctions of the unperturbed equation, Eq. (5.3.13) can be solved within the first-order perturbation approximation. The eigenvalue ω_h in this case is given by

$$\omega_h = A(T - T_c) + 3(B + 4GQ^2)\langle P^2 \rangle + 5C\langle P^4 \rangle, \quad (5.3.38)$$

where T_c is the supercooling temperature obtained in Section 5.3.3.1, and $\langle P^n \rangle$ is the expectation value of P^n given by

$$\langle P^n \rangle = \frac{\int_{-h/2}^{h/2} P^n(z) \varphi_{\omega_c}^2(z) dz}{\int_{-h/2}^{h/2} \varphi_{\omega_c}^2(z) dz}, \quad (5.3.39)$$

where φ_{ω_c} is the eigenfunction of the unperturbed state. Since both P and φ peak at $z = 0$, we may approximate $\langle P^n \rangle$ by $P_0^n = P^n(0)$. This approximation is also valid when P is a slowly varying function. Thus, ω_h may be written as

$$\omega_h = A(T - T_c) + 3(B + 4GQ^2)P_0^2 + 5CP_0^4 = (P_0^2 - p_1^2)(P_0^2 - p_2^2), \quad (5.3.40)$$

where p_1^2 and p_2^2 are the two roots of Eq. (5.3.40) for $\omega_h = 0$. It is clear that p_1^2 and p_2^2 can be both real or complex conjugates. In the simplest case in which they are real, the stability condition of the ferroelectric state is satisfied if P_0^2 lies between p_1^2 and p_2^2 . In the following, several specific cases will be considered.

Let us first consider the case in which $C > 0$. The solutions can be written

down as

$$p_1^2 = -\frac{3(4GQ^2 + B) + d}{10C}; p_2^2 = -\frac{3(4GQ^2 + B) - d}{10C}. \quad (5.3.41)$$

Here d^2 is the discriminant of the quadratic in Eq. (5.3.40), given by

$$d^2 = 9(4GQ^2 + B)^2 - 20AC(T - T_c). \quad (5.3.42)$$

By noting that the roots p_1 and p_2 are both functions of temperature through d , T_{ch} is defined as a temperature by

$$T_{\text{ch}} = T_c + \frac{9(4GQ^2 + B)^2}{20AC}. \quad (5.3.43)$$

For temperatures above T_{ch} , $d^2 < 0$ and ω_h have the same sign as $A(T - T_c)$, i.e., positive definite. Thus, the ferroelectric system will be unstable at temperature above T_{ch} .

In the temperature range $T_c < T < T_{\text{ch}}$, the product $p_1^2 p_2^2$, from Eq. (5.3.40), is equal to $\frac{A(T - T_c)}{5C}$, which is positive definite ($C > 0$). In this temperature regime,

it can be shown that $d^2 > 0$ and the roots in Eq. (5.3.41) are real and have the same sign. In addition, if $4GQ^2 + B < 0$, it can be seen from Eq. (5.3.41) that both roots are positive. In this case, both p_1 and p_2 are real and the ferroelectric state is stable, when P_0 lies between p_1 and p_2 . We note that at $T = T_{\text{ch}}$, $d_2 = 0$ and $p_1 = p_2$ so that P_0 is uniquely defined. At $T = T_c$, the smaller solution becomes zero. Within this temperature regime, which is above the cooling para/ferro-electric transition temperature T_c , the ferroelectric state is still stable and T_{ch} can be identified as the superheating transition temperature. This difference between the heating-up and cooling-down behavior is well known for first-order phase transformations near the critical point, and is caused by the restraint on the realization of the transformation strain in the present case by the rigidity of the substrate through the $4GQ^2$ term.

In the case when $4GQ^2 + B > 0$, both p_1^2 and p_2^2 are negative and P_0 has no real solution in this temperature regime. In this case, the ferroelectric state does not exist for temperatures above T_c and the ferroelectric state cannot be superheated beyond the cooling transition temperature T_c . We note that when $T = T_c$, the largest solution is zero.

At temperatures below T_c , the product of the two roots is negative, and thus only one of them, the largest one, can be positive. For the ferroelectric state to be stable, P_0 must be smaller than the only real root. The system here then behaves as a typical second-order system.

Thus, in the case $C > 0$, the order of para/ferroelectric transition in a thin film depends not on the sign of B , but on the sign of $B + 4GQ^2$. Since $4GQ^2$ is positive, a film made of a first-order bulk material may exhibit a second-order transition behavior if the constraint due to the substrate is sufficiently large.

Plotted in Figs. 5.7a and 5.7b as a function of temperature for various film thicknesses are the regions of stability of P_0 for PbTiO_3 according to the condition $p_2 < P_0 < p_1$ from Eqs. (5.3.41) and (5.3.42). Fig. 5.7a shows the case of a freestanding film, whereas Fig. 5.7b shows the film on a SrTiO_3 substrate. It is obvious that the constraint of the substrate weakens the polarization of the ferroelectric state, reduces the Curie temperature, and changes the first-order transition of the ferroelectric material into a second-order one.

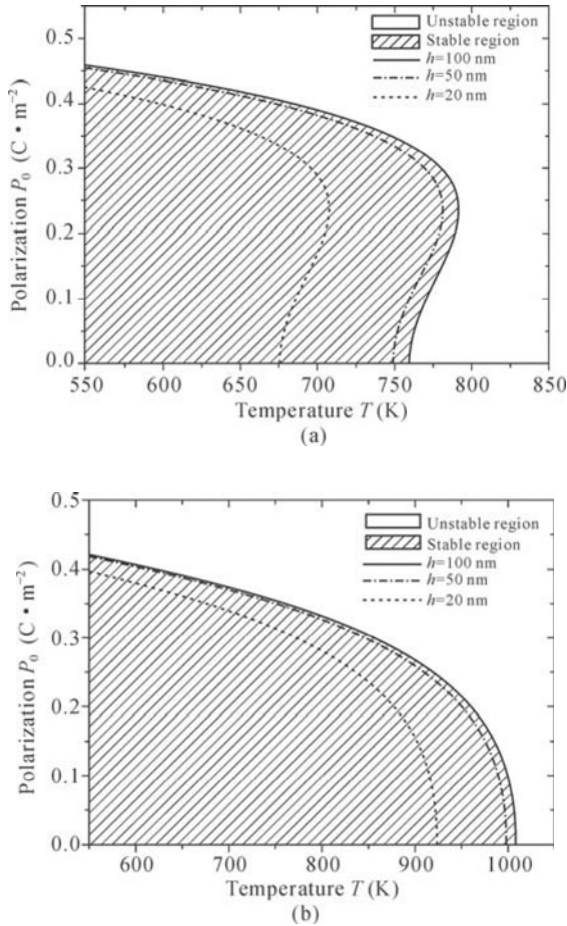


Fig. 5.7 The stable area of ferroelectric to paraelectric transition of free-standing PbTiO_3 thin film (a) and PbTiO_3 thin film (b) on SrTiO_3 substrate

It is interesting that the difference between the superheating and supercooling transition temperatures is insensitive to the film thickness h . This behavior is also seen in the numerical solution of the static Landau-Ginzburg equation for the first-order phase transition of free-standing ferroelectric film. These results show that the two temperatures, plotted as a function of film thickness, are practically parallel.

The case of materials with $C = 0$ can also be easily analyzed using Eq. (5.3.40), which becomes

$$\omega_h = A(T - T_c) + 3(B + 4GQ^2)P_0^2. \quad (5.3.44)$$

If $B + 4GQ^2 > 0$, ω_h is positive definite for $T > T_c$, and the system is unstable. For $T < T_c$, the system is stable as long as $P_0^2 < \frac{A(T_c - T)}{3(4GQ^2 + B)}$. This system behaves like a second-order system. For the rare case in which $B + 4GQ^2 < 0$, the ferroelectric state is stable in the temperature range $T < T_c - \frac{3(B + 4GQ^2)}{A}$, which is above the supercooling transition temperature T_c . This system thus behaves like a first-order system, with a superheating temperature $T_{ch} = T_c - \frac{3(B + 4GQ^2)}{A}$. The case $C < 0$ rarely occurs, and will not be considered. At any rate, the foregoing analysis can be repeated easily for this case also. Analytic expressions of the Curie temperature and the corresponding critical thickness for first-order transitions can thus be obtained by replacing T_{c0} with $T_{c0} + \frac{9(B + 4GQ^2)^2}{20AC}$.

5.3.4 Curie Temperature and Critical Thickness

The following discussions refer to the specific examples of PbTiO_3 and BaTiO_3 . Strictly speaking, the foregoing results apply only to uniaxial ferroelectrics and not to cubic perovskites like PbTiO_3 and BaTiO_3 . However, even for these perovskite ferroelectric thin films under the short-circuit condition, the polarization field is often found to be perpendicular to the surface. The critical thickness of the thin film at $T_c = 0$ can be calculated using Eqs. (5.3.17) and (5.3.18), following the usual assumption of a reduced polarization field on the surface, i.e., $\delta > 0$.

A BaTiO_3 film is considered, which is epitaxially grown on a thick SrTiO_3 substrate. If stress relaxation during fabrication is neglected, the misfit strain can reach 2%, corresponding to a misfit compressive stress of 3.57 GPa. From the second term of Eq. (5.3.17), this large compressive misfit stress may raise the Curie temperature by an unrealistic value of 900 °C. Experimentally, Yanase et al.

(1999) measured a Curie temperature of 350 °C. Compared with an inherent Curie temperature of BaTiO₃ of 130 °C, one may estimate that the Curie temperature is raised only by 220 °C, corresponding to a misfit stress and strain of 852 MPa and 0.00477, respectively. In the first-principle calculation of Junquera and Ghosez (2003), the critical thickness of BaTiO₃ under short-circuit boundaries conditions was found to have a value of about 2.4 nm. The corresponding extrapolating length can be estimated by numerically solving Eqs. (5.3.17) and (5.3.18) simultaneously, giving a value of $\delta \approx 0.5$ nm, which satisfies $\delta \ll h$.

Note that the small extrapolating length physically means that the tetragonal phase on the surface becomes almost central symmetric due to lattice relaxation, resulting in a large reduction in the polarization on the surface. The depolarizing effect does not come from the usual depolarization field due to induced electric charges on the surface. Indeed, Eq. (5.3.26) shows explicitly that this is the case. Thus, even with zero depolarizing charges, i.e., both the third term on the right hand side of Eq. (5.3.17) and the last term in Eq. (5.3.18) vanish, a solution of h_c for these equations may still exist, as can be seen explicitly in Eq. (5.3.23), which is independent of the depolarization field. This conclusion is independent of the contact conditions.

From the numerical solution of Eq. (5.3.19), the relation between the critical thickness and the extrapolation length of BaTiO₃ thin film grown on SrTiO₃ substrate with the short circuit boundary conditions is shown in Fig. 5.8. The decrease in the extrapolation length δ decreases the surface polarization and causes the critical thickness to increase. The corresponding analytic approximations for the cases of $\delta > 0$, and $h_c \ll \delta$ or $h_c \gg \delta$ are derived in Eqs. (5.3.23) and (5.3.26), and can be given by h_c (nm) = 3.5 – 2 δ , and h_c (nm) = 2.2/ δ respectively.

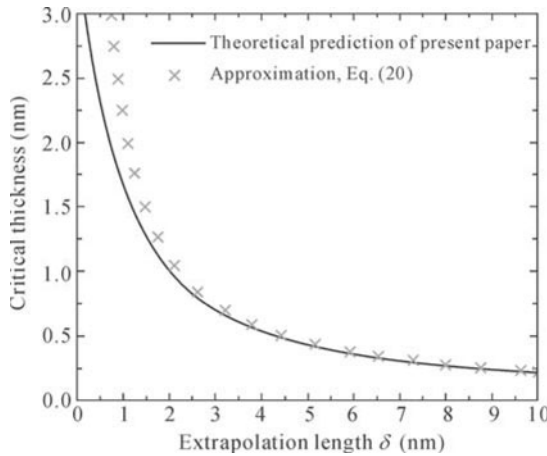


Fig. 5.8 The critical thickness versus the extrapolation length for BaTiO₃ thin film with the short circuit boundary conditions

By assuming that the film is fully strained epitaxially, then with a compressive epitaxial strain of $\varepsilon_{xx} = \varepsilon_{yy} = 1.2\%$, the epitaxial stresses can be calculated to be $\sigma_r = \sigma_{xx} = \sigma_{yy} = -2.4293 \times 10^9 \text{ N} \cdot \text{m}^{-2}$, the electrostrictive coefficient $Q = -0.026$ (in SI unit) and the Curie temperature is found to increase by about 335 K. The transition temperature is a function of film thickness as given by Eqs. (5.3.17) and (5.3.18) (for $\phi = 0$). The results are shown in Fig. 5.9, together with the experimental points of Streiffer et al. The transition temperature can be calculated analytically using Eqs. (5.3.22a) and (5.3.26) for large and small values of h , respectively. For PbTiO_3 these equations can be put into the simple forms $T_c = 1041 - 36000/h^2$ for $h > 20 \text{ nm}$ and $T_c = 1041 - 1440/h$ for $h < 20 \text{ nm}$. Plotted together on the dotted line is the theoretical prediction used for comparison by Streiffer et al. The difference between the two theoretical results in the small film thickness regime is obvious.

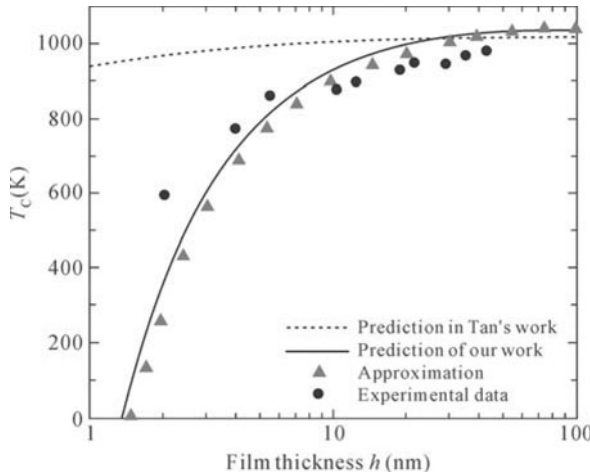


Fig. 5.9 Relation between Curie temperature and film thickness of PbTiO_3 thin film with the short-circuit boundary conditions (Tan et al., 2000; Streiffer et al., 2002; Wang and Woo, 2005)

5.3.5 Curie-Weiss Law of Ferroelectric Thin Film in Nanoscale

5.3.5.1 The Dynamic Equation of the Self-Polarization in the Thin Film

In this section, a ferroelectric thin film is considered only with the open circuit boundary conditions. Effects of the epitaxial stresses are neglected in the following calculations. It is clear that their inclusion has only the equivalent effect of shifting the bulk transition temperature. Following Wang and Woo (2005), with the help of the Ginzburg-Landau functional, the dynamic equation of the self-polarization in the thin film can be written as

$$M \frac{\partial P}{\partial t} = -\frac{\delta F}{\delta P} = -A(T - T_{c0})P - (B + 4GQ^2)P^3 - CP^5 + E_d + D \frac{d^2 P}{dz^2}, \quad (5.3.45)$$

with the boundary conditions $\frac{dP}{dz} = \mp P/\delta$, $z = \pm h/2$. The electric and polarization fields in the ferroelectric are related through conditions on the boundaries via the laws of electrostatics. Here, a ferroelectric thin film in an open-circuit is considered; i.e., the film is insulated from the electrodes and the depolarizing field E_d can be written as Eq. (5.3.8), i.e., $E_d = -P/\epsilon_b$.

5.3.5.2 Steady-State Solution Near the Bifurcation Point

The condition of the dynamic stability of the paraelectric state of the evolution equation Eq. (5.3.45) has been established, from which the critical temperature T_c at which the state becomes unstable can be obtained as a function of the film thickness h

$$T_c = T_{c0} - \frac{1}{A\epsilon_b} - \frac{D}{A} \cdot k_{z\min}^2, \quad (5.3.46)$$

where $k_{z\min}$ is the minimum solution of the following transcendental equation $\cot(k_z h/2) = k_z \delta$.

As explained in Wang and Woo (2005), T_c is the para/ferro-electric transition temperature for the thin film. Putting $\frac{\partial P}{\partial t} = 0$ in Eq. (5.3.45), the corresponding stationary equation of the order parameter P can be written as

$$L_c P \equiv -A(T_c - T_{c0})P - \frac{1}{\epsilon_b} P + D \frac{d^2 P}{dz^2}, \quad (5.3.47)$$

where L_c is the parabolic operator evaluated at the bifurcation point $T = T_c$, in the neighborhood of which the solution P is small, if assumed continuous. Following Nicholis and Prigogine (1977), both P and $\gamma \equiv T - T_c$ can be expanded in a power series in terms of a small perturbation λ from the critical point:

$$\begin{aligned} P &= \lambda P_1 + \lambda^2 P_2 + \dots, \\ \gamma &= T - T_c = \lambda \gamma_1 + \lambda^2 \gamma_2 + \dots. \end{aligned} \quad (5.3.48)$$

This expansion is more flexible than the seemingly more natural one in which P is expanded in a power series of $(T - T_c)$. More importantly, it allows fractional

power dependence of P on $(T - T_c)$. By substituting Eq. (5.3.48) into Eq. (5.3.47) and equating coefficients of equal powers of λ , a set of relations of the following form can be obtained:

$$L_c P_k = a_k, \quad k = 1, 2, \dots, \quad (5.3.49)$$

which have to be satisfied together with the boundary and the initial conditions $\frac{dP_k}{dz} = \mp P_k / \delta$, $z = \pm h/2$ (Wang and Woo, 2006).

The first several coefficients a_k are

$$\begin{aligned} a_1 &= 0, \\ a_2 &= A\gamma_1 P_1, \\ a_3 &= A\gamma_1 P_2 + A\gamma_2 P_1 + (B + 4GQ^2)P_1^3. \end{aligned} \quad (5.3.50)$$

It is easy to check that under the boundary condition, the operator L_c is self-adjoint and

$$(L_c P^*, P_k) = \int_{-h/2}^{h/2} P_k L_c P^* dz = (P^*, L_c P_k) = (P^*, a_k) = \int_{-h/2}^{h/2} P^* a_k dz = 0, \quad (5.3.51)$$

where P^* is the solution of the following homogeneous equation:

$$\begin{aligned} L_c P^* &= 0; \\ \frac{dP^*}{dz} &= \mp \frac{P^*}{\delta}, \quad z = \pm \frac{h}{2}. \end{aligned} \quad (5.3.52)$$

Eq. (5.3.51) can be used to determine the coefficients γ_i . Then, from the second relation of Eq. (5.3.48), one can determine λ as a function of $(T - T_c)$. Substituting the resulting λ into the first relation of Eq. (5.3.48) and solving the inhomogeneous boundary condition equations result in an explicit expression for the solution P .

Let us consider the $\delta > 0$ case, which is satisfied by most ferroelectrics. Eq. (5.3.46) dictates that the transition temperature of the thin film is lower than that of the bulk material. The solution of Eq. (5.3.52) can be written as

$$P^* = M \cos(\beta z) + N \sin(\beta z), \quad (5.3.53)$$

where

$$\beta = \sqrt{A(T_{c0} - T_c) / D - 1 / \epsilon_b D}. \quad (5.3.54)$$

Suppose P is distributed symmetrically along the thickness direction, i.e., $\frac{dp^*}{dz} = 0$ at $z = 0$. Then $N = 0$, and the solution Eq. (5.3.53) becomes

$$P^* = M \cos(\beta z). \quad (5.3.55)$$

Similarly, the solution of the first equation in Eq. (5.3.49) can be obtained in the same form as P^* . Thus,

$$P_1 = M \cos(\beta z). \quad (5.3.56)$$

By using the orthogonal condition in Eq. (5.3.51), γ_1 and γ_2 can be determined,

$$\gamma_1 = 0; \quad \text{since} \quad \int_{-h/2}^{h/2} P_1 P^* dz = \int_{-h/2}^{h/2} P_1^2 dz \neq 0, \quad (5.3.57)$$

$$\gamma_2 = -\frac{B + 4GQ^2}{A} \cdot \frac{\int_{-h/2}^{h/2} P_1^4 dz}{\int_{-h/2}^{h/2} P_1^2 dz}. \quad (5.3.58)$$

If the corresponding bulk ferroelectrics transition is of second-order, i.e., $B > 0$, $B + 4GQ^2$ is also positive and substitution of Eq. (5.3.58) into Eq. (5.3.48) gives

$$\lambda = \pm \left(\frac{T_c - T}{-\gamma_2} \right)^{1/2} \quad \text{for } T < T_c, \quad (5.3.59)$$

i.e., the solution is defined in the supercritical (ferroelectric) regime. According to Nicholis and Prigogine (1977), the bifurcating solutions are both asymptotically stable in the supercritical regime. The phase diagram for this case is shown in Fig. 5.10a. Thus, if the bulk ferroelectrics exhibit second-order phase transition, their thin-film counterparts will also exhibit the same order of phase transition and the universal critical exponents remain unchanged.

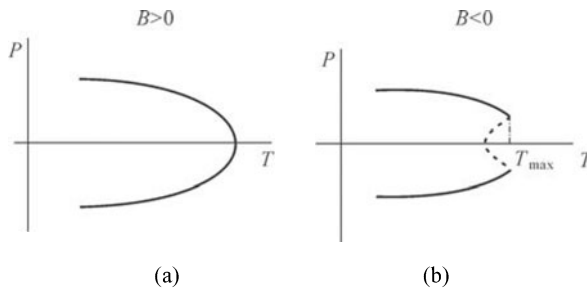


Fig. 5.10 Schematic of bifurcation diagrams for ferroelectric thin films. $B > 0$ (a) and $B < 0$ (b) mean that the phase transitions of ferroelectric thin films are first order and second order, respectively

On the other hand, if the corresponding transition in the bulk material is first-order, i.e. $B < 0$, then depending on the magnitude of the elastic self-energy of the transformation volume $4GQ^2$, which is positive definite and has a magnitude of the same order as B , $B + 4GQ^2$ can be either positive or negative. When $4GQ^2$ is sufficiently large, $B + 4GQ^2$ is positive definite, as in BaTiO_3 and PbTiO_3 . Then a first-order bulk ferroelectric becomes a second-order one in a rigidly constrained thin film. This is consistent with the results of Pertsev et al. (1998)'s work. For the two classical Perovskite ferroelectrics, BaTiO_3 and PbTiO_3 , these authors also found that the two-dimensional clamping of the film could result in a change in transition order. Nevertheless, the results of Wang and Woo (2005) should be interpreted, noting the assumption that the gradient terms in the free energy were negligible.

If the transformation volume satisfies $B + 4GQ^2 = 0$, the solution will include higher order terms of expansion, and thus the critical exponents will differ from their bulk counterpart. If $B + 4GQ^2$ is negative, substitution of Eq. (5.3.58) into Eq. (5.3.48) gives

$$\lambda = \pm \left(\frac{T - T_c}{\gamma_2} \right)^{1/2} \quad \text{for } T > T_c, \quad (5.3.60)$$

i.e., the solution is defined only in the subcritical (paraelectric) regime, and both branches are unstable. Physically, there is a maximum value of $T_{\max}$ above which only one solution is admissible. Therefore, the subcritical branches have to turn, at some value $T_{\max}$, in the direction of decreasing temperature (Fig. 5.11b). Thus, the solution of P near the critical point can be obtained in the form

$$P = P_1 \lambda = \begin{cases} \pm M \cos(\beta z) \left(\frac{T_c - T}{-\gamma_2} \right)^{1/2}, & B + 4GQ^2 > 0 \text{ and } T < T_c, \\ \pm M \cos(\beta z) \left(\frac{T - T_c}{\gamma_2} \right)^{1/2}, & B + 4GQ^2 < 0 \text{ and } T > T_c. \end{cases} \quad (5.3.61)$$

It is clear from Eq. (5.3.61) that, for rigidly constrained thin film ferroelectrics, the order of transitions is independent of its thickness, but depends on the elastic self-energy of the transformation strain under the constraint of the substrate. Furthermore, we note that the second solution in Eq. (5.3.61) is unstable and that the actual transition is discontinuous. The behavior of the self-polarization P in the foregoing discussions agrees completely with that derived earlier using linear stability analysis (Wang and Woo, 2006).

5.3.5.3 Dielectric Susceptibility of Ferroelectric Films Near the Critical Point

The susceptibility χ^T of a ferroelectric material measures the response of the total

(i.e., observable) polarization in an external field, and is defined by

$$\chi^T = \left. \frac{\partial P_{total}}{\partial E_{ext}} \right|_{E_{ext} \rightarrow 0}. \quad (5.3.62)$$

For the present boundary conditions, it can be shown from the laws of electrostatics

$$\chi^T = \left. \frac{\partial P}{\partial E_{ext}} \right|_{E_{ext} \rightarrow 0} + \chi_b \equiv \chi + \chi_b. \quad (5.3.63)$$

It is sufficient to calculate χ , which is much larger than χ_b near the transition temperature. χ can be calculated via the total free energy as usual,

$$\begin{aligned} \chi^{-1} &= \frac{\partial^2 f}{\partial P^2} = A(T - T_{c0}) + \frac{1}{\varepsilon_b} + 3(B + 4GQ^2)P^2 + 5CP^4 + D\beta^2 \\ &= A(T - T_c) + 3(B + 4GQ^2)P^2 + 5CP^4. \end{aligned} \quad (5.3.64)$$

In deriving Eq. (5.3.64), Eqs. (5.3.46) and (5.3.54), the following relations have been used:

$$\begin{aligned} \frac{\partial^2}{\partial P^2} \left(\frac{\partial P}{\partial z} \right)^2 &= \frac{\partial}{\partial P} \left[\frac{\partial z}{\partial P} \cdot \frac{\partial}{\partial z} \left(\frac{\partial P}{\partial z} \right)^2 \right] = 2 \frac{\partial}{\partial P} \left(\frac{\partial^2 P}{\partial z^2} \right) \\ &= 2 \frac{\partial z}{\partial P} \cdot \frac{\partial^3 P}{\partial z^3}. \end{aligned} \quad (5.3.65)$$

For positive $B + 4GQ^2$, substitution of the first equation of Eq. (5.3.61) and β in Eq. (5.3.54) into Eq. (5.3.64), yields the inverse susceptibility near T_c ,

$$\chi^{-1} = \begin{cases} [A + 3(B + 4GQ^2)M^2 \cos^2(\beta z) / \gamma_2](T - T_c), & \text{for } T < T_c, \\ A(T - T_c), & \text{for } T > T_c. \end{cases} \quad (5.3.66)$$

The average value of χ cannot be integrated analytically, like the average value of the reciprocal. Thus,

$$\begin{aligned} \langle \chi^{-1} \rangle &= \frac{1}{h} \int_{-h/2}^{h/2} \chi_s^{-1} dz = A \left(1 - \frac{3\theta(T_c - T)}{h} \frac{[\int_{-h/2}^{h/2} \cos^2(\beta z) dz]^2}{\int_{-h/2}^{h/2} \cos^4(\beta z) dz} \right) (T - T_c) \\ &= A \left[1 - 3\theta(T_c - T) \frac{4 \sin^2(\beta h) + 8\beta h \sin(\beta h) + 4\beta^2 h^2}{\beta h \sin(2\beta h) + 8\beta h \sin(\beta h) + 6\beta^2 h^2} \right] (T - T_c), \end{aligned} \quad (5.3.67)$$

where $\theta(x)$ ($\theta(x) = 0$ for $x < 0$, and $\theta(x) = 1$ for $x > 0$) is the step function. Eq. (5.3.67) may be rewritten in the form

$$\langle \chi^{-1} \rangle^{-1} = \frac{A^{-1} \Theta^{-1}(h)}{T - T_c}. \quad (5.3.68)$$

Eq. (5.3.68) shows that, on both sides of the transition temperature, the Curie-Weiss law holds, and that the average susceptibility of the film diverges with a critical exponent of 1. In addition, the corresponding Curie parameter $A^{-1} \Theta^{-1}(h)$ also changes sign discontinuously across the transition temperature. Both aspects of its behavior are the same as the bulk counterpart. However, for films of finite thickness, the magnitude of the Curie parameter varies according to the thickness, leading to asymmetric divergences across the transition temperature. Indeed, the Curie parameter is independent of the film thickness ($\Theta^{-1} = 1$) on the paraelectric side ($T > T_c$), a well known fact, but depends on the film thickness on the ferroelectric side. Taking the limits in Eq. (5.3.67) and using the relations between β and h derived in Wang and Woo (2005), it can be shown that the value of Θ^{-1} depends on the film thickness h , approaching $-(1-8\delta/(3h))$ for thick films ($h \gg \delta$) and $-0.5(1+h^2/(5\delta^2))$ for thin films ($h \ll \delta$).

Accordingly, for a film thickness of $h = 8\delta$, the Curie parameter Θ^{-1} of the ferroelectric phase would have dropped by 33%. These characteristics are shown in Fig. 5.11, where the normalized susceptibility, i.e., $\Theta^{-1}/(T - T_c)$ as a function of temperature, is compared between thick films ($h \rightarrow \infty$) and thin films ($h \rightarrow 0$).

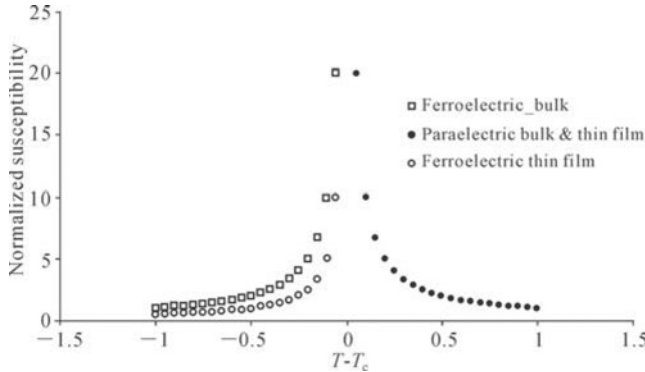


Fig. 5.11 Normalized susceptibility versus the temperature of ferroelectric thin film with the open circuit boundary condition

Note that $\langle \chi^T \rangle$ also has a similar relation to $T - T_c$, as discussed in the foregoing.

From the free energy expression (Wang and Woo, 2005), the contribution from the upper and lower surfaces of the ferroelectric thin film increases the susceptibility according to

$$\Delta\chi = \frac{2\delta}{D}, \quad (5.3.69)$$

which is negligible near the transition temperature.

The example of a well-known ferroelectric PbTiO_3 is considered. The normalized Curie parameter $-\left[\mathcal{C}(h)\right]^{-1}$ is plotted against the film thickness h in Fig. 5.12, in which the Curie temperature was obtained as a function of h . It can be seen that the magnitude of the normalized Curie parameter does decrease rapidly from 1 to 0.5 with decreasing thickness. Indeed, for films of PbTiO_3 about 40 nm ($= 8\delta$), the ferroelectric transition is accompanied with a decrease of the Curie parameter of about 33%.

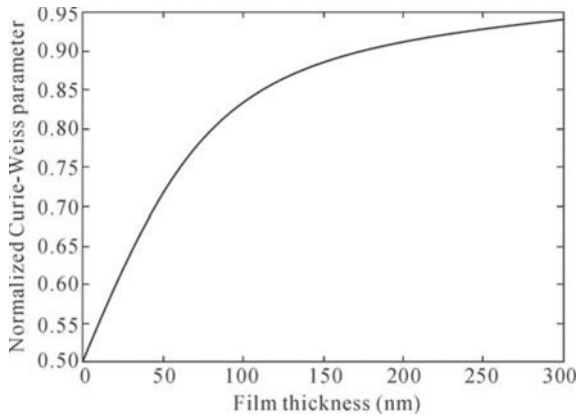


Fig. 5.12 The normalized Curie-Weiss parameter $-\left[\mathcal{C}(h)\right]^{-1}$ versus the film thickness h for PbTiO_3 with the open circuit boundary conditions

5.4 Critical Properties of Ferroelectric Capacitors or Tunnel Junctions

In this section, a symmetric metal-ferroelectric-metal heterostructure in nanoscale, such as a ferroelectric capacitor or tunnel junction, is considered with a single-domain ferroelectric thin film of thicknesses h sandwiched between electrodes each of thickness $(L-h)/2$ (Fig. 5.13a). Under short-circuit boundary conditions, compensation charges $\pm q_c$ exist in the electrodes as shown in Fig. 5.13b. Considering the structure of most metal-ferroelectric-metal capacitors or tunnel junctions grown on compressive substrates, and assuming that all vector fields in the ferroelectric thin film, such as the polarization and electric field, are directed in the z -direction, we can formulate the problem. To be specific, BaTiO_3 thin film grown on SrTiO_3 substrate is considered in the present section, which ferroelectric phase transition changes from first order to second order due to the effect of the isotropic compressive misfit strain.

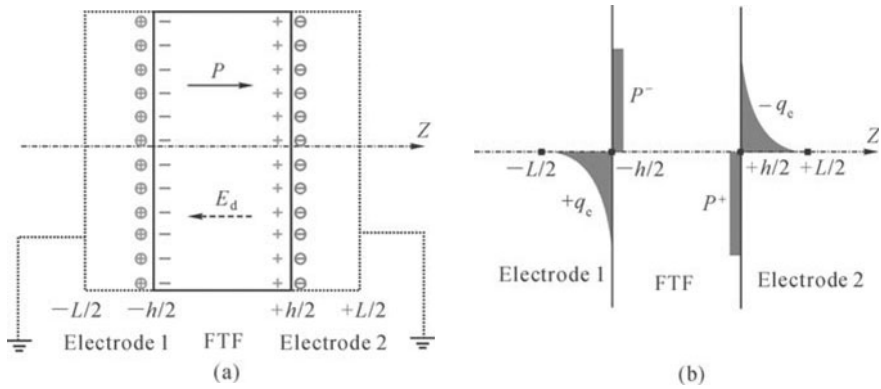


Fig. 5.13 Schematic structures of (a) a metal-ferroelectric-metal with the short circuit boundary conditions and (b) charge distribution in a poled metal-ferroelectric-metal capacitor or tunnel junction

Many previous works have shown that the thermodynamic potential of a metal-ferroelectric-metal structure grown on a compressive substrate with short-circuit boundary conditions should generally have six contributions: (i) the Landau free energy of a bulk ferroelectric, (ii) the elastic interaction between the misfit strain and transformation strain, (iii) the electrostrictive coupling between the misfit strain and the polarization, (iv) the screened electrostatic interaction between the displacement field and the depolarization charge compensation in the electrodes (incomplete), (v) the interaction between the polarization and the built-in field due to the difference in work function steps at the two ferroelectric/electrode interfaces, and (vi) the change in surface energy due to the gradient of the polarization near the interfaces and the interaction between the polarization and interface.

5.4.1 The Thermodynamic Potential of the Ferroelectric Capacitors or Tunnel Junctions

Expressions of the Landau free energy of a bulk ferroelectric and the elastic interactions are well known. Following Mehta et al. (1973), the electric fields and compensation charges in electrodes 1 and 2 (see Fig. 5.13b) are given by

$$\begin{cases} E_{e1}(z) = q_e \sinh[(L+2z)/(2l_{s1})] \{ \epsilon_{e1} \sinh[(L-h)/(2l_{s1})] \}^{-1}, & \text{for } -\frac{L}{2} \leq z \leq -\frac{h}{2}, \\ E_{e2}(z) = -q_e \sinh[(L-2z)/(2l_{s2})] \{ \epsilon_{e2} \sinh[(L-h)/(2l_{s2})] \}^{-1}, & \text{for } \frac{h}{2} \leq z \leq \frac{L}{2}, \end{cases} \quad (5.4.1)$$

where ϵ_{e1} , ϵ_{e2} and l_{s1} , l_{s2} are the dielectric constants and screening lengths in electrodes 1 and 2, respectively.

Consider a ferroelectric thin film of thickness h . Absent an external electric field, the internal electric field in the ferroelectric thin film can be derived via Maxwell's equations, and is related to the depolarization field by $E_d = -(P + q_e)/\epsilon_b$ for $-h/2 \leq z \leq h/2$. To satisfy the short-circuit boundary condition, we must satisfy

$$\int_{-L/2}^{L/2} E dz = \int_{-L/2}^{-h/2} E_{e1} dz + \int_{-h/2}^{h/2} E_d dz + \int_{h/2}^{L/2} E_{e2} dz = 0. \quad (5.4.2)$$

Since the screening length is usually small compared with the thickness of the electrode, q_e can be written as

$$q_e = -\frac{Ph/\epsilon_b}{l_{s1}/\epsilon_{e1} + l_{s2}/\epsilon_{e2} + h/\epsilon_b}. \quad (5.4.3)$$

In the absence of an external electric field, the depolarization field is a function of the ferroelectric thin film thickness, which can be written as

$$E_d(z) = -\frac{1}{\epsilon_b} \left[1 - \frac{h/\epsilon_b}{l_{s1}/\epsilon_{e1} + l_{s2}/\epsilon_{e2} + h/\epsilon_b} \right] P = -\frac{1}{\epsilon_b} (1 - \theta) P. \quad (5.4.4)$$

The built-in electric field due to the difference in work function steps at the two ferroelectric-electrode interfaces has to be considered if the two ferroelectric-electrode interfaces are different. Let $\Delta\phi_1$ and $\Delta\phi_2$ be the work function steps. For $P = 0$, the built-in field can be written as (Tagantsev et al., 2008; Pertsev and Kohlstedt, 2006; Zheng et al., 2010)

$$E_{bi} = -\frac{\Delta\phi_2 - \Delta\phi_1}{h}. \quad (5.4.5)$$

Following Landau and Lifshitz, for a non-zero electric field in the ferroelectrics, the thermodynamic potential can be obtained from the relation $\frac{dg}{dE} = -D$, where D is the displacement field (Landau et al., 1960). As shown recently, the P^8 expansion should be used in the Landau-Lifshitz (LL) free-energy functional for the BaTiO₃ (Li et al., 2005). In this regard, the thermodynamic potential of the metal-ferroelectric-metal system without the surface energy used in the present calculations is written as

$$\Phi = \alpha^* P^2 + \beta^* P^4 + \frac{\gamma}{6} P^6 + \frac{\zeta}{8} P^8 + (s_{11} + s_{12})^{-1} u_m^2 - PE - \frac{1}{2} \epsilon_b E^2, \quad (5.4.6)$$

where the renormalized coefficients α^* and β^* are

$$\begin{aligned}\alpha^* &= \frac{\alpha_0}{2}(T - T_{c0}) - u_m \frac{2Q_{12}}{s_{11} + s_{12}}, \\ \beta^* &= \frac{\beta}{4} + \frac{Q_{12}^2}{s_{11} + s_{12}},\end{aligned}\tag{5.4.7}$$

where α_0 , β , γ and ς are the P^8 expansion coefficients of the free energy of the reference crystal, T_{c0} is the Curie temperature of the bulk reference crystal, Q_{12} is the electrostrictive coefficient, s_{ij} is the elastic compliance tensor, and u_m is the biaxial in-plane misfit strain between the ferroelectric thin film and substrate (Pertsev et al., 1998).

The surface energy can be written as a Taylor expansion in terms of P . A simple expression of the surface energy including only the two lowest order terms of the Taylor expansion can be written as (Tagntsev et al., 2008)

$$\Phi_s = (\zeta_1 - \zeta_2)P + \frac{1}{2}(\eta_1 + \eta_2)P^2,\tag{5.4.8}$$

where ζ_1 , η_1 and ζ_2 , η_2 are the coefficients of the surface energy expansion for the two ferroelectric/electrode interfaces. By noting the screening lengths l_{s1} , l_{s2} of electrodes in Eq. (5.4.4), the work function steps of the ferroelectric-electrode interfaces $\Delta\phi_1$ and $\Delta\phi_2$ in Eq. (5.4.5), the coefficients of the surface energy ζ_1 , ζ_2 and η_1 , η_2 in Eq. (5.4.8) can be estimated from the first-principle calculations (Tagansev et al., 2006), respectively, which will also be used in the following thermodynamic calculations.

From Eqs. (5.4.4) to (5.4.9), the total free energy F of the metal-ferroelectric-metal system can be written as

$$\begin{aligned}F &= \Phi h + \Phi_s \\ &= \left[\alpha^* P^2 + \beta^* P^4 + \frac{\gamma}{6} P^6 + \frac{\varsigma}{8} P^8 + (s_{11} + s_{12})^{-1} u_m^2 - PE - \frac{1}{2} \epsilon_b E^2 \right] h \\ &\quad + (\zeta_1 - \zeta_2)P + \frac{1}{2}(\eta_1 + \eta_2)P^2.\end{aligned}\tag{5.4.9}$$

It is considered that the case for symmetric ferroelectric-electrode interfaces, and set $\zeta_1 = \zeta_2 = \zeta$ and $\eta_1 = \eta_2 = \eta$. Naturally, the built-in electric fields of Eq. (5.4.9) can also be neglected, i.e., $E_{bi} = 0$. If the two ferroelectric/electrode interfaces are different, E_{bi} must be taken into account.

5.4.2 The Evolution Equation and Stability of the Stationary States

Using Eqs. (5.4.6) to (5.4.9), the time evolution of the system is governed by the TDGL equation which can be written as (Wang and Woo, 2005)

$$\frac{\partial P}{\partial t} = -Mh \left[\alpha_0 (T - T_{c0}) P - \frac{4Q_{12}u_m}{s_{11} + s_{12}} P + \frac{2\eta}{h} P + \xi_d P + \left(\beta + \frac{4Q_{12}^2}{s_{11} + s_{12}} \right) P^3 + \gamma P^5 + \zeta P^7 \right], \quad (5.4.10)$$

where ξ_d is the coefficient of the depolarization field in Eq. (5.4.10), and can be written as terms of the thickness h using the LL model (Zheng et al., 2010).

It is known that Eq. (5.4.10) has a trivial solution $P = 0$, representing the stationary paraelectric state. To be able to transform between the paraelectric ($P = 0$) and ferroelectric ($P \neq 0$) states, the initial state has to become dynamically unstable. The dynamical stability of the initial state can be probed by applying a small perturbation Δ to the corresponding stationary solution of Eq. (5.4.10). The dynamics then follow from Eq. (5.4.10) by retaining only terms linear in Δ , which is given by

$$\begin{aligned} \frac{\partial \Delta}{\partial t} = Mh \left\{ -\alpha_0 (T - T_{c0}) \Delta + \frac{4Q_{12}u_m}{s_{11} + s_{12}} \Delta - \frac{2\eta}{h} \Delta - \xi_d \Delta \right. \\ \left. - 3 \left(\beta + \frac{4Q_{12}^2}{s_{11} + s_{12}} \right) P^2 \Delta - 5\gamma P^4 \Delta - 7\zeta P^6 \Delta \right\}. \end{aligned} \quad (5.4.11)$$

Using the method of separation of variables, the TDGL equation can be transformed into an eigenvalue problem. Thus, for the $P = 0$ state, the solution can be written as $\Delta_c = e^{\omega_c t} \Delta_0$, where ω_c is the eigenvalue, and Δ_0 the corresponding eigenfunction of the TDGL equation

$$\omega_c = M \left[-\alpha_0 (T - T_{c0}) + 4Q_{12}u_m (s_{11} + s_{12})^{-1} - \frac{2\eta}{h} - \xi_d \right]. \quad (5.4.12)$$

The $P = 0$ solution is unstable when $\omega_c > 0$ because in this case Δ increases exponentially with time. It can be seen from Eq. (5.4.12) that when the temperature T is sufficiently high, $\omega_c < 0$ and the paraelectric state is stable.

5.4.3 Curie Temperature of the Ferroelectric Capacitors or Tunnel Junctions

When T is sufficiently low, ω_c in Eq. (5.4.12) turns positive and the paraelectric state is no longer stable, since any small perturbation Δ will grow exponentially beyond all bounds. The critical condition $\omega = 0$ yields, in this case, the phase transition temperature T_c of the metal-ferroelectric-metal system on the nanoscale

$$T_c(u_m, \eta, h) = T_{c0} + \frac{4Q_{12}u_m}{\alpha_0(s_{11} + s_{12})} - \frac{2\eta}{\alpha_0 h} - \frac{\xi_d}{\alpha_0} = T_{c0} + T_M + T_S + T_D, \quad (5.4.13)$$

where the first term T_{c0} is the bulk ferroelectric Curie temperature, the second term T_M is related to the misfit of the substrate, the third term T_S , to the induced short-range surface effect, and the fourth term T_D , to the long-range depolarization field effect. It can be seen that the last two terms act against the transformation by lowering the transition temperature T_c , and that the second term may act either way, depending on the misfit strain u_m .

An example, a BaTiO₃ thin film sandwiched between SrRuO₃ electrodes epitaxially grown on SrTiO₃ substrate, is considered. The metal-ferroelectric-metal capacitor or tunnel junction is under short-circuit boundary conditions. Values of the coefficients of the free-energy expansion, electrostriction, elastic properties, background dielectric constant, dielectric constants of electrodes, screening length, coefficients relating to surface and misfit strain are known (Zheng et al., 2010). Eq. (5.4.13) shows that the phase transition temperature T_c of a symmetric metal-ferroelectric-metal system is a sum of four contributions, from the bulk Curie temperature (T_{c0}), the misfit strain (T_M), the short-range surface effect (T_S) and the long-range depolarization field effect (T_D). Fig. 5.14 plots T_{c0} , T_M , T_S and T_D as a function of the ferroelectric thin film thicknesses h . The square solid line (i.e., $T_{c0} + T_M$) indicates an h -independent phase-transition temperature of ~ 1150 K when the surface contribution T_S and depolarization field contribution T_D are neglected. From Fig. 5.14, it can also be seen that both contributions act to suppress spontaneous polarization and lower the transition temperature, with a strength that increases with decreasing FTF thickness h .

The phase-transition temperature T_c calculated as a function of the ferroelectric thin film thickness h is plotted in Fig. 5.15 with the solid black line. The most important feature is the critical thickness of $h_c \approx 2.4$ nm at 0 K, below which ferroelectricity of the ferroelectric thin film disappears permanently. This compares very well with the first-principle results of Junquera and Ghosez, which gives a critical thickness of about six unit cells of BaTiO₃, or 2.4 nm.

To study the effects of the various contributions, in Fig. 5.15 the calculations are shown (1) without the short-range surface and long-range depolarization field, i.e., putting $T_S = 0$, $T_D = 0$, (2) without the long-range depolarization field, i.e., putting $T_S \neq 0$, $T_D = 0$, and (3) without the misfit strain, i.e. putting $T_S \neq 0$, $T_D \neq 0$,

$T_M = 0$, respectively. It is clear by comparison with T_c that the misfit strain is responsible for the rise in the transition temperature as h increases. On the contrary, the surface effect and long-range depolarization field contribution are responsible for lowering the transition temperature, with a magnitude that increases with a decrease in the ferroelectric thin film thickness h . At the same time, Fig. 5.15 also shows that the critical thickness h_c is larger without the misfit strain contribution, and is lowered without the long-range depolarization field contribution. More importantly, the critical thickness disappears when all the surface and depolarization field contributions are neglected, i.e., the dash dot line in Fig. 5.15. This means that the critical thickness of a symmetric metal-ferroelectric-metal system is entirely the result of the short-range surface and long-range depolarization field effects.

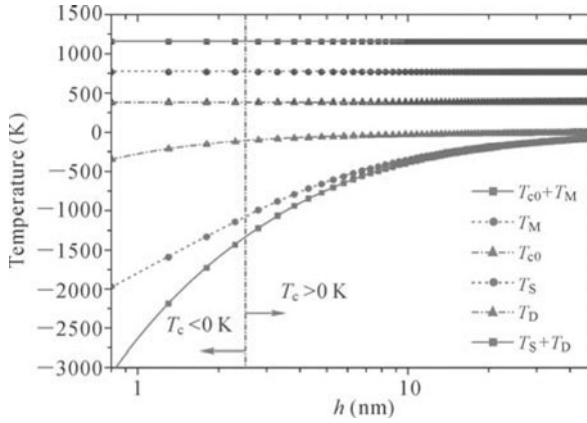


Fig. 5.14 The temperature determined by the bulk material T_{c0} , misfit strain T_M , short-range surface effect T_S , and long-range depolarization field contribution T_D versus the FTF thickness h

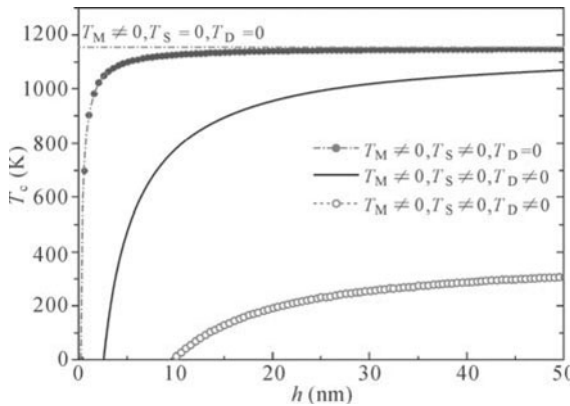


Fig. 5.15 The phase transition temperature T_c versus the ferroelectric thin film thickness h of the metal-ferroelectric-metal capacitor with short circuit boundary conditions

5.4.4 Polarization as a Function of Thickness of the Ferroelectric Capacitors or Tunnel Junctions

It is also important to directly evaluate the strength of polarization P as a function of the ferroelectric thin film thickness h . In terms of the Curie-Weiss relation, the free energy density can be rewritten as

$$\tilde{\Phi} \approx \frac{\alpha_0}{2} [T - T_c(u_m, \eta, h)] P^2 + \beta^* P^4 + \frac{\gamma}{6} P^6 + \frac{\zeta}{8} P^8, \quad (5.4.14)$$

where T_c is the phase transition temperature of the metal-ferroelectric-metal structure from Eq. (5.4.13).

At equilibrium, free energy minimization with respect to P yields the following algebraic equation,

$$\alpha_0 [T - T_c(u_m, \eta, h)] P + 4\beta^* P^3 + \gamma P^5 + \zeta P^7 = 0, \quad (5.4.15)$$

which can be solved to give

$$P^2 = \left[-\frac{A^*}{2} + \sqrt{\left(\frac{A^*}{2}\right)^2 + \left(\frac{B^*}{3}\right)^3} \right]^{1/3} + \left[-\frac{A^*}{2} - \sqrt{\left(\frac{A^*}{2}\right)^2 + \left(\frac{B^*}{3}\right)^3} \right]^{1/3} - C^*, \quad (5.4.16)$$

where A^* , B^* and C^* can be given by

$$\begin{aligned} A^* &= \frac{2\gamma^3}{27\zeta^3} - \frac{4\gamma\beta^*}{3\zeta^2} + \frac{\alpha_0 [T - T_c(u_m, \eta, h)]}{\zeta}, \\ B^* &= \frac{4\beta^*}{\zeta} - \frac{\gamma^2}{3\zeta^2}, \\ C^* &= \frac{\gamma}{3\zeta}, \end{aligned} \quad (5.4.17)$$

from which P can be expressed in terms of the ferroelectric thin film thickness, misfit strain and surface coefficient.

By solving Eqs. (5.4.15) to (5.4.17), the polarization P as a function of the FTF thickness h is plotted in Fig. 5.16a for temperatures $T = 0$ K, 300 K, 600 K. Temperature-dependent critical lengths h_c of 2.4 nm for 0 K, 3.8 nm for 300 K and 6.6 nm for 600 K can be seen, below which the metal-ferroelectric-metal capacitor or tunnel junction is paraelectric, and above which it is ferroelectric. In the ferroelectric region, the spontaneous polarization P behaves almost like a step function, increasing sharply with h and quickly saturating near the bulk polarization values. Fig. 5.16b shows the depolarization field E_d as a function of

sample geometry and temperature. E_d can be negligible when the thickness h of the ferroelectric thin film is either too large or too small. The peak depolarization field at room temperature is only 2/3 of E_d at 0 K, but is more than double the magnitude of E_d at 600 K. According to (Pertsev and Kohlstedt, 2007), the dash lines in Fig. 5.16b are also determined by temperature, and indicate the single domain critical thickness below which the single domain state in an MFMC becomes unstable because the internal depolarization field is very large.

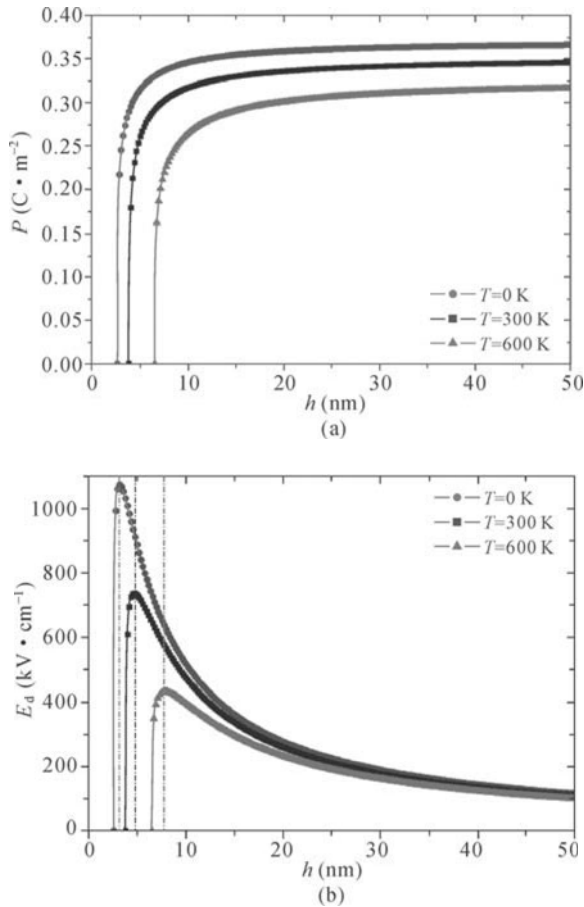


Fig. 5.16 The polarization P (a) and depolarization field (b) versus the ferroelectric thin film thickness h at 0 K, 300 K and 600 K, respectively. The dash lines show the thickness below which the single domain becomes unstable

5.4.5 Critical Thickness of the Ferroelectric Capacitors or Tunnel Junctions

From the solution in Fig. 5.16, the relationship between the critical thickness h_c and temperature of the metal-ferroelectric-metal system in nanoscale can be established. The results presented in Fig. 5.17 show a critical thickness h_c that increases with increasing temperature, caused by a suppression of the internal polarization (see Fig. 5.16a). As a result, the critical thickness h_c at 0 K is only 1/6 that at 900 K.

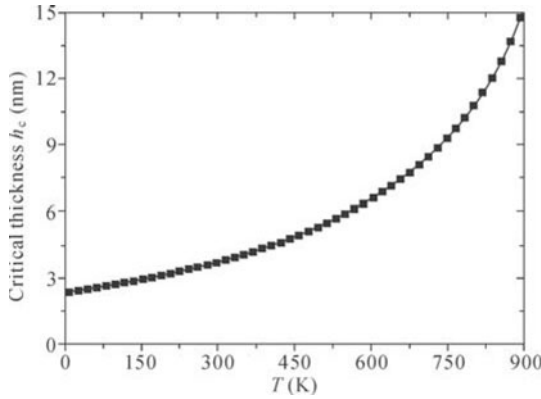


Fig. 5.17 The critical thickness h_c versus temperature changing from 0 K to 900 K of the metal-ferroelectric-metal capacitor or tunnel junction in nanoscale

The critical thickness h_c of a metal-ferroelectric-metal capacitor or tunnel junction in nanoscale is known to be related to the effects of surfaces, depolarization field and misfit strain.

Fig. 5.18 shows h_c as functions of (a) the coefficient of the surface energy η at room temperature, (b) the screening length l_s and (c) the misfit strain u_m . Monotonically increasing functional relationships are seen in all cases. In addition, we note that the critical thickness is very sensitive to the misfit strain. Indeed, in Fig. 5.18c, one can find that the critical thickness can be reduced by a factor of 6 when the misfit strain u_m changed from -0.005 to -0.04 (– means compressive).

Using Eqs. (5.4.15) to (5.4.17), the polarization P can be expressed as a function of temperature. For an FTF with thickness $h = 5$ nm, the transition temperature T_c can reach about 1100 K (see dash line in Fig. 5.19a) with the surface effect, if the long-range depolarization field is neglected. With the depolarization field only and neglecting the surface effect, the transition temperature is reduced to 510 K (see the dash line in Fig. 5.19a). Including both the surface effect and the depolarization field, T_c lowers to 460 K. Fig. 5.19b shows the polarization P as functions of temperature for various film thicknesses. It can be seen that both the polarization P and phase-transition temperature decrease with a decrease in the ferroelectric thin film thickness.

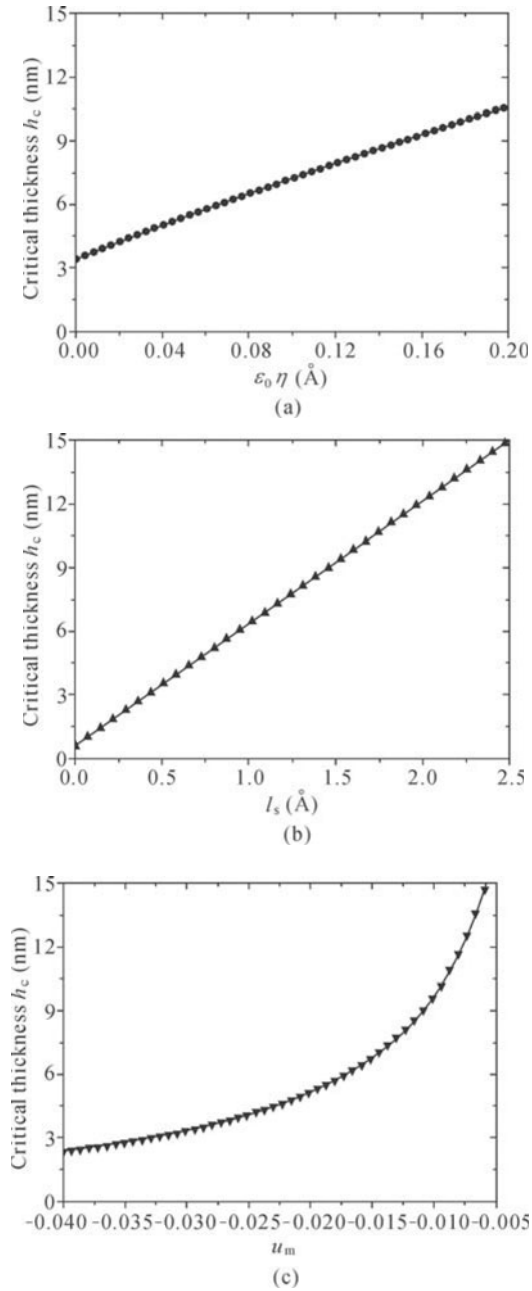


Fig. 5.18 The critical thickness h_c versus the coefficients of the surface effect η (a), the screening length l_s (b), and the misfit strain u_m (c)

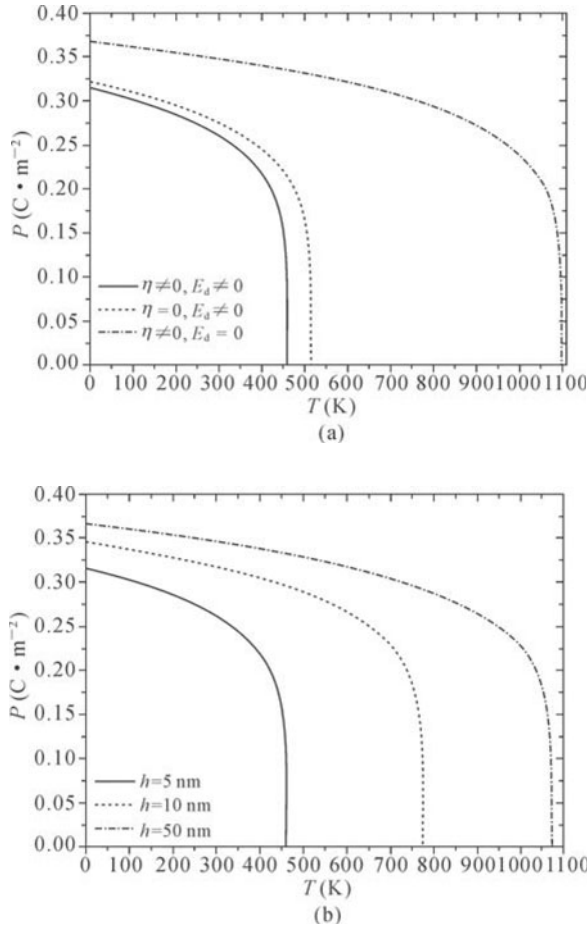


Fig. 5.19 The polarization P as a function of temperature with different conditions (a) and different thicknesses (b) of the ferroelectric thin film

5.4.6 Curie-Weiss Relation of the Ferroelectric Capacitors or Tunnel Junctions

Eq. (5.2.1) shows that an applied electric field E affects P not only extrinsically, the free-energy functional. The total susceptibility χ^T is thus the sum of a linear component χ_0 and a non-linear one χ . The former comes from the extrinsic dependence of P on E , and the latter measures the intrinsic effect of E on the dynamic properties of P , such as expressed in Eq. (5.4.16). Equivalently, χ can also be calculated via the free energy of Eq. (5.4.14)

$$\chi^{-1} = \frac{\partial^2 \tilde{\Phi}}{\partial P^2} \approx \alpha_0 [T - T_c(u_m, \eta, h)] + 12\beta^* P^2 + 5\gamma P^4 + 7\zeta P^6. \quad (5.4.18)$$

Following Wang and Woo (2005) and Zheng et al. (2009), the Curie-Weiss relation of the ferroelectric capacitor or tunnel junctions can be expressed as a function of thickness in the form

$$\chi(h) = \frac{\alpha_0^{-1} \Theta^{-1}(h)}{|T - T_c(u_m, \eta, h)|}. \quad (5.4.19)$$

Here the Curie-Weiss parameter Θ can be written as

$$\begin{aligned} \Theta(h) &= 1 + \frac{12\beta^* P^2 + 5\gamma P^4 + 7\zeta P^6}{\alpha_0 |T - T_c|} \\ &= \begin{cases} 1 + \alpha_0^{-1} (12\beta^* P^2 + 5\gamma P^4 + 7\zeta P^6) / |T - T_c|, & \text{for } T < T_c, \\ 1, & \text{for } T > T_c. \end{cases} \end{aligned} \quad (5.4.20)$$

Accordingly, a Curie-Weiss-type relation for the ferroelectric thin film of the metal-ferroelectric-metal structure can be expressed as

$$\chi^T = \chi + \chi_b = \begin{cases} \frac{\alpha_0^{-1} \Theta^{-1}(h)}{|T - T_c(u_m, \eta, h)|} + \chi_b, & \text{for } T < T_c, \\ \frac{\alpha_0^{-1}}{|T - T_c(u_m, \eta, h)|} + \chi_b, & \text{for } T > T_c. \end{cases} \quad (5.4.21)$$

This Curie-Weiss-type relation also gives the correct limit of the total susceptibility far away from its phase transition temperature T_c , i.e., $\chi^T|_{T \rightarrow \infty} = [\varepsilon_0 C_0 (T - T_c)^{-1} + \chi_b]|_{T \rightarrow \infty} = \chi_b$, which should be experimentally determined. A Curie-Weiss-type relation of the dielectric permittivity can be obtained for a BaTiO₃ ferroelectric capacitor or tunnel junction with different thicknesses of the ferroelectric thin film as shown in Fig. 5.20.

The results of Figs. 5.19 and 5.20 clearly show that the phase transition temperature, polarization and susceptibility are all very sensitive to surface and depolarization field effects, but are all controllable by varying the geometry of the ferroelectric thin film. In the evaluation of the normalized susceptibility $\alpha_0^{-1} \Theta^{-1} / (T - T_c)$, the transition temperature T_c is taken from result of Fig. 5.15. It can be seen that the Curie-Weiss relation is obviously shifted by the surface and depolarization field effects (see Fig. 5.20a). The Curie-Weiss relation can also be calculated as a function of the FTF thickness, as shown in Fig. 5.20b.

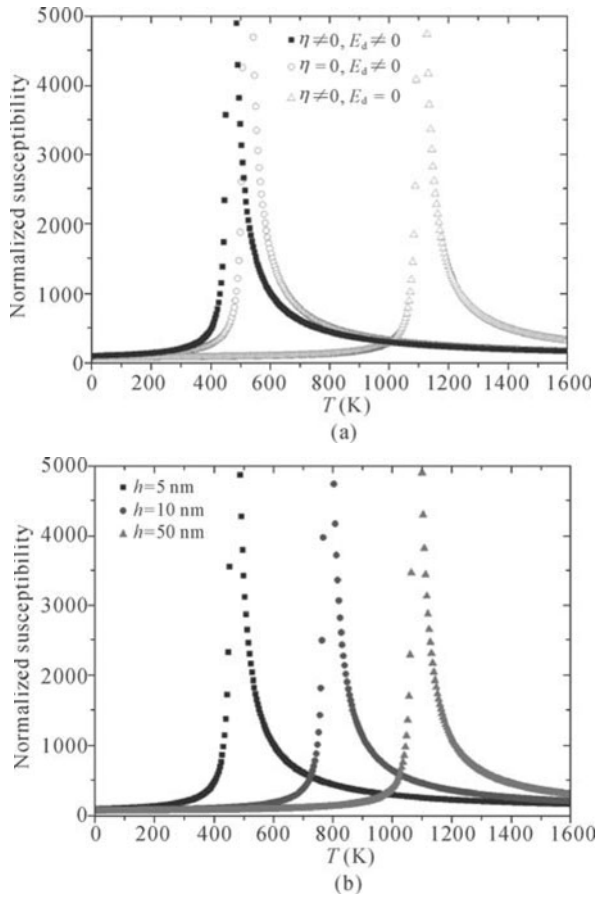


Fig. 5.20 The Curie-Weiss-type relation with different conditions (a) and different ferroelectric thin film thicknesses (b)

5.5 Ferroelectric Superlattices in Nanoscale

Artificially structured ferroelectrics are experiencing a rapid surge of research interest, driven by the demand for new functional materials that provide new opportunities for novel applications. Studies of ferroelectric bilayers, multilayers and superlattices, such as $\text{KNbO}_3/\text{KTaO}_3$, $\text{PbTiO}_3/\text{SrTiO}_3$, $\text{BaTiO}_3/\text{SrTiO}_3$, $\text{PbTiO}_3/\text{PbZrO}_3$, and $\text{PbTiO}_3/\text{BaTiO}_3$, have revealed that properties of ferroelectric multilayers can be controlled by adjusting structural periodicity, misfit strain, boundary conditions, electrostatic interactions etc. (Qu et al., 1997; Shen and Ma, 2001; Zhou and Shin, 2006; Li et al., 2007a; 2007b; Tenne et al., 2006; Ahn et al., 2004; Dawber et al., 2005a; 2007).

In this section, the critical properties, such as the phase transition temperature, polarization and Curie-Weiss-type relation, etc., of the ferroelectric superlattices are discussed as a function of the relative thickness. The results will be compared and discussed in relation to experimental observations and *ab initio* calculations.

5.5.1 The Free Energy Functional of Ferroelectric Superlattices

In what follows, the phase transition characteristics of a ferroelectric superlattice (FS) structure are considered, which is made from multilayer thin-films (MTF) sandwiched between electrodes under short-circuit boundary conditions and grown on a substrate (Fig. 5.21) (Zheng and Woo, 2009a).

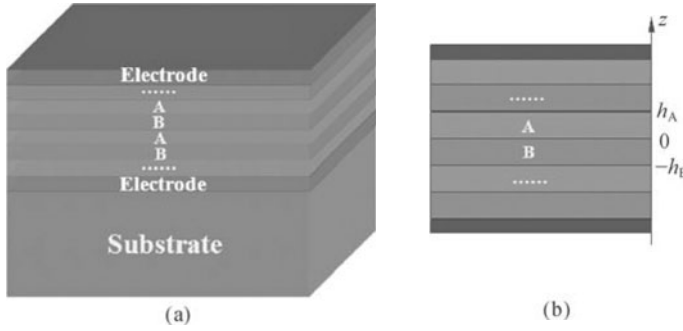


Fig. 5.21 Schematic structures of an epitaxial superlattice on a substrate (a) and the single-period superlattices A and B (b)

The MTF is made of alternating layers of different materials A and B of thicknesses h_A and h_B , respectively, with a total thickness that is large compared with the thickness L of a single period of the superlattice (i.e., $h_A + h_B$) (Fig. 5.21b). Here the periodic boundary conditions are used. In most superlattices grown on compressive substrate (e.g., $\text{BaTiO}_3/\text{SrTiO}_3$, $\text{PbTiO}_3/\text{SrTiO}_3$ and $\text{BaTiO}_3/\text{PbTiO}_3$ on SrTiO_3), all vector fields in the MTF are directed in the z -direction. Accordingly, axial symmetry is assumed in this work to simplify the formulation and computation.

Based on the SPOP theory, the depolarization fields E_{di} ($i = A, B$) in layers i are related to P_i in the absence of E_{ext} , and can be obtained using the Maxwell equations as

$$\begin{aligned} E_{dA} &= -\frac{1}{\epsilon_{bA}}(P_A - \langle P \rangle) = -\frac{1}{\epsilon_{bA}}\left(P_A - \frac{1}{L} \int_{-h_B}^{h_A} P dz\right), \\ E_{dB} &= -\frac{1}{\epsilon_{bB}}(P_B - \langle P \rangle) = -\frac{1}{\epsilon_{bB}}\left(P_B - \frac{1}{L} \int_{-h_B}^{h_A} P dz\right), \end{aligned} \quad (5.5.1)$$

where ε_{bi} are the background dielectric constants of layers i .

The total free energy per unit area of a single-period of the superlattice is defined as $F = F_A + F_B$, where F_A and F_B are the free energy per unit area of the layers A and B, respectively. For each layer, the free energy can be expressed as the sum of F_p , the energy contributions from the polarization, F_s , the elastic energy, F_C , the energy of the electrostrictive coupling between polarization and strain, F_D , the energy of the depolarization field and F_i , the energy due to the superlattice interfaces. The free-energy formulation may be generalized accordingly. In addition, the Landau-Ginzburg expansion can be used to take into account the spatial dependence of P . Thus, for a thick compressive substrate, F_A and F_B can be approximately expressed as

$$\begin{aligned}
 F_A = & \int_0^{h_A} \left[\alpha_A^* P_A^2 + \beta_A^* P_A^4 + \frac{\gamma_A}{6} P_A^6 + \frac{D_A}{2} \left(\frac{\partial P_A}{\partial z} \right)^2 + \frac{1}{2\varepsilon_{bA}} \left(P_A^2 - \langle P \rangle^2 \right) + \frac{u_{mA}^2}{s_{11A} + s_{12A}} \right] dz \\
 & + \frac{1}{2} D_A \delta_A^{-1} [P_A^2(0) + P_A^2(h_A)] - \frac{1}{2} \xi [P_A(0) \cdot P_B(0) + P_A(h_A) \cdot P_B(-h_B)], \\
 F_B = & \int_{-h_B}^0 \left[\alpha_B^* P_B^2 + \beta_B^* P_B^4 + \frac{\gamma_B}{6} P_B^6 + \frac{D_B}{2} \left(\frac{\partial P_B}{\partial z} \right)^2 + \frac{1}{2\varepsilon_{bB}} \left(P_B^2 - \langle P \rangle^2 \right) + \frac{u_{mB}^2}{s_{11B} + s_{12B}} \right] dz \\
 & + \frac{1}{2} D_B \delta_B^{-1} [P_B^2(0) + P_B^2(-h_B)] - \frac{1}{2} \xi [P_B(0) \cdot P_A(0) + P_B(-h_B) \cdot P_A(h_A)],
 \end{aligned} \tag{5.5.2}$$

where the renormalized Landau coefficients α_A^* , β_A^* , α_B^* and β_B^* can be expressed as

$$\begin{aligned}
 \alpha_A^* &= \frac{\alpha_{0A}}{2} (T - T_{c0A}) - \frac{2Q_{12A} u_{mA}}{s_{11A} + s_{12A}}, \quad \beta_A^* = \frac{\beta_A}{4} + \frac{Q_{12A}^2}{s_{11A} + s_{12A}}, \\
 \alpha_B^* &= \frac{\alpha_{0B}}{2} (T - T_{c0B}) - \frac{2Q_{12B} u_{mB}}{s_{11B} + s_{12B}}, \quad \beta_B^* = \frac{\beta_B}{4} + \frac{Q_{12B}^2}{s_{11B} + s_{12B}},
 \end{aligned} \tag{5.5.3}$$

where α_{0i} , β_i and γ_i are the expansion coefficients of the free energy of layer i . $\alpha_{0A} = 1/(\varepsilon_0 C_A)$ and $\alpha_{0B} = 1/(\varepsilon_0 C_B)$, and C_i is the Curie constant of layer i . T_{c0i} is the Curie-Weiss temperature of their bulk materials, Q_{12i} are the electrostriction tensors, s_{11i} and s_{12i} are components of the elastic compliance tensor, u_{mi} are the misfit strains between layers i and substrate. The δ_i are also called extrapolation lengths, which describe surface relaxation effects of layers i . Together with the short-range coupling strength ξ between layer A and layer B, the change to the free energy due to the presence of the interface can be described. Note that the depolarization field and free energy of the TPOP approach have the same form as Eqs. (5.5.2) and (5.5.3), but with ε_{bi} and P_i replaced by the permittivity of free-space ε_0 and the total polarization P_i^T , respectively. In TPOP calculations, the

absolute dominance of the depolarization field will lead to the quenching of critical and any other properties in the presence of PTO/STO superlattice and interfaces.

The dynamic behavior of P_i is governed by the E-L equation associated with F in Eq. (5.5.2), which in the present case takes the form of the time-dependent Landau-Ginzburg (TDLG) equation or Landau-Khalatnikov (LK) equation,

$$\begin{aligned}\frac{\partial P_A(z,t)}{\partial t} &= -M_A \frac{\delta F}{\delta P_A(z,t)} \\ &= -M_A \left[2\alpha_A^* P_A + 4\beta_A^* P_A^3 + \gamma_A P_A^5 - D_A \frac{\partial^2 P_A}{\partial z^2} + \frac{P_A}{\epsilon_{bA}} - \frac{1}{\epsilon_{bA}} \langle P \rangle \right], \\ \frac{\partial P_B(z,t)}{\partial t} &= -M_B \frac{\delta F}{\delta P_B(z,t)} \\ &= -M_B \left[2\alpha_B^* P_B + 4\beta_B^* P_B^3 + \gamma_B P_B^5 - D_B \frac{\partial^2 P_B}{\partial z^2} + \frac{P_B}{\epsilon_{bB}} - \frac{1}{\epsilon_{bB}} \langle P \rangle \right].\end{aligned}\quad (5.5.4)$$

Here M_A and M_B are the kinetic coefficients that describe the time delay in evolution of the polarization fields.

Following recent work, the interface between A and B is assumed to be coherent. The interfacial effects approximately represented in Eq. (5.5.2) can be replaced in (5.5.4) with the following boundary conditions at $z = -h_B$, 0 and h_A ,

$$\begin{aligned}\left. \frac{\partial P_B(z,t)}{\partial z} \right|_{z=-h_B} &= \delta_B^{-1} P_B(-h_B, t) - D_B^{-1} \xi P_A(h_A, t), \\ \left. \frac{\partial P_B(z,t)}{\partial z} \right|_{z=0^-} &= -\delta_B^{-1} P_B(0, t) + D_B^{-1} \xi P_A(0, t), \\ \left. \frac{\partial P_A(z,t)}{\partial z} \right|_{z=0^+} &= \delta_A^{-1} P_A(0, t) - D_A^{-1} \xi P_B(0, t), \\ \left. \frac{\partial P_A(z,t)}{\partial z} \right|_{z=h_A} &= -\delta_A^{-1} P_A(h_A, t) + D_A^{-1} \xi P_B(-h_B, t).\end{aligned}\quad (5.5.5)$$

Analytic expressions of the critical properties have been derived for the Curie temperature and critical thickness of single ferroelectric thin films using the TDLG (Wang and Woo, 2005), with results showing that the phase transition temperatures are sensitive to the film thickness, surface boundary conditions and depolarization field. However, for the MTF considered here, the mathematical analysis involved would be much more complex. In the ferroelectric regime, an initially local polarized region anywhere in the MTF will spread and grow into a stable polarization configuration over the whole region across the interfaces. On the other hand, in the paraelectric region, the same initially local polarized region anywhere in the MTF is unstable and will shrink away. In this way, the phase transition temperature, stable polarization and other properties of the superlattice

can be determined. At the same time, based on Wang and Woo (2005), stability analysis can be used to calculate the susceptibility and Curie-Weiss-type relation of the ferroelectric superlattices.

In the following, the specific example of a $\text{PbTiO}_3/\text{SrTiO}_3$ superlattice (PTO/STO) grown on an STO substrate, which has been successfully fabricated and characterized recently, is considered. Material constants, such as the electrostrictive and elastic coefficients of PTO and STO, are known. The susceptibility and dielectric constants of background materials, the extrapolation lengths, that the interfacial short-range coupling coefficients can be approximately obtained from previous works. Note that entirely accurate parameters for different ferroelectric materials, such as the dielectric constant of background materials and extrapolation lengths, need future work based on experiments, first principle calculations or molecular dynamics.

5.5.2 The Phase Transition Temperature of PTO/STO Superlattice

The transition temperature of ferroelectric ultra-thin superlattices is known to depend significantly on the thicknesses of the component layers. In the present section, the ferro/paraelectric transition temperature T_c of the FS is calculated as a function of the thicknesses of the STO and PTO layers. Note that the Curie temperatures of STO and PTO are 35.5 K and 763 K, respectively. In Fig. 5.22, the calculated T_c is plotted against the relative thickness of the PTO layer in an FS in which the total thickness of the PTO and STO layers is kept constant at 10 nm. It can be seen that T_c increases with the PTO thickness, and decreases with an increase in STO thickness. It can also be seen that a very significant shift of up to a thousand degrees of T_c is achievable by varying the relative thicknesses of STO and PTO.

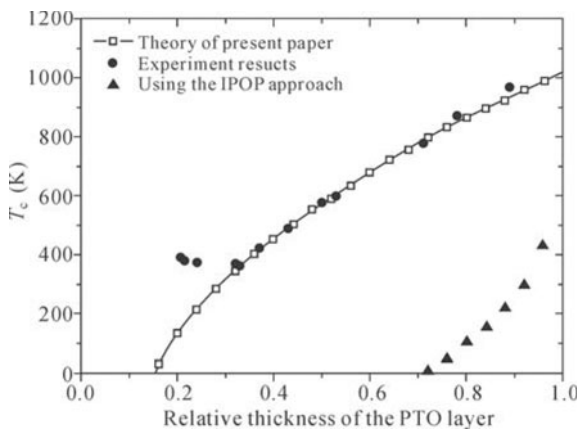


Fig. 5.22 The phase transition temperature as a function of the relative thickness of the PTO layer

Fig. 5.22 shows that a minimum relative thickness (MRT) of the PTO layer exists, which in the present case is $\sim 17\%$, below which ferroelectricity disappears permanently. For comparison, the corresponding experimental measurements using X-ray diffraction are also shown (Dawber et al., 2007). It can be seen that there is excellent agreement between the experiment and calculated results. This agreement is even more remarkable when we consider that there is not any fitting parameter in the calculation. Note that this dependence on the PTO relative fraction can provide a convenient means of control of T_c via the relative thicknesses. In addition, one may be interested in comparing the foregoing results with those obtained from a similar treatment based on the TPOP approach. The results are shown in Fig. 5.22 as triangles, which can be seen to be shifted downwards by $\sim 500^\circ\text{C}$ from experimental results and from results calculated following the SPOP approach. At the same time, the MRT of the PTO layer obtained is over 400% larger than that obtained from the SPOP approach. As explained earlier, these results represent a significant under-estimation of the total polarization due to an insufficient screening of the depolarization effects in the TPOP approach.

Indeed, even if the deviation from uniformity is as small as $\sim 0.4\%$ of the polarization in bulk PbTiO_3 at room temperature ($\sim 0.7 \text{ C}\cdot\text{m}^{-2}$) under short-circuit boundary conditions, the total free energy is still dominated by the contribution from the depolarization field, which is about two orders of magnitude larger than the sum of all other contributions in the total free energy. This dominance of the depolarization field may cause a total quenching of the properties of individual thin film layers. Similar results are obtained in other theoretical calculations on $\text{BaTiO}_3/\text{SrTiO}_3$ bilayers using the TPOP approach, where the MRT obtained for BTO is over 85%. Indeed, we have also repeated such calculations and confirm the results. The TPOP results obviously deviate significantly from the experimental ones of Tenne et al. (2006), where the ferroelectricity of a $\text{BaTiO}_3/\text{SrTiO}_3$ superlattice is measurable even when the relative BTO thickness is below 30%.

5.5.3 Polarization and Critical Thickness of PTO/STO Superlattice

The variation in the critical properties of the FS as a function of the relative thicknesses of the STO and PTO layers suggests that it is an effect coming from the influence of the interface and depolarization field. Similar effects are also expected to affect the total polarization P^T , calculated as the sum of the local spontaneous polarization P_i and that induced by the depolarization field, i.e., $P_i^E = \chi_{bi} E_{di}$.

The behavior of P^T at 0 K is plotted in Fig. 5.23 as a function of n_p/n_s , where n_p and n_s are the thicknesses of the PTO and STO layers, respectively, measured in terms of the number of atomic planes. To facilitate comparison with the results of

ab initio calculations, n_s is kept constant at a value of 3, and n_p varies between 1 and 48. The excellent agreement between the two results shown in Fig. 5.23 further reinforces the credibility of the SPOP approach. Results in Fig. 5.23 also show the increasing strength of the interfacial depolarization charges as the PTO thickness decreases. Below a critical ratio of n_p/n_s the FS cannot sustain any ferroelectric characteristics. Consistent with the results of the *ab initio* calculations, the critical thickness of the PTO layer is $h_A^c \approx 1.2$ nm (Dawber et al., 2005b). For completeness, Fig. 5.23 also presents results obtained using the TPOP approach as triangles. Comparison with results from the *ab initio* calculations of Teen et al. (2005) shows substantial disagreement. They showed that the ferroelectricity of FS disappeared at a relative thickness of the STO layer of about 15% (instead of 50%), and the corresponding critical thickness calculated using TPOP was much larger, at $h_A^c \approx 6.8$ nm (instead of $h_A^c \approx 1.2$ nm). This is consistent with a significantly overestimated depolarization field without inadequate screening. The modification adopted by using the electrostatic model helps push the results into better agreement in the regime of small n_p/n_s . However, the polarization does not seem to help when the FS is dominated by the PTO layer.

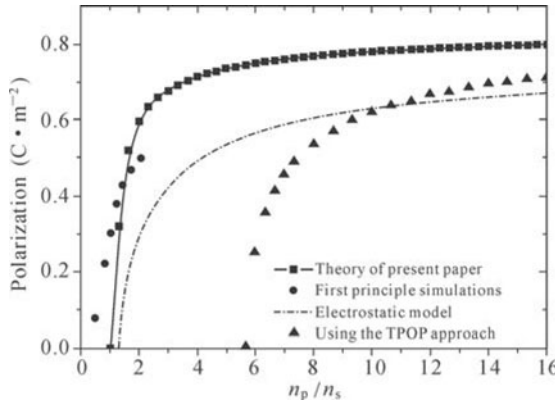


Fig. 5.23 The average total polarization as a function of the ratio n_p/n_s of PTO/STO superlattices

The average spontaneous polarization at 300 K as a function of the relative thickness of the PTO layer for a fixed $L = 10$ nm is shown in Fig. 5.24a. The data points are derived from those in the inset in Fig. 5.24a, which shows the through-thickness (i.e., as function of z) values of P_A and P_B for various relative thicknesses of the PTO versus STO layers. It is clear that as the PTO layer gets thinner, the spontaneous polarizations decrease. For PTO in the PTO/STO superlattice, the MRT at 300 K is $\sim 28\%$ (Zheng and Woo, 2009a).

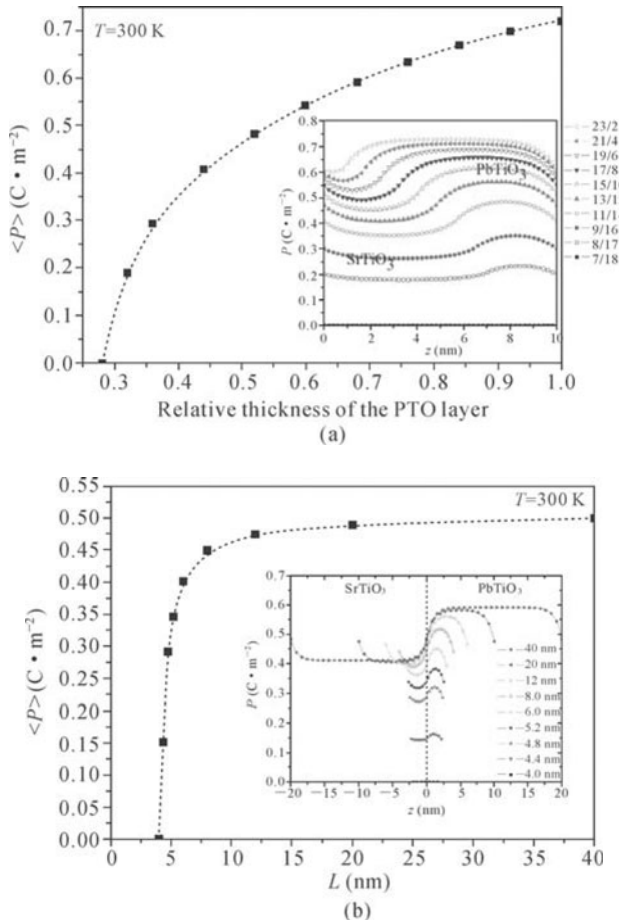


Fig. 5.24 (a) The relation between $\langle P \rangle$ and the relative thickness of the PTO layer for a fixed L of 10 nm. The inset shows the distribution of P_A and P_B in a single-period superlattice with the relative thickness of the PTO layer from 72% (shown as “ $\cdots$ ”) to 8% (shown as “ $\blacksquare$ ”). (b) The relation between $\langle P \rangle$ and L (e.g., squares on a dashed line “ $-\blacksquare-$ ”). The inset shows the distribution of P_A and P_B in a single-period superlattice

To consider the effects of the interfacial and depolarization field effects, we also plot the sum $\langle P \rangle$ of the average spontaneous polarizations $\langle P_A \rangle$ and $\langle P_B \rangle$ in individual STO and PTO layers as a function of the single-period thickness L ($= h_A + h_B$) at 300 K for equal thicknesses of the two layers in Fig. 5.24b. Here $\langle P_A \rangle$ and $\langle P_B \rangle$ are expressed as $\int_0^{h_A} P_A h_A^{-1} dz$ and $\int_{-h_B}^0 P_B h_B^{-1} dz$, where the through-thickness spontaneous polarizations P_A and P_B are shown in the inset. It can be seen that $\langle P \rangle$, P_A and P_B , all decrease with decreasing L as the interface effects

grow stronger with decreasing L .

If the depolarization field is absent (i.e., if $E_{dA} = E_{dB} = 0$), the average spontaneous polarization $\langle P_B \rangle$ of the STO layer is near zero at room temperature. Noting that bulk PTO and STO are respectively in ferroelectric ($P_A \neq 0$) and paraelectric ($P_B = 0$) states at room temperature, the increase in P_B and the decrease in P_A shown in Fig. 5.24b are mainly the result of the depolarization charges induced by the spontaneous polarization of the PTO layer under the short-circuit boundary conditions. Based on results in the inset of Fig. 5.24b, the calculated relation between $\langle P \rangle$ and L at 300 K is shown as squares on a dashed line in Fig. 5.24b. It can be seen that there is a critical value of $L_c \approx 4$ nm, below which the spontaneous polarization disappears permanently.

5.5.4 The Curie-Weiss-Type Relation of PTO/STO Superlattice

As mentioned, the application of the applied electric field E affects the free-energy functional and produces an indirect effect on the function P when the variational minimum is sought. Accordingly, the dielectric properties of FS are governed by a Curie-Weiss-type relation that can be expressed as

$$\langle \chi_T \rangle = \langle \chi \rangle + \langle \chi_b \rangle = \begin{cases} \frac{\epsilon_0 \langle C \rangle}{2|T - T_c|} + \langle \chi_b \rangle, & \text{for } T < T_c, \\ \frac{\epsilon_0 \langle C \rangle}{|T - T_c|} + \langle \chi_b \rangle, & \text{for } T > T_c. \end{cases} \quad (5.5.6)$$

Here the average Curie constant $\langle C \rangle$ and background susceptibility $\langle \chi_b \rangle$ of the FS are

$$\langle C \rangle = \frac{h_A C_A}{h_A + h_B} + \frac{h_B C_B}{h_A + h_B}, \quad \langle \chi_b \rangle = \frac{h_A \chi_{bA}}{h_A + h_B} + \frac{h_B \chi_{bB}}{h_A + h_B}. \quad (5.5.7)$$

$\langle \chi_T \rangle$ in the PTO layer for different temperatures is plotted in Fig. 5.25 against $T - T_c$ where T_c is taken from Fig. 5.22. It can be seen clearly that the susceptibility is very well controllable via the relative thicknesses of the layers. Eq. (5.5.6) also gives the correct limit of the susceptibility far away from T_c , i.e., $\langle \chi_T \rangle|_{T \rightarrow \infty} = \langle \chi_b \rangle$, which should be investigated and confirmed by experiments.

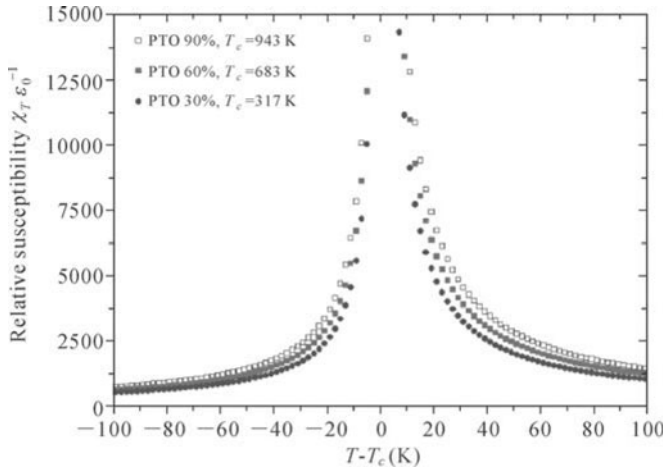


Fig. 5.25 The Curie-Weiss-type relation of PTO/STO superlattice with different relative thicknesses of the PTO layer for a fixed $L = 10$ nm

5.6 Ferroelectric Nanowires and Nanotubes

Theory and experiments on nanoparticles (NP), nanowires (NW), nanorods (NR), nanodisks (ND) and nanotubes (NT) all indicate strong surface effects on their polarization. This is expected because the large surface to volume ratio in such cases naturally leads to strong surface effects (Uchino et al., 1989; Ma et al., 1998; Morozovska, 2006; Liu et al., 2006; Alexe et al., 2006; Jang et al., 2005; Luo et al., 2003; Zhou et al., 2007; Naumov and Glinchuk, 2003). Indeed, the tendency to minimize the surface energy introduces a large intrinsic surface stress, i.e., surface tension in samples of nano-dimension, and is one of the most important surface effects considered.

In this section, the transition temperature and polarization of ferroelectric NWs and NTs are derived and analyzed, paying particular attention to the effects of the surface tension and surface relaxation. The transition temperature and polarization as a function of wire and tube dimensions is derived through the actions of the surface tension, polarization gradient coefficient, extrapolation length and electrostriction coefficient. The implications of the results are considered and discussed.

5.6.1 Surface Tension of Ferroelectric Nanowires and Nanotubes

Nanocrystals generally have a large surface to volume ratio. The tendency to minimize the surface energy density is often sufficiently large as to significantly

affect the properties of the nanocrystal. Experimentally, Ma et al. and Uchino et al. noted that the effect of surface tension on a nanocrystal was analogous to hydrostatic pressure on a bulk single crystal. For ferroelectric NW, NR or NPs, surface effects expressed in terms of the extrapolation length can change the polarization near surfaces and associated properties. At the same time, the surface tension produces a radial pressure, which is expected to affect properties. Based on the thermodynamic model, Morozovska et al. found that the surface tension in finite-size NR/NW/ND plays a role similar to an epitaxial stress in thin films (Morozovska et al., 2006; Zheng et al., 2010a). The uniform radial stress due to the surface tension μ/R in an NR, with surface energy density μ and radius R , compresses the NR in the transverse direction and stretches it along the polar axis z . Recently, Zhou et al. (2007) also illustrated the effect of the surface tension and internal compressive stress on the properties of ferroelectric NW. By a similar reasoning, surface tension in an NT should also experience similar effects.

The linear elastic problem of a thick-wall tube under radial pressures p_a and p_b at the inner and outer surface (see Fig. 5.26) has been solved by Lamé (Timoshenko and Goodier, 1970). The resulting stress field can be expressed in cylindrical coordinates as follows:

$$\begin{aligned}\sigma_{rr}(r) &= \frac{a^2}{b^2 - a^2} \left(1 - \frac{b^2}{r^2}\right) p_a - \frac{b^2}{b^2 - a^2} \left(1 - \frac{a^2}{r^2}\right) p_b, \\ \sigma_{\varphi\varphi}(r) &= \frac{a^2}{b^2 - a^2} \left(1 + \frac{b^2}{r^2}\right) p_a - \frac{b^2}{b^2 - a^2} \left(1 + \frac{a^2}{r^2}\right) p_b, \\ \sigma_{r\varphi} &= 0, \sigma_{rz} = 0, \sigma_{zz} = 0, \sigma_{z\varphi} = 0.\end{aligned}\quad (5.6.1)$$

Here, for the NW, it is noted that the inner radius a is zero. Eq. (5.6.1) reduces to the stress field of an NW (i.e., $a = 0$ and $b = R$, and $p = -\mu/R$), which has been discussed (Zheng et al., 2007). For the NT, the radial stresses p_a and p_b in Eq. (5.6.1) can be approximately expressed in terms of the surface energy densities μ_a and μ_b , i.e., $p_a = -\mu_a/a$ and $p_b = \mu_b/a$.

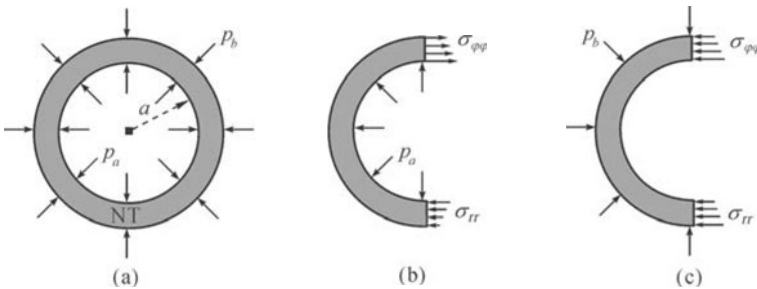


Fig. 5.26 Schematics of the stress field in the nano-tube problem under applied inner and outer radial pressures

5.6.2 Size and Surface Effects on Ferroelectric Nanowires

5.6.2.1 Contribution of the Surface Tension on the Free Energy of Ferroelectric Nanowires

Generally speaking, the free energy density expansion of ferroelectrics on the spontaneous polarization $\mathbf{P} = (0, 0, P)$ and stresses σ_i in the Voigt notation can be given by

$$f = \frac{\alpha_0}{2}(T - T_{c0})P^2 + \frac{\beta}{4}P^4 + \frac{\gamma}{6}P^6 - Q_{11}\sigma_{33}P^2 - Q_{12}(\sigma_1 + \sigma_2)P^2 - \frac{1}{2}s_{11}(\sigma_1^2 + \sigma_2^2 + \sigma_3^2) - s_{12}(\sigma_1\sigma_2 + \sigma_2\sigma_3 + \sigma_1\sigma_3) - \frac{1}{2}s_{44}(\sigma_4^2 + \sigma_5^2 + \sigma_6^2), \quad (5.6.2)$$

where α_0 , β , γ are the expansion coefficients of the Landau free energy, T_{c0} is the Curie-Weiss temperature of the bulk material, and Q_{ij} is the electrostriction tensor.

The conditions of mechanical equilibrium on the surface and in bulk of the cylindrical solid body have the form in the cylindrical coordinates as

$$\begin{aligned} \sigma_{rr}|_{r=R} = -p, \quad \sigma_{r\varphi}|_{r=R} = 0, \quad \sigma_{rz}|_{r=R} = 0, \\ \sigma_{zz}|_{z=\pm h/2} = 0, \quad \sigma_{zr}|_{z=\pm h/2} = 0, \quad \sigma_{z\varphi}|_{z=\pm h/2} = 0, \end{aligned} \quad (5.6.3)$$

$$\begin{cases} \frac{\partial \sigma_{zz}}{\partial z} + \frac{\partial \sigma_{zr}}{\partial r} + \frac{\sigma_{zr}}{r} + \frac{1}{r} \frac{\partial \sigma_{z\varphi}}{\partial \varphi} = 0, \\ \frac{\partial \sigma_{rr}}{\partial r} + \frac{\sigma_{rr} - \sigma_{\varphi\varphi}}{r} + \frac{1}{r} \frac{\partial \sigma_{r\varphi}}{\partial \varphi} + \frac{\partial \sigma_{zr}}{\partial z} = 0, \\ \frac{1}{r} \frac{\partial \sigma_{\varphi\varphi}}{\partial \varphi} + \frac{\partial \sigma_{z\varphi}}{\partial z} + \frac{\partial \sigma_{r\varphi}}{\partial r} + 2 \frac{\partial \sigma_{r\varphi}}{\partial r} = 0. \end{cases} \quad (5.6.4)$$

From Eqs. (5.6.3) and (5.6.4), the following relationship can be given as

$$\sigma_1 = \sigma_2 = -p, \quad \sigma_3 = \sigma_4 = \sigma_5 = \sigma_6 = 0. \quad (5.6.5)$$

The tensor components in Cartesian coordinates can be found from relations,

$$\begin{pmatrix} \sigma_{xx} & \sigma_{xy} & \sigma_{xz} \\ \sigma_{yx} & \sigma_{yy} & \sigma_{yz} \\ \sigma_{zx} & \sigma_{zy} & \sigma_{zz} \end{pmatrix} = \begin{pmatrix} \cos \varphi & -\sin \varphi & 0 \\ \sin \varphi & \cos \varphi & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} \sigma_{rr} & \sigma_{r\varphi} & \sigma_{rz} \\ \sigma_{\varphi r} & \sigma_{\varphi\varphi} & \sigma_{\varphi z} \\ \sigma_{zr} & \sigma_{z\varphi} & \sigma_{zz} \end{pmatrix} \cdot \begin{pmatrix} \cos \varphi & \sin \varphi & 0 \\ -\sin \varphi & \cos \varphi & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (5.6.6)$$

Allowing for Eq. (5.6.5), Eq. (5.6.6) leads to

$$\begin{cases} \sigma_{xx} = \cos^2 \varphi \sigma_{rr} + \sin^2 \varphi \sigma_{\varphi\varphi} = -\frac{\mu}{R}, \\ \sigma_{yy} = \sin^2 \varphi \sigma_{rr} + \cos^2 \varphi \sigma_{\varphi\varphi} = -\frac{\mu}{R}, \\ \sigma_{xy} = \sin \varphi \cos \varphi (\sigma_{rr} - \sigma_{\varphi\varphi}) = 0, \\ \sigma_{xz} = \sigma_{yz} = \sigma_{zz} = 0. \end{cases} \quad (5.6.7)$$

Stress components in Eq. (5.6.4) can be given by (Morozovska et al., 2006)

$$\sigma_1 = \sigma_2 = -\frac{\mu}{R}, \quad \sigma_{33} = \sigma_{23} = \sigma_{13} = \sigma_{12} = 0. \quad (5.6.8)$$

5.6.2.2 Free Energy of Ferroelectric Nanowires with Size and Surface Effects

For an FNW, the surface effect of FNW is in terms of the extrapolation length and can change the polarization near surfaces and associated properties, which has been investigated in Naumov et al. (2005). Meanwhile, the surface tension produces an effective radial pressure (Fig. 5.27) along the radial direction which would affect the Curie temperature, polarization and dielectric constant of FNWs, which is similar to an effective compressive stress along the longitudinal polar axis. Note that the depolarization field E_d along the polar axial direction can be neglected because the NW is infinite (Zheng et al., 2008a; 2008b).

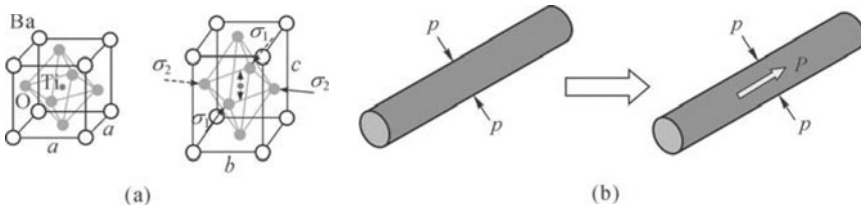


Fig. 5.27 (a) Effect of radial effective surface tension on lattice of BaTiO₃ nanowire; (b) Ferroelectric nanowire with the effective radial pressures p

In fact, the surface tension expression Eq. (5.6.8) of FNW is inapplicable when $R \rightarrow 0$ because the magnitude of the stresses would have exceeded the validity of elasticity theory and plastic yielding would have occurred. For the FNW, by assuming a simple elastic-plastic model with a yield stress σ_Y , an effective stress due to the surface tension should be defined as

$$\sigma_{\text{eff}} = \begin{cases} -\mu/R, & \text{for } R > R_Y, \\ -\sigma_Y, & \text{for } R \leq R_Y, \end{cases} \quad (5.6.9)$$

where R_Y is the radius of the NW below which plastic yield occurs.

Based on Landau-Ginzburg-Devonshire theory, the total free energy of a ferroelectric NW can be expressed as $F = F_{\text{bulk}} + F_{\text{grad}} + F_{\text{Ele}} + F_s$, where F_{bulk} , F_{grad} , F_{Ele} and F_s are the bulk energy, gradient energy, the electric field and surface energy, respectively. According to Morozovska et al. (2006) and Zheng et al. (2008), the total free energy of a ferroelectric nanowire can be expressed as

$$F = 2\pi \int_0^h dz \int_0^R r dr \left\{ \frac{\alpha_0}{2} (T - T_{c0}) P^2(r, z) - 2Q_{12} \sigma_{\text{eff}} P^2(r, z) + \frac{\beta}{4} P^4(r, z) + \frac{\gamma}{6} P^6(r, z) + \frac{D}{2} [\nabla P(r, z)]^2 - (s_{11} + s_{12}) \sigma_{\text{eff}}^2 \right\} + F_s, \quad (5.6.10)$$

where the effective uniform radial stress due to the surface tension in an NW with surface energy density μ and radius R can compress the FNW in the transverse direction and stretch it along the polar axis z . The last term F_s accounts for the near-surface depolarizing effect contributions to the total free energy. For infinite FNWs, Eq. (5.6.10) only takes into account contributions from sidewall of the FNW. F_s can be expressed as $F_s = D \int_{-h/2}^{h/2} \frac{2\pi R}{\delta} P^2(r = R, z) dz$, where δ are the extrapolation lengths at the sidewall surface of NWs.

5.6.2.3 Stability of the Ferroelectric State and Transition Characteristics

The time evolution of the system is governed by the time-dependent Ginzburg-Landau equation. Using Eqs. (5.6.9) and (5.6.10), this can be written as

$$\begin{aligned} \frac{\partial P}{\partial t} &= -M \frac{\delta F}{\delta P} \\ &= -M \left\{ \alpha_\sigma(T, R) P(r) + \beta P^3(r) + \gamma P^5(r) - D \frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial P(r)}{\partial r} \right) \right\}, \end{aligned} \quad (5.6.11)$$

where the renormalized coefficient α_σ can be given by $\alpha_\sigma = \alpha_0(T - T_{c0}) + 4Q_{12}\sigma_{\text{eff}}$. The surface term in Eq. (5.6.10) yields the boundary conditions, $\frac{\partial P}{\partial z} = -p/\delta$, $r = R$ and $\frac{\partial P}{\partial z} = 0$, $r = 0$.

Based on Wang and Woo's works, Eq. (5.6.10) has a trivial solution $P=0$, representing the stationary paraelectric state (Wang and Woo, 2005). To be able to transform between the paraelectric ($P = 0$) and the ferro-electric ($P \neq 0$) states, the initial state must become dynamically unstable. The dynamical stability of the initial state can be probed by applying a small perturbation Δ to the corresponding stationary solution of Eq. (5.6.10). The dynamics of Δ follow from Eq. (5.6.10) by retaining only terms linear in Δ , which is given by

$$\frac{\partial \Delta}{\partial t} = -M \left[\alpha_\sigma \Delta + 3\beta P^3 \Delta + 6\gamma P^5 \Delta - D \frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial \Delta}{\partial r} \right) \right]. \quad (5.6.12)$$

In the general case, Eq. (5.6.12) should be solved by using the Bessel functions (Zheng et al., 2008b). For simplicity, the following calculations only approximately assume that the gradient term of FNW is the same as that in the thin film. The boundary condition is the same as Eq. (5.6.11), only with P replaced by Δ . For a symmetric configuration, the condition $\frac{d\Delta}{dr} = 0$, at $r = 0$ must also hold.

Using the method of separation of variables, Eq. (5.6.12) can be transformed into an eigenvalue problem. For the $P = 0$ state, the solution can be written as $\Delta(r, t) = e^{\omega_c t} \varphi_{\omega_c}(r)$, where ω_c is the eigenvalue and $\varphi_{\omega_c}(z)$ the corresponding eigenfunction of equation

$$\omega_c = M[-\alpha_\sigma - Dk^2]. \quad (5.6.13)$$

The stationary solution $P = 0$ is unstable when $\omega_c > 0$ because, in this case, the perturbation Δ increases exponentially with time. It can be seen from Eq. (5.6.13) that when the temperature T is sufficiently high, $\omega_c < 0$ and the paraelectric state is stable. When T is sufficiently low, ω_c turns positive and this state is no longer stable, since any small perturbation Δ will grow exponentially beyond all bounds.

5.6.2.4 Curie Temperatures of Ferroelectric Nanowires

The critical condition $\omega_c = 0$ yields, in this case, the transition temperature T_c of the FNW

$$T_c = T_{c0} + \frac{2Q_{12}}{\alpha_0} \sigma_{\text{eff}} - \frac{D}{\alpha_0} k_c^2, \quad (5.6.14)$$

where the second term is related to the surface tension, and the third term to the lattice relaxation on the surface. While the last term acts against the transformation by lowering the transition temperature, the second term may act either way, depending on the effective radial pressures, i.e., p . Here k_c depends on the radius R and the extrapolating length δ , as the smallest non-zero root of the transcendental equation.

The study is conducted for a PbTiO_3 NW which has been successfully fabricated and implemented recently. Since typical values of the effective surface tension coefficient μ vary between 5 and 50 $\text{N}\cdot\text{m}^{-1}$ one can determine the relation between the phase transition temperature and the radius of the FNW by using linear stability analysis. The phase transition temperature T_c as a function of the radius R is as shown in Fig. 5.28. The effect of the surface tension can be immediately discerned from the peak of the transition temperature $T_c^{\text{max}} \approx 1.75T_{c0}$ at $R \approx 2$ nm. From Fig. 5.28, the critical radius of the NW at which T_c vanishes, i.e., ferroelectricity disappears permanently, can also be obtained. Similar observations for the existence of the critical ferroelectric size in FNWs and thin films have been discussed.

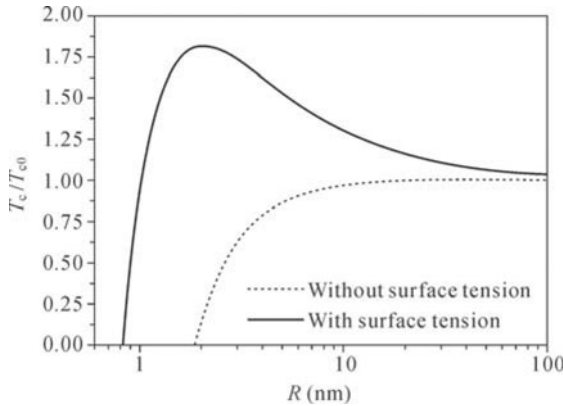


Fig. 5.28 The phase transition temperature T_c versus the radius R of the PbTiO_3 ferroelectric nanowire

5.6.2.5 Polarization of Ferroelectric Nanowires

It is also of interest to directly evaluate the strength of the spontaneous polarization. In terms of the Curie-Weiss relation, the spatially averaged free energy density can be written as

$$f_{\text{total}} = \frac{\alpha_0}{2}(T - T_c)P^2 + \frac{\beta}{4}P^4 + \frac{\gamma}{6}P^6 - (s_{11} + s_{12})\sigma_{\text{eff}}^2, \quad (5.6.15)$$

where T_c is the phase transition temperature of the FNW from Eq. (5.6.14), as a function of its radius and extrapolation length. Here note that P is a spatial average.

The free energy minimization of Eq. (5.6.15) with respect to P yields the following equation

$$\alpha_0 [T - T_c(R, \delta)]P + \beta P^3 + \gamma P^5 = 0. \quad (5.6.16)$$

Then the approximated polarization P can be written as

$$P = \left[\frac{-\beta \pm \sqrt{\beta^2 - 4\alpha_0\gamma [T - T_c(R, \delta)]}}{2\gamma} \right]^{1/2}. \quad (5.6.17)$$

Interestingly, besides increasing the transition temperature, the surface tension also affects the polarization in the FNW significantly. The average polarization P of a PbTiO_3 NW at room temperature as a function of radius R has been obtained as shown in Fig. 5.29. The similarity between Fig. 5.28 and Fig. 5.29 is clear. Fig. 5.29 also demonstrates the existence of a critical radius R_c of about 0.8 nm, below which the polarization disappears.

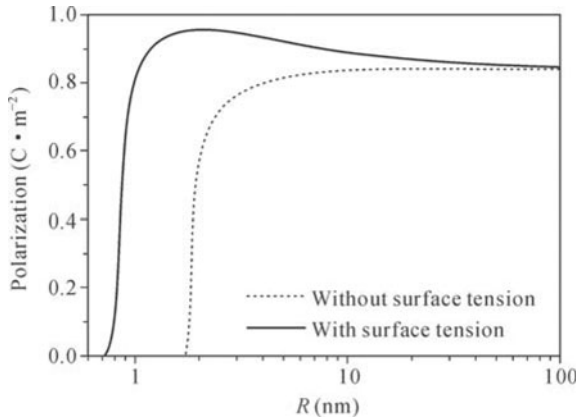


Fig. 5.29 The average polarization P versus the radius R of the PbTiO_3 ferroelectric nanowire

5.6.2.6 Curie-Weiss-type Relation of Ferroelectric Nanowires

The susceptibility χ of the FNW can be expressed as a function of P via the total free energy density f of Eq. (5.6.15)

$$\chi^{-1} = \left(\frac{\partial^2 f_{\text{total}}}{\partial P^2} \right) = \alpha_0 (T - T_c) + 3\beta P^2 + 5\gamma P^4, \quad (5.6.18)$$

where T_c is the phase transition temperature of the NW as given by Eq. (5.6.14).

Using Eq. (5.6.18), a Curie-Weiss-type relation of the spatial average $\langle \chi \rangle$ of a ferroelectric NW near Curie temperature can be expressed in the following form (Zheng et al., 2009)

$$\langle \chi \rangle = \begin{cases} \frac{\Theta^{-1}(\delta, w)}{\alpha_0 |T - T_c|}, & \text{for } T < T_c, \\ \frac{1}{\alpha_0 |T - T_c|}, & \text{for } T > T_c, \end{cases} \quad (5.6.19)$$

where $\Theta(\delta, w)$ is the Curie-Weiss parameter as a function of the extrapolation length and wall-thickness of the tube. The approximated equation of the Curie-Weiss parameter can be obtained as shown in Fig. 5.30. Of course, the total susceptibility of ferroelectric NWs should be given by $\langle \chi^T \rangle = \langle \chi \rangle + \chi_b$.

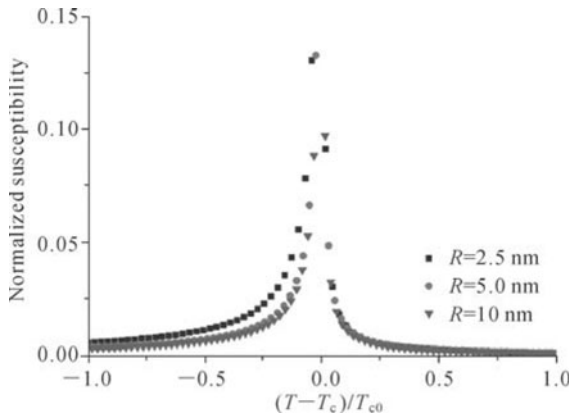


Fig. 5.30 Curie-Weiss relation of PbTiO₃ nanowire as function of the radius

5.6.3 Ferroelectric Nanotubes

5.6.3.1 Free Energy of Ferroelectric Nanotubes with Size and Surface Effects

In the following, the specific case of a ferroelectric NT with inner radius a , outer radius b , average radius $R = (a + b)/2$, length h and wall thickness w , is shown in Fig. 5.31. The total free energy of a ferroelectric NT can also be expressed as the sum of the bulk energy F_{bulk} , gradient energy F_{grad} , the electric field F_{Ele} and surface energy F_s , respectively.

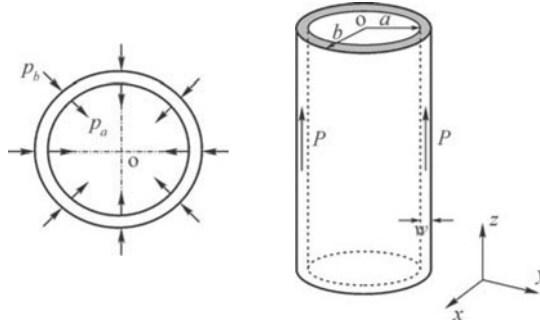


Fig. 5.31 Schematics of an NT under the effective radial pressures $p_a = -\mu_a/a$ and $p_b = -\mu_b/b$ induced by the surface tension

The surface tension creates a radial compression of the NT. The first term F_{bulk} in cylindrical coordinates (r, φ, z) becomes

$$F_p = 2\pi \int_{-h/2}^{h/2} dz \int_a^b r dr \left\{ \frac{\alpha_\sigma}{2} P^2(r, z) + \frac{\beta}{4} P^4(r, z) + \frac{\gamma}{6} P^6(r, z) - f(\sigma_{ij}) \right\}, \quad (5.6.20)$$

where $f(\sigma_{ij}) = \frac{1}{2} s_{ijkl} \sigma_{ij} \sigma_{kl}$ is a function of only external stresses. The expression of α_σ is given by

$$\alpha_\sigma = \alpha_0 (T - T_{c0}) - 2Q_{12} (\sigma_{rr} + \sigma_{\varphi\varphi}) - Q_{11} \sigma_{zz}, \quad (5.6.21)$$

where σ_{rr} and $\sigma_{\varphi\varphi}$ are the radial and tangential components of the stress field in the NT induced by the surface tension.

The Ginzburg gradient term F_{grad} in the free energy expression of the NT can be written as

$$F_{\text{grad}} = 2\pi \int_{-h/2}^{h/2} dz \int_a^b r dr \left\{ \frac{D}{2} [\nabla P(r, z)]^2 \right\}. \quad (5.6.22)$$

In the present section, the short-circuit boundary condition is considered, and the depolarization field along z direction can be neglected because the NT is long. So the polarization P^E induced by the depolarization field can be neglected. The last term F_s accounts for surface contributions to the total free energy. Taking into account contributions from the top, bottom and sidewalls, F_s can be expressed as

$$\begin{aligned} F_s = & D \int_a^b \frac{2\pi r}{\delta_{\text{end}}} \left[P^2(r, z = h/2) + P^2(r, z = -h/2) \right] dr \\ & + D \int_{-h/2}^{h/2} \frac{2\pi a}{\delta_{s-w}} P^2(r = a, z) dz + D \int_{-h/2}^{h/2} \frac{2\pi b}{\delta_{s-w}} P^2(r = b, z) dz, \end{aligned} \quad (5.6.23)$$

where δ_{end} and δ_{s-w} are the extrapolation lengths at the end surfaces and the sidewall of the nanotube, respectively.

From Eqs. (5.6.1) and (5.6.20), the normalized coefficient α_σ can be rewritten in terms of the surface tension σ_{st} ,

$$\alpha_\sigma = \alpha_0 (T - T_{c0}) - 2Q_{12} \sigma_{\text{st}}, \quad (5.6.24)$$

where σ_{st} is a function of the inner and outer radii, and can be expressed as

$$\sigma_{\text{st}} = \sigma_{rr} + \sigma_{\varphi\varphi} = \frac{a^2}{b^2 - a^2} p_a - \frac{b^2}{b^2 - a^2} p_b = -\frac{a\mu_a + b\mu_b}{b^2 - a^2} = -\frac{\bar{\mu}}{w} - \frac{\Delta\mu}{2R}, \quad (5.6.25)$$

where $\bar{\mu}$ is the average surface energy density and $\Delta\mu = (\mu_b - \mu_a)/2$. The minus sign in Eq. (5.6.25) indicates a compression. Note that Eq. (5.6.25) is inapplicable when $b \approx a$ because the magnitude of the stresses would have exceeded the validity of elasticity theory and plastic yielding would have occurred. Assuming a simple elastic-plastic model with a yield stress σ_Y , we may rewrite α_σ in the form

$$\alpha_\sigma(T, a, b) = \alpha_0 (T - T_{c0}) - 2Q_{12} \sigma_{\text{st}}^{\text{eff}}, \quad (5.6.26)$$

with an effective stress due to the surface tension defined as

$$\sigma_{\text{st}}^{\text{eff}} = \begin{cases} -\frac{\bar{\mu}}{w} \left(1 + \frac{w\Delta\mu}{2R\bar{\mu}} \right), & \text{for } w > w_Y, \\ -\sigma_Y, & \text{for } w \leq w_Y, \end{cases} \quad (5.6.27)$$

where $\bar{\mu}$ is the average surface energy density and $\Delta\mu = (\mu_b - \mu_a)/2$. w_Y is the thickness below which plastic yield occurs. Of course, Zheng et al. (2008) constructed an approximated function based on Eq. (5.6.26) to fit the effective stress, which could be given by $\sigma_{st}^{\text{eff}} = -\sigma_Y [1 - \exp(-w_Y/w)]$. The use of σ_{st}^{eff} helps us to avoid the unphysical divergence at $b \approx a$. If $|w\Delta\mu/(2R\bar{\mu})| \ll 1$, which is usually satisfied, σ_{st}^{eff} is independent of the average tube radius R , and any size-dependence of the ferroelectric properties of the tube is dominated by the wall-thickness and the extrapolation length. Note that effects on the polarization due to the strain fields on the dislocations generated at the yield may be important, but is a complicated problem that has to be considered as an issue on its own.

5.6.3.2 Stability of the Ferroelectric State and Transition Characteristics

The time evolution of the system is governed by the time-dependent Ginzburg-Landau equation. Using Eqs. (5.6.20) and (5.6.27), this can be written as

$$\begin{aligned} \frac{\partial P}{\partial t} &= -M \frac{\delta F}{\delta P} \\ &= -M \left\{ \alpha_\sigma(T, r) P(r) + \beta P^3(r) + \gamma P^5(r) - D \frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial P(r)}{\partial r} \right) \right\}. \end{aligned} \quad (5.6.28)$$

Here the surface term in Eq. (5.6.28) yields the boundary conditions

$$\frac{\partial P}{\partial r} = P/\delta_{s-w}, \quad \text{for } r = a, \quad \text{and} \quad \frac{\partial P}{\partial r} = -P/\delta_{s-w}, \quad \text{for } r = b.$$

Eq. (5.6.28) has a trivial solution $P=0$, representing the stationary paraelectric state. To be able to transform between the paraelectric ($P=0$) and the ferroelectric ($P \neq 0$) states, the initial state must become dynamically unstable. The dynamical stability of the initial state can be probed by applying a small perturbation Δ to the corresponding stationary solution of Eq. (5.6.28). The dynamics of Δ follow from Eq. (5.6.28) by retaining only terms linear in Δ , which is given by

$$\frac{\partial \Delta}{\partial t} = -M \left[\alpha_\sigma \Delta + 3\beta P^3 \Delta + 6\gamma P^5 \Delta - D \frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial \Delta}{\partial r} \right) \right]. \quad (5.6.29)$$

In the general case, Eq. (5.6.28) should be solved by using the Bessel functions. For simplicity, we only approximately assume the gradient term of NT is the same as it is in thin film. The boundary condition is the same as Eq. (5.6.11), only with P replaced by Δ . For a symmetric configuration, the condition $\frac{d\Delta}{dr} = 0$,

at $r = a + w/2$, must also hold.

Using the method of separation of variables, the TDGL equation can be transformed into an eigenvalue problem. For the $P = 0$ state, the solution can be written as $\Delta_c(r, t) = e^{\omega_c t} \varphi_{\omega_c}(r)$, where ω_c is the eigenvalue and $\varphi_{\omega_c}(z)$ the corresponding eigenfunction of equation (TDGL)

$$\omega_c = M[-\alpha_\sigma - Dk^2]. \quad (5.6.30)$$

The stationary solution $P = 0$ is unstable when $\omega_c > 0$, because in this case the perturbation Δ increases exponentially with time. It can be seen from Eq. (5.6.30) that when the temperature T is sufficiently high, $\omega_c < 0$, and the paraelectric state is stable. When T is sufficiently low, ω_c turns positive and this state is no longer stable, since any small perturbation Δ will grow exponentially beyond all bounds.

5.6.3.3 Curie Temperature and Critical wall Thickness of Ferroelectric Nanotubes

The critical condition $\omega_c = 0$ yields, in this case, the transition temperature T_c of the ferroelectric nanotube

$$T_c = T_{c0} + \frac{2Q_{12}}{\alpha_0} \sigma_{st}^{\text{eff}} - \frac{D}{\alpha_0} k_c^2 \quad (5.6.31)$$

where the last term acts against the transformation by lowering the transition temperature. The second term may act either way, depending on the effective radial pressures, i.e., p_a and p_b . Here k_c depends on the wall thickness w and the extrapolating length $\delta_{s-w} = \delta$, as the smallest non-zero root of the transcendental equation $\cot(k_c w/2) - k_c \delta = 0$.

Analytic solutions of Eq. (5.6.31) have been obtained in Wang's calculations for various important limits and approximations. Thus, for thin-walled tubes, i.e., $w \ll \delta$, $k_c^2 = 2w^{-1}(\delta + w/6)^{-1}$, Eq. (5.6.31) can be written as

$$T_c = T_{c0} + \frac{2Q_{12}}{\alpha_0} \sigma_{st}^{\text{eff}} - \frac{2D}{\alpha_0 w(\delta + w/6)}. \quad (5.6.32)$$

The corresponding critical wall thickness for a positive T_c is given by

$$w_c \approx \frac{2D}{\delta(4Q_{12}\sigma_{st}^{\text{eff}} + \alpha_0 T_{c0})}. \quad (5.6.33)$$

When $w \gg \delta$ solution of the transcendental equation gives $k_c = \frac{\pi}{w + 2\delta}$ and Eq. (5.6.31) for the transition gives

$$T_c = T_{c0} + \frac{2Q_{12}}{\alpha_0} \sigma_{st}^{\text{eff}} - \frac{D}{\alpha_0} \left(\frac{\pi}{w + 2\delta} \right)^2. \quad (5.6.34)$$

Neglecting δ in comparison with w , the corresponding critical wall thickness for a positive T_c can be easily solved, and is given by

$$w_c \approx \sqrt{\frac{\pi^2 D}{2Q_{12}\sigma_{st}^{\text{eff}} + \alpha_0 T_{c0}}}. \quad (5.6.35)$$

Note that in both limits of $w \ll \delta$ (small surface relaxation) and $w \gg \delta$ (large relaxation), the surface effect due to δ (the third term on the RHS) causes the transition temperature to decrease monotonically as the wall thickness w of the tube decreases, while the surface tension has the opposite effect, except when the wall thickness is below w_y . As $w < 2R$ always holds, for $\Delta\mu/\bar{\mu} \ll 1$, σ_{st}^{eff} is independent of R and both T_c and w_c are independent of R .

The analytical and numerical calculations in this section are for BaTiO₃ nanotubes which have been successfully fabricated and implemented in various devices, such as capacitors and nonvolatile memories etc. Only the case of a long tube with length h much larger than its inner/outer radius is considered, so that the depolarization field can be neglected, thus allowing us to concentrate on the effects of the surface tension. The ideal case is that the difference between the surface energy densities of the inner and outer surfaces is negligible, i.e., $\Delta\mu \approx 0$. Since typical values of the effective surface tension coefficient vary between 5 and 50 N·m⁻¹, a value of $\mu_a = \mu_b = \mu = 20$ N·m⁻¹ is used.

Phase-transition temperatures T_c are calculated for various wall thicknesses w of an NT with inner radius $a = 10$ nm and surface energy density $\mu = 20$ N·m⁻¹ by solving Eq. (5.6.31). The results are plotted in Fig. 5.32a, as a function of w . The important feature that can immediately be discerned is the maximum shown by T_c at $T_c^{\text{max}} \approx 580$ K for a wall thickness of $w \approx 8.5$ nm. Furthermore, for a wall thickness larger than ~ 4 nm, T_c is actually higher than that of the bulk material at 400 K.

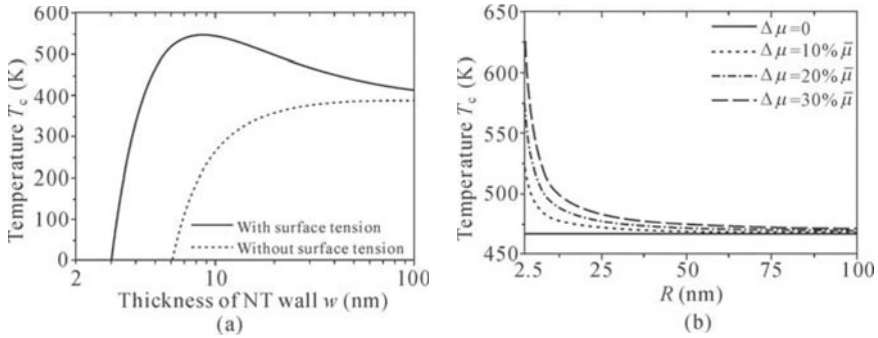


Fig. 5.32 (a) The phase transition temperature T_c versus the wall thickness w of NT, where the inner radius is $a = 10$ nm, without considering the surface tension effect and considering the surface tension effect; (b) The phase transition temperature T_c versus the average radius R , where the inner and outer surface tension coefficient is considered as the difference

For comparison, the calculation was repeated without the surface tension contributions, i.e., by setting $\mu = 0 \text{ N}\cdot\text{m}^{-1}$ and plotted as the line in Fig. 5.32a. It can be seen from the comparison that the surface tension is responsible for the raising of the transition temperature, with a magnitude that tends to increase as the increasing surface tension w decreases. As the compression due to the surface tension stops increasing beyond the elastic limit, the effects of the near-surface eigen-strain relaxation described by the extrapolation length δ start to dominate and lower the transition temperature as w decreases. As a result, the rise in T_c shows a maximum of $T_c^{\max} \approx 580 \text{ K}$ at a wall thickness of $w \approx 8.5 \text{ nm}$. According to these results, in addition to materials constants such as the Landau free-energy parameters, the surface energy density and extrapolation length, the transition temperature of the NT is also sensitive to geometry parameters such as the wall thickness w .

In comparison with thin films, the operation of the surface tension leads to the following important differences: (1) contrary to thin films in which the transition temperature monotonically decreases with decreasing film thickness, the transition temperature for NTs attains a maximum as the wall thickness decreases, and (2) the critical thickness for the disappearance of ferroelectricity in an NT is significantly larger than that of the thin film, and sometimes even the bulk material. Similar observations for ferroelectric NW and thin films have been discussed (Zheng et al., 2008b). In pure materials, the surface energy densities between the inner and outer wall surfaces of the NT are the same, so it is reasonable to assume that $\Delta\mu \approx 0$. However, practical considerations sometimes require this condition to be relaxed, such as in cases where the environments encountered by the inner and outer surfaces are different. It is considered that properties of ferroelectric NT can be changed by adjusting the difference in the surface tension coefficients between the inner and outer surfaces. It is clear that in such cases the transition temperature also depends on the average radius R of the

NT as well as $\Delta\mu$. In Fig. 5.32b, the calculated transition temperature T_c as a function of the average radius R for different values of $\Delta\mu/\bar{\mu}$ is plotted. It can be seen that positive values of $\Delta\mu/\bar{\mu}$ cause an increase in T_c , which increases as the average tube radius R decreases. The increase is most rapid for the small-radii tubes with an average tube radius of less than 10 nm. For a value of $\Delta\mu/\bar{\mu} = 30\%$, the T_c for a tube with an average radius of 2.5 nm can be higher by 100 K than one at 10 nm. For a tube with radii larger than 25 nm, T_c is relatively insensitive to the tube size.

5.6.3.4 Polarization of Ferroelectric Nanotubes

It is also of interest to directly evaluate the strength of the polarization. In terms of the Curie-Weiss relation, the spatially averaged free energy density can be written as

$$f_{\text{total}} = \frac{\alpha_0}{2}(T - T_c)P^2 + \frac{\beta}{4}P^4 + \frac{\gamma}{6}P^6 - f(\sigma_{ij}), \quad (5.6.36)$$

where T_c is the phase transition temperature of the tube from Eq. (5.6.31), as a function of its wall thickness, radius and extrapolation length. Here we note that P is a spatial average.

Using the variational method, free energy minimization with respect to P yields the following equation

$$\alpha_0 [T - T_c(R, w, \delta)]P + \beta P^3 + \gamma P^5 = 0, \quad (5.6.37)$$

yielding

$$P = \left[\frac{-\beta \pm \sqrt{\beta^2 - 4\alpha_0\gamma [T - T_c(R, w, \delta)]}}{2\gamma} \right]^{1/2}. \quad (5.6.38)$$

In addition to raising the transition temperature, the enhancement of ferroelectricity in NTs via the effects of the surface tension can also be seen directly from the strength of the average polarization. By solving Eq. (5.6.38), the average polarization for a BTO NT can be obtained as a function of wall thickness w , which is plotted in Fig. 5.33a.

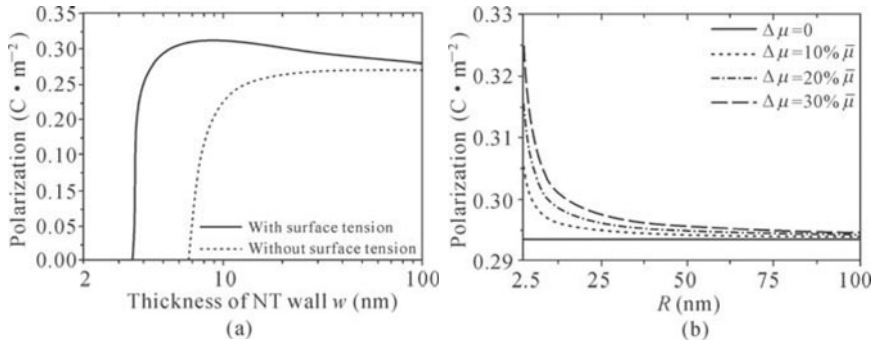


Fig. 5.33 (a) The polarization versus the wall thickness of NT at room temperature, where the inner radius is $a = 10$ nm, without considering the surface tension effect and considering the surface tension effect; (b) The polarization versus the average radius R , where the inner and outer surface tension coefficient is considered as the difference

The line shows similar results obtained without considering surface tension (i.e., $\mu = 0 \text{ N} \cdot \text{m}^{-1}$). The enhancement effect of the polarization due to the surface tension is obvious, and follows the same trend as the transition temperature, reaching a maximum near $w = 8$ nm, at which point the stress in the tube caused by the surface tension maximizes at the yield point. The results in Fig. 5.33a also demonstrate directly the existence of a critical wall thickness w_c of about 3.2 nm at 0 K, below which the polarization disappears. Fig. 5.33b shows the corresponding plot of the polarization as a function of the average tube radius R for various values of $\Delta\mu$. The similarity between Figs. 5.29b and 5.33b is obvious.

5.6.3.5 Curie-Weiss-type relation of Ferroelectric Nanotubes

The susceptibility χ can be expressed as a function of P via the total free energy density f_{total} of the NT

$$\chi^{-1} = \left(\frac{\partial^2 f_{\text{total}}}{\partial P^2} \right) = \alpha_0 (T - T_c) + 3\beta P^2 + 5\gamma P^4, \quad (5.6.39)$$

where T_c is the phase transition temperature of the NT as given by Eq. (5.6.31).

A Curie-Weiss-type relation of the spatial average $\langle \chi \rangle$ of a ferroelectric NT near Curie temperature can be expressed in the following form

$$\langle \chi \rangle = \begin{cases} \frac{\Theta^{-1}(\delta, w)}{\alpha_0 |T - T_c|}, & \text{for } T < T_c, \\ \frac{1}{\alpha_0 |T - T_c|}, & \text{for } T > T_c, \end{cases} \quad (5.6.40)$$

where $\Theta(\delta, w)$ is the Curie-Weiss parameter as a function of the extrapolation length and wall thickness of the tube.

Based on previous calculations in Wang et al. (2006), an approximated equation of the Curie-Weiss parameter can be obtained. Fig. 5.34 shows a Curie-Weiss-type relation of the dielectric permittivity for a BTO NT with different wall thickness near the transition temperature. It can be seen that when the ferroelectric NT is in the paraelectric state ($T > T_c$), the Curie-Weiss relation is independent of the geometry of the tube. When ferroelectric NT is in a ferroelectric state, the Curie-Weiss relation is a function of the wall thickness, radius and extrapolation length.

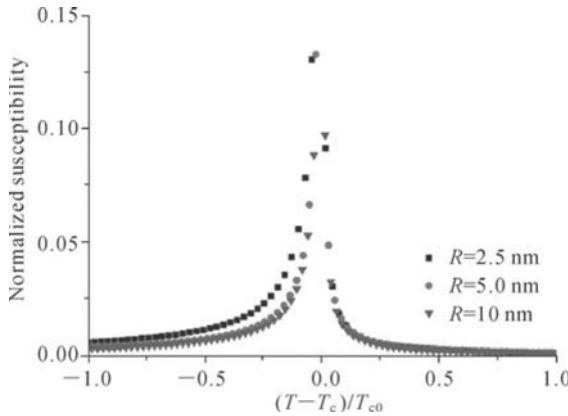


Fig. 5.34 The Curie-Weiss-type relation in ferroelectric NT, where the wall thickness is $w = 2.5, 5$ and 10 nm, respectively

5.7 Ferroelectric Nanograins or Nanoislands

Recently, the polarization instability of ferroelectric nanograins or nanoislands has attracted much attention due to the potential application of nanoscale ferroelectrics in non-volatile memories. It is well known that due to the surface constraints, generally speaking, the Curie temperature will decrease with a decrease in the size of the ferroelectrics, and it may drop to below 0 K for ferroelectric materials in nanoscale. Thus the ferroelectricity may be totally suppressed if the size of ferroelectrics is below some limits. Wang et al. (1995) estimated the lateral

limiting size of a three-dimensional ferroelectric cell using phenomenological analysis. But they did not take the misfit stress into consideration. Roelofs et al. (2002) fabricated ferroelectric lead titanate nanoislands down to 15 nm on Si/SiO₂/TiO₂/Pt substrate, and they found that the limiting size for the ferroelectric phase is about 20 nm. In a more recent paper, Chu et al. (2004) investigated the effect of lattice dislocations at the interface between the ferroelectric nanometer-sized island and electrodes on the size limit, and they concluded that controlling such defects to keep the epitaxial stress will be crucial in pushing the storage density of ferroelectric memories to such a limit. In fact, the misfit dislocations play the role of releasing the misfit stresses. In this section, the effect of the epitaxial stress on the size limit of such a single ferroelectric nanoisland or nanograin on a substrate is discussed based on stability analysis.

5.7.1 Free Energy of Ferroelectric Nanograins or Nanoislands

A single ferroelectric nanoisland was modeled as a cuboid with size $x_0 \times x_0 \times h$, where h is the thickness of the element (Fig. 5.35). The bottom plane is connected with a substrate. The misfit strain between the film and the substrate is given by $\varepsilon_{xx}^M = \varepsilon_{yy}^M = \varepsilon_m = (a - a_0)/a_0$, where a and a_0 are the lattice constants of the substrate and the film, respectively (Wang et al., 2010).



Fig. 5.35 PTO nanoisland grown on STO substrate

If the thickness is not very small compared with x_0 , the misfit strain will induce an inhomogeneous epitaxial stress distribution in the film. The maximum epitaxial stresses appear on the bottom plane of the cuboid, whereas the minimum epitaxial stresses appear on the top plane of the cuboid. The epitaxial stress can be described by the position-dependent function as $\sigma^M = \sigma_{xx} = \sigma_{yy} = \sigma_0 e^{-\beta z}$, where σ_0 is the film/substrate interface stress and $\sigma_0 = G\varepsilon_{xx}^M = [C_{11} + C_{12} - C_{13}(C_{13} + C_{23})/C_{33}]\varepsilon_{xx}^M$, where C_{11} , C_{12} , C_{13} , C_{23} , and C_{33} are components of the elastic modulus of the film. The stress components along the z -direction will not affect the ferroelectric transition if we assume the polarization is along the z -axis. Accompanying the ferroelectric transition from the cubic to tetragonal phase along the z -axis, the transformation strain can be written as $\varepsilon_0 = \varepsilon_{xx}^T = \varepsilon_{yy}^T = QP^2$, where P is the spontaneous polarization and Q is the electrostrictive coefficient. Including the effects of the depolarization field and the epitaxial stresses, the free energy of the cell can be written as

$$\begin{aligned}
F = & \iiint_V \left\{ \frac{A}{2} (T - T_{c0}) P^2 + \frac{B}{4} P^4 + \frac{C}{6} P^6 + \frac{D_{44}}{2} \left(\frac{\partial P}{\partial x} \right)^2 + \frac{D_{44}}{2} \left(\frac{\partial P}{\partial y} \right)^2 \right. \\
& + \frac{D_{11}}{2} \left(\frac{\partial P}{\partial z} \right)^2 - \frac{1}{2} E_d P - 2 \sigma^M \epsilon_0 \} dV + \iint_{S_x} \frac{D_{44} \delta_1^{-1}}{2} P^2 dz dy \\
& + \iint_{S_y} \frac{D_{44} \delta_1^{-1}}{2} P^2 dx dz + \iint_{S_z} \frac{D_{11} \delta_3^{-1}}{2} P^2 dx dy,
\end{aligned} \tag{5.7.1}$$

where D_{11} and D_{44} are expansion coefficients of the gradient term, S_x, S_y, S_z represent the surface planes that cover the cuboid and δ_1, δ_3 are the extrapolation lengths that measure the strength of the surface effects normal to the x - (and y -) axis and z -axis on the polarization, respectively. In Eq. (5.7.1), it is assumed that the properties of the island are the same along the x -direction and y -direction. The depolarization field is related to the polarization field and the boundary conditions. There are two cases: (a) with the ferroelectric island sandwiched between two metallic electrodes in short-circuit conditions and (b) with the island sandwiched between two dielectric substrates.

5.7.2 Stability of the Ferroelectric State and Transition Characteristics

The time evolution of the system from an initial state is governed by the time-dependent Ginzburg-Landau equation

$$\begin{aligned}
M \frac{\partial P}{\partial t} = & - \frac{\delta F}{\delta P} \\
= & -A(T - T_{c0})P + 4Q\sigma^M P - \frac{1}{\alpha} P - BP^3 - CP^5 \\
& + D_{44} \frac{\partial^2 P}{\partial x^2} + D_{44} \frac{\partial^2 P}{\partial y^2} + D_{11} \frac{\partial^2 P}{\partial z^2} + \frac{\phi}{\alpha h} \int_0^h P dz.
\end{aligned} \tag{5.7.2}$$

The surface term in Eq. (5.7.1) yields the boundary conditions

$$\begin{aligned}
\frac{\partial P}{\partial z} = \frac{P}{\delta_3}, \text{ for } z = 0 \text{ and } \frac{\partial P}{\partial z} = -\frac{P}{\delta_3}, \text{ for } z = h, \\
\frac{\partial P}{\partial x} = \mp \frac{P}{\delta_1}, \text{ for } x = \pm \frac{x_0}{2}, \\
\frac{\partial P}{\partial y} = \mp \frac{P}{\delta_1}, \text{ for } y = \pm \frac{x_0}{2}.
\end{aligned} \tag{5.7.3}$$

Eqs. (5.7.2) and (5.7.3) have a trivial solution of $P = 0$, representing the stationary non-ferroelectric state. To be able to transform to the ferroelectric ($P \neq 0$) state, the stationary non-ferroelectric state must become unstable. To determine the critical bifurcation point, one needs only to derive a solution of the following linear characteristic equation

$$\begin{aligned} \omega\Delta = & -A(T - T_{c0})\Delta + 4Q\sigma^M\Delta - \frac{1}{\alpha}\Delta + D_x \frac{\partial^2 \Delta}{\partial x^2} + D_x \frac{\partial^2 \Delta}{\partial y^2} \\ & + D_z \frac{\partial^2 \Delta}{\partial z^2} + \frac{\phi}{\alpha h} \int_0^h \Delta dz, \end{aligned} \quad (5.7.4)$$

under the boundary condition Eq. (5.7.3). The solution can be expressed in the form

$$\Delta = \cos(k_x x) \cos(k_y y) [C_1 \Delta_1(z) + C_2 \Delta_2(z)], \quad (5.7.5)$$

where k_x is determined by the following equation,

$$k_x \tan\left(\frac{k_x x_0}{2}\right) = \frac{1}{\delta_x}, \quad (5.7.6)$$

$$\Delta_1(z) = J_v(bx) + R_1 \chi(z); \Delta_2 = Y_v(bx) + R_2 \chi(z) \quad (5.7.7)$$

$$x = e^{-\frac{1}{2}\beta z},$$

$$\nu^2 = \frac{4A(T_c - T_{c0}) + \frac{1}{\alpha} + 2D_{44}k_x^2}{D_{11}\beta^2},$$

$$b^2 = \frac{16Q\sigma_0}{D_{11}\beta^2},$$

$$R_1 = \frac{1}{\alpha h} \left\{ 1 - \frac{\pi}{2\alpha h} \left[\int_0^h J_v(bx) dz \int bx Y_v(bx) dx - \int_0^h Y_v(bx) dz \int bx J_v(bx) dx \right] \right\}^{-1} \int_0^h J_v(bx) dz,$$

$$R_2 = \frac{1}{\alpha h} \left\{ 1 - \frac{\pi}{2\alpha h} \left[\int_0^h J_v(bx) dz \int bx Y_v(bx) dx - \int_0^h Y_v(bx) dz \int bx J_v(bx) dx \right] \right\}^{-1} \int_0^h Y_v(bx) dz,$$

$$\chi(z) = \frac{\pi}{2} \left[J_v(bx) \int bx Y_v(bx) dx - Y_v(bx) \int bx J_v(bx) dx \right].$$

(5.7.8)

$J_\nu(bx)$ and $Y_\nu(bx)$ are two Bessel functions. Substitution of Eq. (5.7.5) into the boundary conditions $\frac{\partial \Delta}{\partial z} = \frac{\Delta}{\delta_z}$, at $z = 0$, yields a relation between the constant C_1 and C_2 as follows

$$C_2 = \gamma C_1, \quad (5.7.9)$$

where

$$\begin{aligned} \gamma = & \left\{ \left[J'_\nu(bx) + R_1 X'(z) \right] \right\}_{z=0} - \frac{1}{\delta_z} \left[J_\nu(bx) + R_1 X(z) \right] \Big|_{z=0} \Bigg\} \\ & \times \left\{ \frac{1}{\delta_z} \left[Y_\nu(bx) + R_2 X(z) \right] \right\}_{z=0} - \left[Y'_\nu(bx) + R_2 X'(z) \right] \Big|_{z=0} \Bigg\}^{-1}, \end{aligned} \quad (5.7.10)$$

Then substitution of the solution into the boundary condition $\frac{\partial \Delta}{\partial z} = -\frac{\Delta}{\delta_z}$, at $z = h$, yields

$$\begin{aligned} & \left[J'_\nu(bx) + R_1 X'(z) \right] \Big|_{z=h} + \gamma \left[Y'_\nu(bx) + R_2 X'(z) \right] \Big|_{z=h} \\ & = -\frac{1}{\delta_z} \left\{ \left[J_\nu(bx) + R_1 X(z) \right] \Big|_{z=h} + \gamma \left[Y_\nu(bx) + R_2 X(z) \right] \Big|_{z=h} \right\}. \end{aligned} \quad (5.7.11)$$

5.7.3 Critical Properties of Nanograins or Nanoislands

By solving Eq. (5.7.11), one can obtain the solution of v^2 , from which the transition temperature as a function of the thickness and width of the ferroelectric nanoislands can be obtained. Generally speaking, the critical sizes of the ferroelectric nanoislands refer to the smallest sizes of the island at which the ferroelectric transition was totally suppressed even when the temperature drops to 0 K; i.e., the ferroelectric nanoislands become non-ferroelectric. To assign the transition temperature as 0 K, the critical sizes (h , x_0) can be determined from Eq. (5.7.11). In the following, a PbTiO_3 nanoisland on SrTiO_3 substrate was calculated numerically. The polarization is assumed to be perpendicular to the film/substrate interface. The critical sizes of the ferroelectric nanoislands are shown in Fig. 5.36 (Wang et al., 2010).

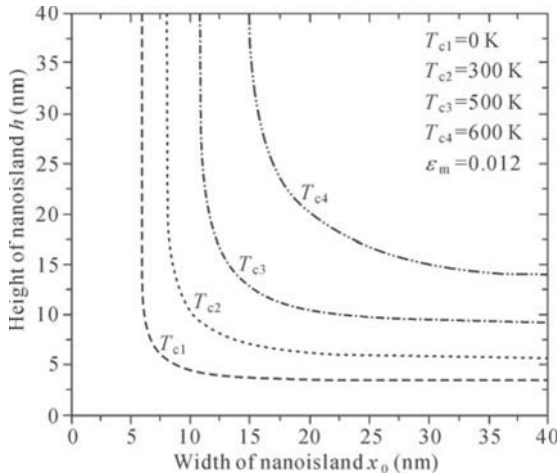


Fig. 5.36 The critical size of nanoisland under different transition temperatures

Based on their experimental observations, Chu et al. (2004) claimed that the misfit dislocations may play a dominant role in enlarging the stable ferroelectric nanoisland by eliminating the misfit stress. To check the effect of the misfit stress, Fig. 5.37a shows the critical sizes of the ferroelectric nanoislands at three different misfit strains. Indeed, the critical size of the ferroelectric grain is getting larger by eliminating the misfit stress, but the effect of the misfit stress is not so crucial.

The extrapolation length δ is a very important material constant in such a phenomenological model. For three different extrapolation lengths, Fig. 5.38b gives the critical sizes of the grain. One can find that the critical size decreases with an increase in the extrapolation length. That is reasonable since the smaller the extrapolation length, the stronger the effect of the surface.

5.8 Remarks

Modeling ferroelectrics in nanoscale using the Landau free-energy functional approach requires incorporating contributions from effects of the size, surface, interfacial and depolarization field, etc. Chapter 5 has shown derivation of the free-energy functional based on the Landau expansion in straight adherence to thermodynamic principles, i.e., the SPOP approach. Expressions of the depolarization field and free energy by using the TPOP approach may have effectively double-counted the depolarization field, and thus result in a serious overestimation of its effects on properties of ferroelectrics in nanoscale.

More importantly, based on stability analysis (Wang and Woo, 2005), Chapter 5 comprehensively calculated and discussed effects of size and surface on the critical properties of ferroelectrics in nanoscale, such as ferroelectric thin films,

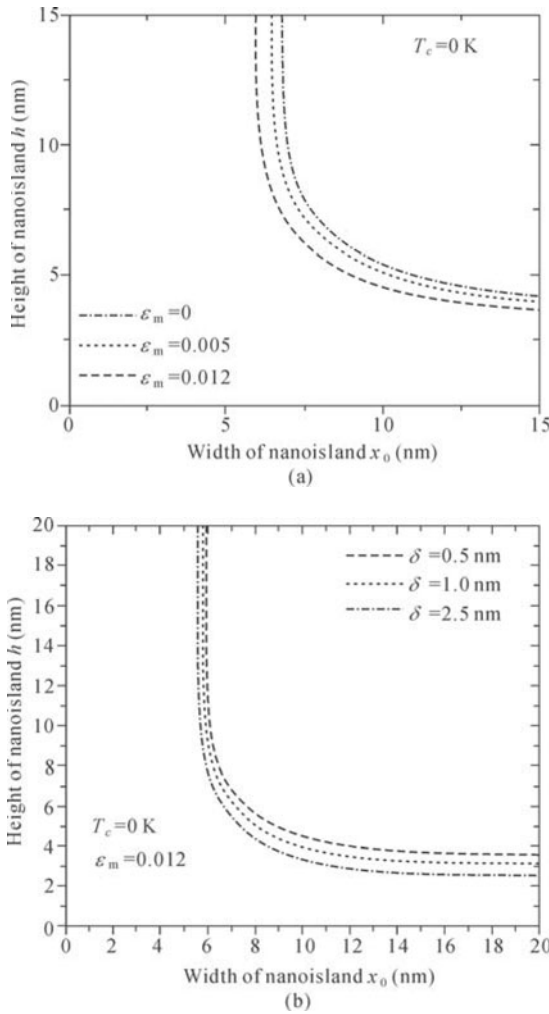


Fig. 5.37 (a) The critical size of nanoisland with different misfit strains ($T_c = 0$); (b) The critical size of nanoisland with different extrapolation lengths

capacitors or tunnel junctions, superlattices, nanowires, nanotubes, nanograins and nanoislands. By analyzing the conditions of dynamic instability of the time-dependent Ginzburg-Landau equation, analytic expressions have been developed to describe critical properties of nanoscale ferroelectric materials, such as the phase transition temperature, critical size, depolarization field, polarization, Curie-Weiss-type relation, etc.

Nevertheless, despite the more reasonable thermodynamic model of nanoscale ferroelectrics in Chapter 5, we still believe that there are many other outstanding challenges to thermodynamic models based on the Landau-type theory. A comprehensive treatment of the surface, interface, background dielectric

properties, etc. of ferroelectrics in nanoscale still requires a better understanding of materials behavior, which may require the combined efforts of phenomenological theory, molecular dynamics, first-principle-based calculations, etc.

References

- [1] Ahn, C.H., Rabe, K.M., Triscone, J.M.: Ferroelectricity at the nanoscale: Local polarization in oxide thin films and heterostructures. *Science* **303**, 488-491 (2004)
- [2] Alexe, M., Hesse, D., Schmidt, V., Senz, S., Fan, H.J., Zacharias, M., Gosele, U.: Ferroelectric nanotubes fabricated using nanowires as positive templates. *Appl. Phys. Lett.* **89**, 172907 (2006)
- [3] Choi, T., Lee, S., Choi, Y.J., Kiryukhin, V., Cheong, S.W.: Switchable ferroelectric diode and photovoltaic effect in BiFeO₃. *Science* **324**, 63-66 (2009)
- [4] Chu, M.W., Szafraniak, I., Scholz, R., Harnager, C., Hesse, D., Alexe, M., Gosele, U.: Impact of misfit dislocations on the polarization instability of epitaxial nanostructured ferroelectric perovskites. *Nat. Mater.* **3**, 87-90 (2004)
- [5] Dawber, M., Rabe, K.M., Scott, J.F.: Physics of thin-film ferroelectric oxides. *Rev. Mod. Phys.* **77**, 1083-1130 (2005a)
- [6] Dawber, M., Lichtensteiger, C., Cantoni, M., Veithen, M., Ghosez, P., Johnston, K., Rabe, K.M., Triscone, J.M.: Unusual behavior of the ferroelectric polarization in PbTiO₃/SrTiO₃ superlattices. *Phys. Rev. Lett.* **95**, 177601 (2005b)
- [7] Dawber, M., Stucki, N., Lichtensteiger, C., Gariglio, S., Ghosez, P., Triscone, J.M.: Tailoring the properties of artificially layered ferroelectric superlattices. *Adv. Mater.* **19**, 4153-4159 (2007)
- [8] Duan, C.G., Jaswal, S.S., Tsymbal, E.Y.: Predicted magnetoelectric effect in Fe/BaTiO₃ multilayers: Ferroelectric control of magnetism. *Phys. Rev. Lett.* **97**, 047201 (2006a)
- [9] Duan, C.G., Sabirianov, R.F., Mei, W.N., Jaswai, S.S., Tsymbal, E.Y.: Interface effect on ferroelectricity at the nanoscale. *Nano Lett.* **6**, 483-487 (2006b)
- [10] Eliseev, E.A., Morozovska, A.N.: General approach for the description of size effects in ferroelectric nanosystems. *J. Mater. Sci.* **44**, 5149-5160 (2009)
- [11] Eliseev, E.A., Morozovska, A.N., Glinchuk, M.D., Blinc, R.: Spontaneous flexoelectric/flexomagnetic effect in nanoferroics. *Phys. Rev. B* **79**, 165433 (2009)
- [12] Fong, D.D., Stephenson, G.B., Streiffer, S.K., Eastman, J.A., Auciello, O., Fuoss, P.H., Thompson, C.: Ferroelectricity in ultrathin perovskite films. *Science* **304**, 1650-1653 (2004)
- [13] Gerra, G., Tagantsev, A.K., Setter, N.: Ferroelectricity in asymmetric metal-ferroelectric-metal heterostructures: A combined first-principles-phenomenological approach. *Phys. Rev. Lett.* **98**, 207601 (2007)
- [14] Glinchuk, M.D., Morozovskaya, A.N.: Effect of surface tension and depolarization field on ferroelectric nanomaterial properties. *Phys. Status*

- Soli. B **238**, 81-91 (2003)
- [15] Gonzalo, J.A., Jimenez, B.: *Ferroelectricity: The Fundamental Collection*. Wiley-VCH, Weinheim (2005)
 - [16] Haun, M.J., Furman, E., Jang, S.J., McKinstry, H.A., Cross, L.E.: Thermodynamic theory of PbTiO_3 . *J. Appl. Phys.* **62**, 3331-3338 (1987)
 - [17] Hu, S.Y., Li, Y.L., Chen, L.Q.: Effect of interfacial dislocations on ferroelectric phase stability and domain morphology in a thin film—a phase-field model. *J. Appl. Phys.* **94**, 2542-2547 (2003)
 - [18] Jang, J.E., Cha, S.N., Choi, Y., Amaratunga, G.A.J., Kang, D.J., Hasko, D.G., Jung, J.E., Kim, J.M.: Nanoscale capacitors based on metal-insulator-carbon nanotube-metal structures. *Appl. Phys. Lett.* **87**, 263103 (2005)
 - [19] Junquera, J., Ghosez, P.: Critical thickness for ferroelectricity in perovskite ultrathin films. *Nature* **422**, 506-509 (2003)
 - [20] Kim, D.J., Jo, J.Y., Kim, Y.S., Chang, Y.J., Lee, J.S., Yoon, J.G., Song, T.K., Noh, T.W.: Polarization relaxation induced by a depolarization field in ultrathin ferroelectric BaTiO_3 capacitors. *Phys. Rev. Lett.* **95**, 237602 (2005)
 - [21] Kittel, C.: *Introduction to Solid State Physics*, 8th ed. pp.473 John Wiley & Sons, New York (2005)
 - [22] Kohlstedt, H., Pertsev, N.A., Contreras, J.R., Waser, R.: Theoretical current-voltage characteristics of ferroelectric tunnel junctions. *Phys. Rev. B* **72**, 125341 (2005)
 - [23] Kretschmer, R., Binder, K.: Surface effects on phase transitions in ferroelectrics and dipolar magnets. *Phys. Rev. B* **20**, 1065-1076 (1979)
 - [24] Landau, L.D., Lifshitz, E.M., Pitaevskii, L.P.: *Electrodynamics of Continuous Media*. Pergamon Press, Oxford (1960)
 - [25] Li, S.P., Eastman, J.A., Vetrone, J.M., Foster, C.M., Newnham, R.E., Cross, L.E.: Dimension and size effects in ferroelectrics. *Jpn. J. Appl. Phys.* **36**, 5169-5174 (1997)
 - [26] Li, Y.L., Hu, S.Y., Liu, Z.K., Chen, L.Q.: Effect of electrical boundary conditions on ferroelectric domain structures in thin films. *Appl. Phys. Lett.* **81**, 427-429 (2002)
 - [27] Li, Y.L., Cross, L.E., Chen, L.Q.: A phenomenological thermodynamic potential for BaTiO_3 single crystals. *J. Appl. Phys.* **98**, 064101 (2005)
 - [28] Li, Y.L., Hu, S.Y., Tenne, D., Soukiassian, A., Schlom, D.G., Chen, L.Q., Xi, X.X., Choi, K.J., Eom, C.B., Saxena, A., Lookman, T., Jia, Q.X.: Interfacial coherency and ferroelectricity of $\text{BaTiO}_3/\text{SrTiO}_3$ superlattice films. *Appl. Phys. Lett.* **91**, 252904 (2007a)
 - [29] Li, Y.L., Hu, S.Y., Tenne, D., Soukiassian, A., Schlom, D.G., Xi, X.X., Choi, K.J., Eom, C.B., Saxena, A., Lookman, T., Jia, Q.X., Chen, L.Q.: Prediction of ferroelectricity in $\text{BaTiO}_3/\text{SrTiO}_3$ superlattices with domains. *Appl. Phys. Lett.* **91**, 112914 (2007b)
 - [30] Lines, M.E., Glass, A.M.: *Principles and Applications of Ferroelectrics and Related Materials*. Oxford University Press, New York (1979)
 - [31] Liu, W., Sun, X.H., Han, H.W., Li, M.Y., Zhao, X.Z.: Surface polarization enhancement in $(\text{Pb}_{0.25}\text{Ba}_{0.15}\text{Sr}_{0.6})\text{TiO}_3$ nanotubes. *Appl. Phys. Lett.* **89**,

- 163122 (2006)
- [32] Luo, X., Lin, S.P., Wang, B., Zheng, Y.: Impact of applied strain on the electron transport through ferroelectric tunnel junctions. *Appl. Phys. Lett.* **97**, 012905 (2010)
 - [33] Luo, Y., Szafraniak, I., Zakharov, N.D., Nagarajan, V., Steinhart, M., Wehrspohn, R.B., Wendorff, J.H., Ramesh, R., Alexe, M.: Nanoshell tubes of ferroelectric lead zirconate titanate and barium titanate. *Appl. Phys. Lett.* **83**, 440-442 (2003)
 - [34] Lupascu, D.C.: *Fatigue in Ferroelectric Ceramics and Related Issues*. Springer, Berlin Heidelberg (2004)
 - [35] Ma, W.H., Zhang, M.S., Lu, Z.H.: A study of size effects in PbTiO_3 nanocrystals by Raman spectroscopy. *Phys. Status Soli. A* **166**, 811-815 (1998)
 - [36] Mehta, R.R., Silverman, B.D., Jacobs, J.T.: Depolarization fields in thin ferroelectric films. *J. Appl. Phys.* **44**, 3379-3385 (1973)
 - [37] Morozovska, A.N., Eliseev, E.A., Glinchuk, M.D.: Ferroelectricity enhancement in confined nanorods: Direct variational method. *Phys. Rev. B* **73**, 214106 (2006)
 - [38] Morozovska, A.N., Kalinin, S.V., Eliseev, E.A., Gopalan, V., Svechnikov, S.V.: Interaction of a 180° ferroelectric domain wall with a biased scanning probe microscopy tip: Effective wall geometry and thermodynamics in Ginzburg-Landau-Devonshire theory. *Phys. Rev. B* **78**, 125407 (2008)
 - [39] Naumov, I.I., Fu, H.X.: Spontaneous polarization in one-dimensional $\text{Pb}(\text{ZrTi})\text{O}_3$ nanowires. *Phys. Rev. Lett.* **95**, 247602 (2005)
 - [40] Nicolis, G., Prigogine, I.: *Self-organization in Nonequilibrium Systems*. John Wiley & Sons, New York (1977)
 - [41] Pertsev, N.A., Kohlstedt, H.: Depolarizing-field effect in strained nanoscale ferroelectric capacitors and tunnel junctions. *Arxiv preprint cond-mat/0603762* (2006)
 - [42] Pertsev, N.A., Kohlstedt, H.: Elastic stabilization of a single-domain ferroelectric state in nanoscale capacitors and tunnel junctions. *Phys. Rev. Lett.* **98**, 257603-257606 (2007)
 - [43] Pertsev, N.A., Zembilgotov, A.G., Tagantsev, A.K.: Effect of mechanical boundary conditions on phase diagrams of epitaxial ferroelectric thin films. *Phys. Rev. Lett.* **80**, 1988-1991 (1998)
 - [44] Pintilie, L., Vrejoiu, I., Rhun, G.L., Alexe, M.: Short-circuit photocurrent in epitaxial lead zirconate-titanate thin films. *J. Appl. Phys.* **101**, 064109 (2007)
 - [45] Pintilie, L., Boldyreva, K., Alexe, M., Hesse, D.: Capacitance tuning in antiferroelectric-ferroelectric PbZrO_3 - $\text{Pb}(\text{Zr}_{0.8}\text{Ti}_{0.2})\text{O}_3$ epitaxial multilayers. *New J. Phys.* **10**, 013003 (2008)
 - [46] Qu, B.D., Zhong, W.L., Prince, R.H.: Interfacial coupling in ferroelectric superlattices. *Phys. Rev. B* **55**, 11218-11224 (1997)
 - [47] Rabe, K.M., Ahn, C.H., Triscone, J.M.: *Physics of Ferroelectrics: A Modern Perspective*. Springer, Berlin Heidelberg (2007)
 - [48] Contreras, J.R., Kohlstedt, H., Poppe, U., Waser, R., Buchal, C., Pertsev, N.A.: Resistive switching in metal-ferroelectric-metal junctions. *Appl. Phys.*

- Lett. **83**, 4595-4597 (2003)
- [49] Roelofs, A., Schneller, T., Szot, K., Waser, R.: Piezoresponse force microscopy of lead titanate nanograins possibly reaching the limit of ferroelectricity. *Appl. Phys. Lett.* **81**, 5231-5233 (2002)
 - [50] Rupprecht, G., Bell, B.O.: Dielectric constant in paraelectric perovskites. *Phys. Rev.* **135**, A748-A752 (1964)
 - [51] Scott, J.F.: Applications of modern ferroelectrics. *Science*, **315**, 954-959 (2007)
 - [52] Shen, J., Ma, Y.Q.: Long-range coupling interactions in ferroelectric superlattices. *Phys. Rev. B* **61**, 14279-14282 (2000)
 - [53] Shen, J., Ma, Y.Q.: Long-range coupling interactions in ferroelectric sandwich structures. *J. Appl. Phys.* **89**, 5031-5035 (2001)
 - [54] Spanier, J.E., Kolpak, A.M., Urban, J.J., Grinberg, I., Ouyang, L., Yun, W.S., Rappe, A.M., Park, H.K.: Ferroelectric phase transition in individual single-crystalline BaTiO₃ nanowires. *Nano Lett.* **6**, 735-739 (2006)
 - [55] Streiffer, S.K., Eastman, J.A., Fong, D.D., Thompson, C., Munkholm, A., Murty, M.V.R., Auciello, O., Bai, G.R., Stephenson, G.B.: Observation of nanoscale 180° stripe domains in ferroelectric PbTiO₃ thin films. *Phys. Rev. Lett.* **89**, 067601 (2002)
 - [56] Tagantsev, A.K.: Landau expansion for ferroelectrics: Which variable to use? *Ferroelectrics* **375**, 19-27 (2008)
 - [57] Tagantsev, A.K., Gerra, G., Setter, N.: Short-range and long-range contributions to the size effect in metal-ferroelectric-metal heterostructures. *Phys. Rev. B* **77**, 174111 (2008)
 - [58] Tan, E.K., Osman, J., Tilley, D.R.: First-order phase transitions in ferroelectric films. *Solid State Commun.* **116**, 61-65 (2000)
 - [59] Tenne, D.A., Bruchhausen, A., Lanzillotti-Kimura, N.D., Fainstein, A., Katiyar, R.S., Cantarero, A., Soukiassian, A., Vaithyanathan, V., Haeni, J.H., Tian, W., Schlom, D.G., Choi, K.J., Kim, D.M., Eom, C.B., Sun, H.P., Pan, X.Q., Li, Y.L., Chen, L.Q., Jia, Q.X., Nakhmanson, S.M., Rabe, K.M., Xi, X.X.: Probing nanoscale ferroelectricity by ultraviolet Raman spectroscopy. *Science* **313**, 1614-1616 (2006)
 - [60] Timoshenko, S.P., Goodier, J.N.: *Theory of Elasticity*, 3rd ed. McGraw-Hill Publications, International Edition, Singapore (1970)
 - [61] Tsymbal, E.Y., Kohlstedt, H.: Tunneling across a ferroelectric. *Science* **313**, 181-183 (2006)
 - [62] Tybell, T., Ahn, C.H., Triscone, J.M.: Ferroelectricity in thin perovskite films. *Appl. Phys. Lett.* **75**, 856-858 (1999)
 - [63] Uchino, K.J., Sadanaga, E.J., Hirose, T.: Dependence of the crystal structure on particle size in barium titanate. *J. Am. Ceram. Soc.* **72**, 1555-1558 (1989)
 - [64] Umeno, Y., Meyer, B., Elsasser, C., Gumbsch, P.: Ab initio study of the critical thickness for ferroelectricity in ultrathin Pt/PbTiO₃/Pt films. *Phys. Rev. B* **74**, 060101 (2006)
 - [65] Velez, J.P., Duan, C.G., Belashchenko, K.D., Jaswal, S.S., Tsymbal, E.Y.: Effect of ferroelectricity on electron transport in Pt/BaTiO₃/Pt tunnel

- junctions. *Phys. Rev. Lett.* **98**, 137201 (2007)
- [66] Wang, B., Woo, C.H.: Curie temperature and critical thickness of ferroelectric thin films. *J. Appl. Phys.* **97**, 084109 (2005)
- [67] Wang, B., Woo, C.H.: Curie-Weiss law in thin-film ferroelectrics. *J. Appl. Phys.* **100**, 044114 (2006)
- [68] Wang, B., Woo, C.H., Zheng, Y.: Size limit of ferroelectric nanoislands on substrates. (Finished)
- [69] Wang, Y.G., Zhong, W.L., Zhang, P.L.: Lateral size effects on cells in ferroelectric films. *Phys. Rev. B* **51**, 17235-17238 (1995)
- [70] Woo, C.H., Zheng, Y.: Depolarization in modeling nano-scale ferroelectrics using the Landau free energy functional. *Appl. Phys. A* **91**, 59-63 (2008)
- [71] Yanase, N., Abe, K., Fukushima, N., Kawakubo, T.: Thickness dependence of ferroelectricity in heteroepitaxial BaTiO₃ thin film capacitors. *Jpn. J. Appl. Phys.* **38**, 5305-5308 (1999)
- [72] Zheng, Y., Woo, C.H.: Tuning the properties of a vertically composition-graded ferroelectric film using a mechanical load. *J. Phys. D: Appl. Phys.* **41**, 175403 (2008)
- [73] Zheng, Y., Woo, C.H.: Thermodynamic modeling of critical properties of ferroelectric superlattices in nano-scale. *Appl. Phys. A* **97**, 617-626 (2009a)
- [74] Zheng, Y., Woo, C.H.: Giant piezoelectric resistance in ferroelectric tunnel junctions. *Nanotechnology* **20**, 075401 (2009b)
- [75] Zheng, Y., Woo, C.H.: Hyper-sensitive piezophotovoltaic effects in ferroelectric nanocylinders. *J. Appl. Phys.* **107**, 104120 (2010)
- [76] Zheng, Y., Woo, C.H., Wang, B., Zhu, Z.Y.: Ferroelectric rods with adjustable dielectric tunability. *Appl. Phys. Lett.* **90**, 092905 (2007)
- [77] Zheng, Y., Woo, C.H., Wang, B.: Pulse-loaded ferroelectric nanowire as an alternating current source. *Nano Lett.* **8**, 3131-3136 (2008a)
- [78] Zheng, Y., Woo, C.H., Wang, B.: Surface tension and size effect in ferroelectric nanotubes. *J. Phys: Condens. Matter.* **20**, 135216 (2008b)
- [79] Zheng, Y., Wang, B., Woo, C.H.: Thermodynamic modeling of nanoscale ferroelectric systems. *Acta Mech. Solida Sin.* **22**, 524-549 (2009)
- [80] Zheng, Y., Cai, M.Q., Woo, C.H.: Critical properties of symmetric nanoscale metal-ferroelectric-metal capacitors. *Acta Mater.* **58**, 3050-3058 (2010)
- [81] Zhong, W.L., Wang, Y.G., Zhang, P.L., Qu, B.D.: Phenomenological study of the size effect on phase transitions in ferroelectric particles. *Phys. Rev. B* **50**, 698-703 (1994)
- [82] Zhou, Y., Shin, F.G.: Magnetoelectric effect of mildly conducting magnetostrictive/piezoelectric particulate composites. *J. Appl. Phys.* **100**, 043910 (2006)
- [83] Zhou, Z.H., Gao, X.S., Wang, J., Fujihara, K., Ramakrishna, S., Nagarajan, V.: Giant strain in PbZr_{0.2}Ti_{0.8}O₃ nanowires. *Appl. Phys. Lett.* **90**, 052902 (2007)
- [84] Zhuravlev, M.Y., Sabirianov, R.F., Jaswal, S.S., Tsymbal, E.Y.: Giant electroresistance in ferroelectric tunnel junctions. *Phys. Rev. Lett.* **94**, 246802 (2005)

Strain Engineering: Ferroelectric Films on Compliant Substrates

6.1 Background

In Chapter 5, we have discussed the unique physical properties of ferroelectric thin films, which generally exhibit characteristics that are different from their bulk counterparts. Due to the surface constraint on polarization, the para/ferroelectric transition temperatures of the film may be substantially lower than those of the bulk, so much so that the ferroelectric transformation may be totally suppressed in very thin films. The magnitude of this critical thickness can be calculated as shown in the last chapter. However, the theoretical calculations were performed for free standing or totally constrained films. As to the epitaxial films grown on compliant substrates, the influences of substrates over the critical thickness and Curie temperature of the film should be considered.

Substrates impart mechanical constraints on thin films, and may thus affect their phase transition characteristics (Wang and Woo, 2004), through the elastic interaction with the transformation strain. The effect of the substrates comes from these sources. Firstly, as shown in Fig. 6.1, there is the epitaxial stress generated during the fabrication process due to the misfit of the two lattices. This stress may assist, or obstruct, the transformation, and raise or lower the Curie temperature of the film, accordingly (Zembilgotov et al., 2002). Secondly, there is also the constraint on the transformation strain due to the presence of the substrate, which always obstructs the proceedings of the transformation, and affects the phase-transition characteristics, such as the order of the transition, the universal exponents, etc. At last, the substrate also affects the critical thickness for misfit dislocation generation.

In fact, there are some interesting experimental results showing that the substrates have the ability to affect the phase transition and critical characteristics of thin film. For example, Tybell et al. (1999) reported that PbZrTiO_3 films grown on metallic Nb-SrTiO_3 substrate exhibit a stable polarization orthogonal to the surface, even when the film is as thin as 4 nm. As shown in Fig. 6.2 and Fig. 6.3,

using EFM (electric field microscopy) and piezoelectric microscopy, the team observed the polarization signal. These results were particularly striking for 4-nm-thick films, where the measurements showed that a ferroelectric state was stable for over 140 h, the duration of these measurements, in accord with recent calculations that predicted ferroelectricity down to monolayer thicknesses (Tybell et al., 1999).

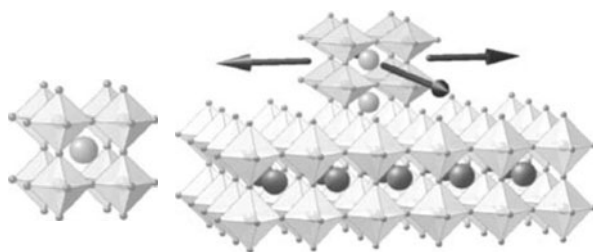


Fig. 6.1 Sketch map of epitaxial stress (Lee et al., 2010) Copyright (2010) by the Nature Publishing Group

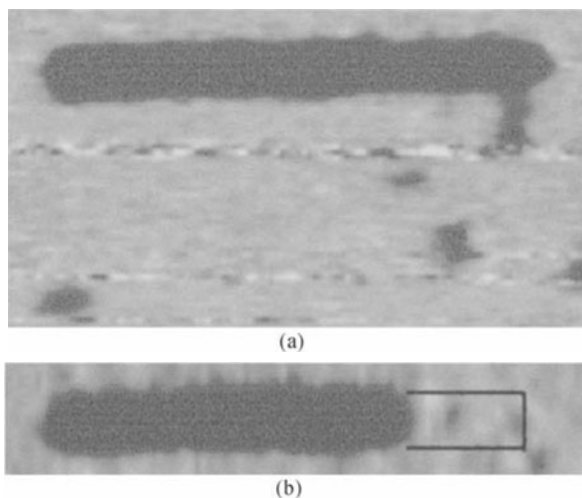


Fig. 6.2 Piezoelectric images of an 80 Å film surface. Originally, the polarization of the sample was uniform [light gray in the images] except for some regions [dark in (a)] larized oppositely. In (a) a line region was switched from “light to dark” polarization. (b) shows a zoom after switching back a part of the written line to its original polarization state (Tybell et al., 1999) Copyright (1999) by the American Institute of Physics

Moreover, the experimental results reported by Yanase et al. (1999) in the same year showed that BaTiO_3 films with different thicknesses grown on SrTiO_3 displayed different hysteresis loops, which are shown in Fig. 6.4. And even when the thickness of the films was reduced to 12 nm, the hysteresis loop can be clearly observed (Yanase et al., 1999). Pertsev et al. (1998) reported on the effect of mechanical boundary conditions on the polarization, dielectric constants and phase

diagram of ferroelectric film grown on a substrate. As shown in Fig. 6.5, they found that the phase diagrams of BaTiO_3 and PbTiO_3 epitaxially grown on different cubic substrates were very different with respect to the transitions between various low-temperature ferroelectric phases.

Zembilgotov et al. (2002) used mean-field Landau-type theory to analyze the polarization properties of epitaxial ferroelectric thin film grown on dissimilar cubic substrates, in which they induced biaxial compressive stresses in the film plane. The theory simultaneously allowed for the influence of the misfit strain imposed on the film lattice by a thick substrate. The mean polarization in ferroelectric thin film was calculated as a function of film thickness, temperature, and misfit strain. It showed that the stress generated due to this misfit strain might assist, or obstruct, the transformation and raise or lower the transition temperature of the film, accordingly, which is shown in Fig. 6.6.

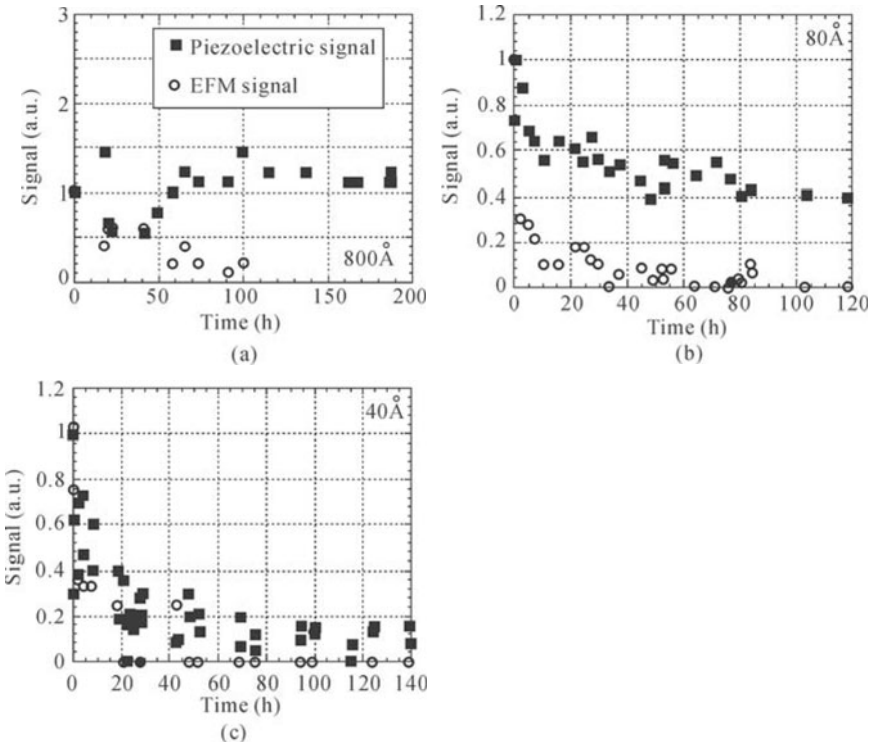


Fig. 6.3 Time dependencies of the EFM and piezoelectric signals measured on lines (a) 800-Å-thick film, (b) 80-Å-thick film and (c) 40-Å-thick film. We notice that while the EFM signal goes to zero, the piezoelectric signal saturates at a nonzero level (Tybell et al., 1999). Copyright (1999) by the American Institute of Physics

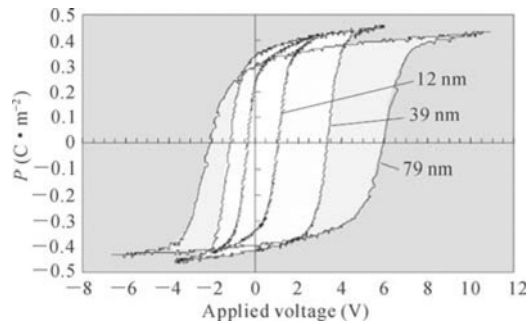


Fig. 6.4 P - V hysteresis loops observed for heteroepitaxial BaTiO_3 films with thickness of 12 nm, 39 nm and 79 nm (Yanase et al., 1999). Copyright (1999) by the Japan Society of Applied Physics

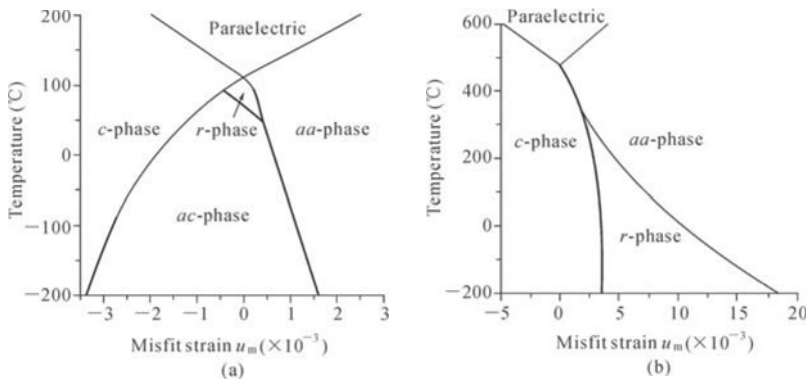


Fig. 6.5 Phase diagrams of (001) single-domain BaTiO_3 (a) and PbTiO_3 (b) thin films epitaxially grown on different cubic substrates providing various misfit strains u_m in the heterostructures. The second- and first-order phase transitions are shown by thin and thick lines, respectively (Pertsev and Zembilgotov, 1998). Copyright (1998) by the American Physics Society

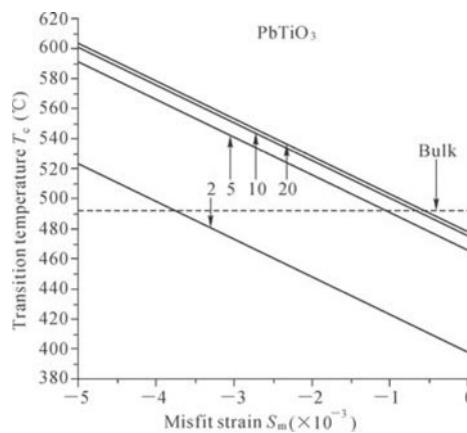


Fig. 6.6 Ferroelectric transition temperature T_c calculated as a function of the misfit strain for PbTiO_3 films of several different thicknesses. Numbers indicate the normalized film thickness for respective curves. Dashed line shows the transition temperature of a bulk crystal (Zembilgotov and Pertsev, 2002). Copyright (2002) by the American Institute of Physics

A series of studies have been done by Wang, Woo and Zheng, and this chapter is mainly based on their work. By treating the para/ferroelectric transition as a point of instability of the time-dependent Ginzburg-Landau equation, Wang and Woo (2004) found that the order of the transition is sensitively dependent on the relative thickness of the film and substrate. They also predicted the change in the transition temperature versus the film and substrate thickness, and determined the critical thickness of the film in Wang et al. (2005).

Later, Zheng et al. (2008) also considered the dynamic piezoelectric bending response and switching behavior of ferroelectric/paraelectric bilayers due to an applied electric field. The bending response was found to be very large, and was highly adjustable by varying the relative thicknesses of the two layers. The results showed ferro/paraelectric bilayers might have very good potential for such applications as transducers, sensors and actuators.

The existence of a critical thickness for dislocation generation in elastic thin films was first proposed by Frank and Merwe (1949) and it was subsequently confirmed experimentally and investigated theoretically. Wang et al. (2003) extended the analysis to piezoelectric thin films, and discussed the effect of piezoelectric properties on the critical thickness.

In addition, based on the Landau-Devonshire formalism and Timosheko's method for thermal stresses, the critical thickness for misfit dislocation generation in a ferroelectric thin film on a compliant substrate was established by Zheng et al. (2006). The effects of the polarization and the misfit stresses on the critical thickness for misfit dislocation generation were discussed.

In this chapter, we will focus on ferroelectric thin films on a compliant substrate. Combining solid mechanics with the Landau phase transition theory, we can see that the characteristics of the phase transition, such as *the transition temperature, the critical thickness, the order of the phase transition*, and even *the critical thickness for dislocation generation*, can be tuned simply by changing the relative thickness of the film and substrate.

6.2 Manipulation of Phase Transition Behavior of Ferroelectric Thin Films on Compliant Substrates

In this section, we will consider ferroelectric films on a compliant substrate (Fig. 6.7) and discuss the behavior of the film near the transition temperature. The main idea is to establish the evolution equation and treat the para/ferro electric transition as its point of instability. By determining the bifurcation point of the evolution equation, the phase transition temperature can be obtained. Then, by using a perturbation approach, the effect of film dimensions and surface conditions on the system behavior of ferroelectric thin films near the paraelectric-ferroelectric phase transition can be investigated.



Fig. 6.7 Schematic of ferroelectric thin film and substrate system

6.2.1 Free Energy Expressions

Consider a ferroelectric thin film and substrate system (Fig. 6.7). Both the film and the substrate are treated as a cubic elastic body with elastic constants $C_{11}, C_{12}; C_{11}^0, C_{12}^0$ and thickness h, H , respectively. The polarization is assumed to be normal to the surface of the film. Due to the constraint of the substrate and the inhomogeneous distribution of the polarizations in the film, characteristics of the phase transition may be affected, such as the transition temperature, or even the order of the transitions, etc.

To establish the evolution equation, we must find the total free energy of the system. The total free energy consists of three parts, elastic strain energy U_s induced by the transformation strain, interaction energy U_c and intrinsic ferroelectric free energy U_p .

• Elastic strain energy

Generally speaking, the ferroelectric phase transition is caused by a change in the crystallographic structures associated with a transformation strain. If the polarization P is along one direction, the transverse transformation strain can be expressed as:

$$\varepsilon^T = QP^2, \quad (6.2.1)$$

where Q is the electrostrictive coefficient. For cubic free-standing film, the transformation strain on the plane of the film is

$$\varepsilon_{xx}^T = \varepsilon_{yy}^T = QP^2, \quad (6.2.2)$$

from which the elastic stress can be derived. First fix the square film element to allow the transformation to happen and it will create a normal stress in the film

$$\sigma_{xx}^T = \sigma_{yy}^T = -GQP^2, \quad (6.2.3)$$

where

$$G = (C_{11} + C_{12} - 2C_{12}^2 / C_{11}) . \quad (6.2.4)$$

Then, to satisfy the force-free conditions, one must apply the negative axial force $-\frac{1}{h} \int_{-h/2}^{h/2} \sigma_{xx}^T dz$ on the y -boundary and x -boundary of the element. Thus, the total stress in the film becomes

$$\sigma_{xx} = \sigma_{yy} = -GQP^2 + \frac{GQ}{h} \int_{-h/2}^{h/2} P^2 dz . \quad (6.2.5)$$

Then, the elastic strain energy can be obtained. It is due to the inhomogeneous distribution of P along the thickness direction. For 2D clamping film, the second term in the stress expression is zero.

The elastic strain energy induced by the transformation strain can then be calculated,

$$\begin{aligned} U_s &= \frac{1}{2} \int_0^{H+h} \frac{\sigma_{xx}^2 + \sigma_{yy}^2}{G} dz \\ &= \frac{1}{G} \int_0^h \sigma^2(z) dz + \frac{1}{G_0} \int_h^{H+h} \sigma^2(z) dz \\ &= GQ^2 \int_0^h P^4 dz + g_1 \left[\int_0^h P^2 dz \right] + g_2 \left[\int_0^h P^2(z-a) dz \right]^2 \\ &\quad + g_3 \int_0^h P^2 dz \int_0^h P^2(z-a) dz, \end{aligned} \quad (6.2.6)$$

where

$$\begin{aligned} g_1 &= Q^2 \left[\frac{Gh + HG^2 / G_0}{(H+h)^2} - \frac{2G}{H+h} \right], \\ g_2 &= Q^2 \left\{ -\frac{2G}{I} + \frac{G}{3I^2} [a^3 + (h-a)^3] + \frac{G^2}{3G_0 I^2} [(H+h-a)^3 - (h-a)^3] \right\}, \\ g_3 &= \frac{Q^2}{(H+h)I} \left\{ G[-a^2 + (h-a)^2] + \frac{G^2}{G_0} [(H+h-a)^2 - (h-a)^2] \right\}, \end{aligned} \quad (6.2.7)$$

where $I = (H+h)^3 / 12$ is the area moment of inertia, and the neutral axis can be determined as follows:

$$a = \frac{\int_0^{H+h} G(z)z dz}{\int_0^{H+h} G(z) dz} = \frac{Gh^2 + G_0 H^2 + 2G_0 Hh}{2(Gh + G_0 H)} . \quad (6.2.8)$$

• Interaction energy

During fabrication, the misfit between the film and the substrate creates a misfit stress $\sigma^0 = \sigma_{xx} = \sigma_{yy}$ in the film. If the misfit strain is $\varepsilon^0 = (a_s - a_f) / a_f$ where a_s and a_f are the lattice constants of the substrate and the film, respectively, the misfit stress in the film can be derived as

$$\sigma^0 = G\varepsilon^0 \left[\frac{h}{H+h} + \frac{h^2 - 2ah}{2I} (z-a) - 1 \right]. \quad (6.2.9)$$

For thick substrates, the first two terms in Eq. (6.2.9) vanish as $H \rightarrow \infty$ and only the last term remains.

The misfit stresses here play the role of an applied load. And the interaction energy between them and the transformation strains is given as

$$U_e = - \int_0^h (\sigma_{xx} \varepsilon^T + \sigma_{yy} \varepsilon^T) dz = -2Q \int_0^h \sigma^0 P^2 dz. \quad (6.2.10)$$

• Intrinsic free energy

According to the Landau phase transition theory, the free energy associated with the ferroelectric thin film could be written as

$$U_p = \int_0^h \left[\frac{A}{2} (T - T_{c0}) P^2 + \frac{B}{4} P^4 + \frac{C}{6} P^6 + \frac{D}{2} \left(\frac{\partial P}{\partial z} \right)^2 \right] dz + \int_S \frac{D\delta^{-1}}{2} P^2 dx dy, \quad (6.2.11a)$$

where A , B , C , and D are expansion coefficients of the corresponding bulk material. T_{c0} is the transition temperature of the bulk crystal, S represents the upper and lower unit surface planes of the film, and δ is the extrapolation length that measures the strength of the surface effect on the polarization (Wang and Woo, 2005). If there exists an external electric field E_{ext} and a depolarizing field E_d , the energy caused by the electric fields should be added to U_p and Eq. (6.2.11a) can be written as

$$U_p = \int_0^h \left[\frac{A}{2} (T - T_{c0}) P^2 + \frac{B}{4} P^4 + \frac{C}{6} P^6 + \frac{D}{2} \left(\frac{\partial P}{\partial z} \right)^2 - \frac{1}{2} E_d P - E_{ext} P \right] dz + \int_S \frac{D\delta^{-1}}{2} P^2 dx dy, \quad (6.2.11b)$$

where E_d depends on the electric boundary condition, which can be expressed in the following form (Pertsev et al., 1988; Wang and Woo, 2004):

$$E_d = -\frac{1}{\alpha} \left(P - \frac{\phi}{h} \int_0^h P dz \right), \quad (6.2.12)$$

where α is the dielectric constant of the film and ϕ takes on the value of 0 or 1. The ferroelectric film is between two metallic electrodes in short-circuit conditions; $\phi = 0$ refers to the fact that the ferroelectric film is between two dielectrics.

6.2.2 Evolution Equations

By using the energy Eqs. (6.2.6), (6.2.10) and (6.2.11), the evolution equation can be established through the following variational derivative,

$$\begin{aligned} -K \frac{\partial P}{\partial t} &= \frac{\delta(U_s + U_e + U_p)}{\delta P} \\ &= (A_1 z + A_2)P + A_3 P^3 + CP^5 - D \frac{d^2 P}{dz^2} - R, \end{aligned} \quad (6.2.13)$$

where K is the kinetic coefficient related to the domain wall mobility and

$$\begin{aligned} A_1 &= 4g_2 A_2 + 2g_3 A_1 - \frac{2QG\varepsilon^0}{I} (h^2 - 2ah), \\ A_2 &= A(T - T_{c0}) + \frac{1}{\alpha} - 4QG\varepsilon^0 \left(\frac{h}{H+h} - \frac{ah^2 - 2a^2 h}{2I} - 1 \right) \\ &\quad + (4g_1 - 2ag_3)A_1 + (2g_3 - 4ag_2)A_2, \\ A_3 &= B + 4GQ^2, \\ R &= \frac{\phi}{\alpha h} \int_0^h P dz, \end{aligned} \quad (6.2.14)$$

with

$$A_1 = \int_0^h P^2 dz, \quad A_2 = \int_0^h P^2 (z-a) dz. \quad (6.2.15)$$

The external electric field E_{ext} is not considered here, and the boundary conditions are

$$\frac{\partial P}{\partial z} = \begin{cases} \frac{P}{\delta} & \text{for } z = 0; \\ -\frac{P}{\delta} & \text{for } z = h. \end{cases} \quad (6.2.16)$$

At steady state, Eq. (6.2.13) satisfies $\frac{\partial P}{\partial t} = 0$, thus giving

$$\begin{aligned} L_c P &\equiv -(\lambda_1 z + \lambda_{2c})P + D \frac{d^2 P}{dz^2} + \frac{\phi}{\alpha h} \int_0^h P dz \\ &= A(T - T_c)P + \lambda_3 P^3 + CP^5 + (4g_2 \Delta_2 + 2g_3 \Delta_1)zP \\ &\quad + [(4g_1 \Delta_2 - 2ag_3)\Delta_1 + (2g_3 - 4ag_2)\Delta_2]P, \end{aligned} \quad (6.2.17)$$

where

$$\begin{aligned} \lambda_1 &= -\frac{2QG\varepsilon^0}{I}(h^2 - 2ah), \\ \lambda_{2c} &= A(T_c - T_{c0}) + \frac{1}{\alpha} - 4QG\varepsilon^0 \left(\frac{h}{H+h} - \frac{ah^2 - 2a^2h}{2I} - 1 \right), \end{aligned} \quad (6.2.18)$$

where T_c is the transition temperature of the film.

In the following, through determining the bifurcation point of Eq. (6.2.13), one can obtain the phase transition temperature. Then, by using a perturbation approach, the effect of film dimensions and surface conditions on the system behavior of ferroelectric thin films near the paraelectric-ferroelectric phase transition can be investigated. Based on these results, one can find that the phase transition characteristics of ferroelectric films, such as the transition temperature, the order of transition, etc., can be tuned only through changing the ‘extrinsic conditions’ such as the thickness of the film and substrate.

6.2.3 Manipulation of Ferroelectric Transition Temperature and Critical Thickness

In this section we will introduce the manipulation of transition temperature and critical thickness of ferroelectric thin films on compliant substrates. Because the critical thickness is the thickness that makes the transition temperature $T_c=0$, if we can determine the transition temperature as a function of the relative thickness of the film and substrate, we can find the critical thickness. So we shall focus on the transition temperature, or the supercooling temperature for the first-order transition. Therefore, it is enough to consider only the linearized Ginzburg-Landau equation. We have carried out extensive calculation, and found that the effect of the last three terms of the elastic energy related to the phase transition in Eq. (6.2.6) on the transition temperature can be neglected, and the main contributions of the elastic energy are on the higher order terms of the Landau free energy expansion.

Therefore, the simplified linearized evolution equation corresponding to the energy Eqs. (6.2.6), (6.2.10) and (6.2.11) can be established through the following variational derivative,

$$-K \frac{\partial P}{\partial t} = \frac{\delta(U_s + U_e + U_p)}{\delta P} = (A_1 z + A_2)P - D \frac{d^2 P}{dz^2} - R, \quad (6.2.19)$$

where K is the kinetic coefficient related to the domain wall mobility

$$\begin{aligned} A_1 &= -\frac{2QG\epsilon^0}{I}(h^2 - 2ah), \\ A_2 &= A(T - T_{c0}) + \frac{1}{\alpha} - 4QG\epsilon^0 \left(\frac{h}{H+h} - \frac{ah^2 - 2a^2h}{2I} - 1 \right), \\ R &= \frac{\phi}{\alpha h} \int_0^h P \, dz, \end{aligned} \quad (6.2.20)$$

and the boundary conditions are the same as Eq. (6.2.16).

It is obvious that $P = 0$ is a trivial solution of Eq. (6.2.17), representing the stationary non-ferroelectric state. If the temperature is reduced below some critical value, the trivial zero solution will bifurcate yielding $P \neq 0$ solutions, corresponding to the transformation of the paraelectric state to the ferroelectric state. Therefore, the transition temperature can be established through the bifurcation point. From the bifurcation theory of the partial differential equation, the bifurcation point can be determined from the eigenvalue λ of the linearized equation

$$(A_1 z + A_2 + \lambda)P - D \frac{d^2 P}{dz^2} - R = 0, \quad (6.2.21)$$

where the eigenvalue means λ values for which non-zero solutions of Eq. (6.2.21) and boundary condition Eq. (6.2.16) exist. The bifurcation point, at the transition temperature, is the point at which $\lambda = 0$, where the stable $P=0$ solution starts to become unstable. Since R is a constant, we can derive the solution for the homogenous equation first, and then determine R . The equation is simplified with the changed variable

$$\begin{aligned} \eta &= \left(\frac{D}{A_1}\right)^{2/3} \frac{1}{D} (A_1 z + A_2 + \lambda), \\ \frac{d^2 P}{d\eta^2} - \eta P &= -R. \end{aligned} \quad (6.2.22)$$

The general solution of Eq. (6.2.36) is a linear combination of the two Airy functions $Ai(\eta)$, $Bi(\eta)$, therefore

$$P(\eta) = C_1 P_1(\eta) + C_2 P_2(\eta), \quad (6.2.23)$$

where

$$\begin{aligned} P_1(\eta) &= Ai(\eta) + \Omega [F_{Ai}(\eta)Bi(\eta) - F_{Bi}(\eta)Ai(\eta)] \int_{\eta_1}^{\eta_2} Ai(\xi) d\xi, \\ P_2(\eta) &= Bi(\eta) + \Omega [F_{Ai}(\eta)Bi(\eta) - F_{Bi}(\eta)Ai(\eta)] \int_{\eta_1}^{\eta_2} Bi(\xi) d\xi, \end{aligned} \quad (6.2.24)$$

where

$$\begin{aligned} \Omega &= \frac{\pi}{\alpha h} \left(\frac{D}{A_1} \right)^{1/3} \left\{ 1 - \frac{\pi}{\alpha h A_1} \int_{\eta_1}^{\eta_2} [F_{Ai}(\xi)Bi(\xi) - F_{Bi}(\xi)Ai(\xi)] d\xi \right\}^{-1}, \\ \eta_1 &= \left(\frac{D}{A_1} \right)^{2/3} \frac{1}{D} (A_2 + \lambda), \eta_2 = \left(\frac{D}{A_1} \right)^{2/3} \frac{1}{D} (A_1 h + A_2 + \lambda), \end{aligned} \quad (6.2.25)$$

and $F_{Ai}(\eta)$, $F_{Bi}(\eta)$ are original functions of $Ai(\eta)$, $Bi(\eta)$, respectively.

Substitution of Eq. (6.2.23) into the first one of the boundary conditions Eq. (6.2.16) yields a relation between the constants C_1 and C_2 as follows:

$$C_2 = \gamma C_1, \quad (6.2.26)$$

where

$$\gamma = \left[\frac{1}{\delta} P_2(\eta_1) - \left(\frac{A_1}{D} \right)^{1/3} P_2'(\eta_1) \right]^{-1} \left[\left(\frac{A_1}{D} \right)^{1/3} P_1'(\eta_1) - \frac{1}{\delta} P_1(\eta_1) \right], \quad (6.2.27)$$

where $P_1'(\eta_1)$, $P_2'(\eta_1)$ are derivatives with respect to η at $\eta = \eta_1$, respectively. Therefore Eq. (6.2.23) becomes

$$P(\eta) = C_1 [P_1(\eta) + \gamma P_2(\eta)]. \quad (6.2.28)$$

Substitution of Eq. (6.2.28) into the second boundary condition of Eq. (6.2.16) yields

$$P_1'(\eta_2) + \gamma P_2'(\eta_2) = -\frac{1}{\delta} [P_1(\eta_2) + \gamma P_2(\eta_2)]. \quad (6.2.29)$$

To obtain the transition temperature, one can set the eigenvalue

$$\lambda = 0, \quad \text{and} \quad -\omega = A(T - T_{c0}) + \frac{1}{\alpha}. \quad (6.2.30)$$

By solving ω versus the thickness of the film numerically, the transition temperature can be obtained as

$$T_c = T_{c0} - \frac{1}{A\alpha} - \frac{\omega_{\min}}{A}, \quad (6.2.31)$$

where $\omega_{\min}$ means the minimum solution of ω . Taking $T_c = 0$, one can determine the critical thickness.

Within the present formulation, if the substrate thickness $H \rightarrow \infty$, Eq. (6.2.20) gives

$$A_1 = 0, \quad A_2 = A(T - T_{c0}) + \frac{1}{\alpha} + 4QG\varepsilon^0, \quad (6.2.32)$$

and Eq. (6.2.21) reduces to the same form as introduced in Chapter 5, which is also based on Wang and Woo (2005). For a free standing thin film, and when the polarization distribution is asymmetric with respect to its center axis, the misfit stress is zero in such cases. One can determine the transition temperature in the same way, as shown in Chapter 5. These two cases represent the two limiting cases for the substrate effects. With a substrate, usually one cannot get rid of the misfit stress in the fabrication process. Depending on the combination of the film and the substrate, the misfit stress may help the ferroelectric transformation or suppress the transformation, which means the misfit stress may increase or decrease the transition temperature of the film. Therefore, one can tune the transition temperature of the film by changing these external factors, such as the thicknesses of the film and substrate, in addition to changing the internal factors, such as the material's composition of the film.

To consider the effect of the substrate on the transition temperature of the ferroelectric thin film, the transition temperature of PbTiO_3 thin film on SrTiO_3 substrate was calculated using Eq. (6.2.31) as a function of the thickness of the film and substrate, respectively, and the polarization was assumed to be perpendicular to the film surfaces. The calculation result of the transition temperature versus the thickness of the film is shown in Fig. 6.8 under different thicknesses of the substrate.

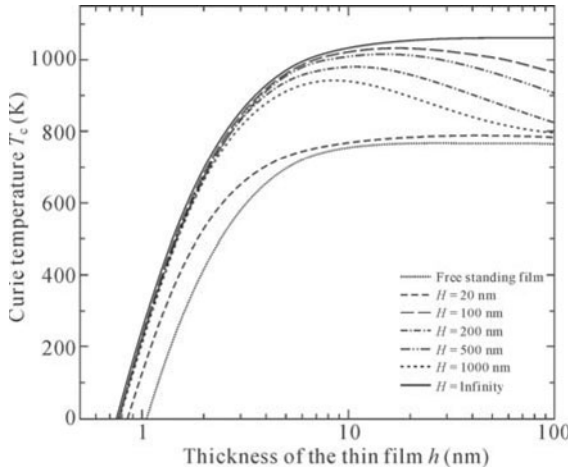


Fig. 6.8 The transition temperature of PbTiO_3 film on SrTiO_3 substrate versus the film thickness under different substrate thicknesses

It can be found that the transition temperature is changing between the two limiting cases: the upper limit is for the totally constrained film with largest misfit stress, whereas the lower limit is for the free standing film with almost zero misfit stress. One can tune the transition temperature in such a range just by controlling the substrate thickness. With the consideration that the transition temperature will generally decrease with a decrease in the film thickness, by adjusting the thickness of the film and substrate one can tune the transition temperature over quite a large range.

6.2.4 Manipulation of the Order of Transition

In this section, by analyzing the solution of the evolution Eq. (6.2.13), or the time-dependent Ginzburg-Landau (TDGL) equation, we will find that the order of the transition can be determined as a function of the relative thickness of the film and substrate. We will also discuss the critical ratio of the thickness of the film to that of the substrate, below which thin films of first-order transition materials will undergo a second-order transition.

We can see that $P=0$ is a trivial solution of Eq. (6.2.17), and T_c marks a point of instability of that solution. We are interested in finding the condition for the existence of a real, stable, continuous and non-zero solution in the neighborhood of the point of instability. We note that this is also the condition for the transition to be second order in nature.

If the solution P exists, continuity guarantees that in a finite neighborhood of T_c , it can be assumed to be small. Both P and $\gamma = T - T_c$ can be expanded in a

power series in terms of a small perturbation ε from the critical point (Nicolis and Prigogine, 1977):

$$\begin{aligned} P &= \varepsilon P_1 + \varepsilon^2 P_2 + \dots, \\ \gamma &= T - T_c = \varepsilon \gamma_1 + \varepsilon^2 \gamma_2 + \dots \end{aligned} \quad (6.2.33)$$

This expansion is more flexible than the seemingly more natural one in which P is expanded in a power series of $(T - T_c)$. More importantly, it allows fractional power dependence of P on $(T - T_c)$. By substituting Eq. (6.2.33) into Eq. (6.2.17), and equating coefficients of equal powers of ε , a set of relations of the following form can be obtained:

$$L_c P_k = a_k, \quad k = 1, 2, \dots, \quad (6.2.34)$$

which have to be satisfied together with the boundary conditions Eq. (6.2.16). The first several coefficients a_k are:

$$\begin{aligned} a_1 &= 0, \quad a_2 = A \gamma_1 P_1, \\ a_3 &= A \gamma_1 P_2 + A \gamma_2 P_1 + (B + 4GQ^2) P_1^3 \\ &\quad + (4g_1 - 2ag_3 + 2g_3 z) P_1 \int_0^h P_1^2 \, dz \\ &\quad + (2g_3 - 4ag_2 + 4g_2 z) P_1 \int_0^h P_1^2 (z - a) \, dz. \end{aligned} \quad (6.2.35)$$

It is easy to check that under the boundary condition Eq. (6.2.16), the operator L_c is self-adjoint and

$$\begin{aligned} (L_c P^*, P_k) &= \int_{-h/2}^{h/2} P_k L_c P^* \, dz = (P^*, L_c P_k) \\ &= (P^*, a_k) = \int_{-h/2}^{h/2} P^* a_k \, dz = 0, \end{aligned} \quad (6.2.36)$$

where P^* is the solution of the following homogeneous equation:

$$\begin{aligned} L_c P^* &= 0; \\ \frac{dP^*}{dz} &= \mp \frac{P^*}{\delta}, \quad z = \pm \frac{h}{2}. \end{aligned} \quad (6.2.37)$$

Eq. (6.2.36) can be used to determine the coefficients γ_i . Then, from the second relation of Eq. (6.2.33), one can determine ε as a function of $(T - T_c)$, substituting the resulting ε into the first relation of Eq. (6.2.33) and solving the inhomogeneous

Eq. (6.2.34) result in an explicit expression for the solution P .

Eq. (6.2.18) is simplified with the changed variable

$$\eta = \left(\frac{D}{\lambda_1} \right)^{2/3} \frac{1}{D} (\lambda_1 z + \lambda_{2c}),$$

$$\frac{d^2 P^*}{d\eta^2} - \eta P^* = -R \lambda_1^{-2/3} D^{-1/3}. \quad (6.2.38)$$

The general solution of Eq. (6.2.38) is a linear combination of the two Airy functions $Ai(\eta)$, $Bi(\eta)$. Therefore

$$P^*(\eta) = C_1 P_1(\eta) + C_2 P_2(\eta), \quad (6.2.39)$$

where

$$P_1(\eta) = Ai(\eta) + R_0, \quad P_2(\eta) = Bi(\eta) + R_0, \quad (6.2.40)$$

and R_0 is a constant given by

$$R_0 = \frac{\pi}{\alpha h} \left(\frac{\lambda_1}{D} \right)^{-1/3} \int_{\eta_1}^{\eta_2} Ai(\eta) d\eta \left[Ai(\eta) \int_{\eta_1}^{\eta_2} Bi(\xi) d\xi - Bi(\eta) \int_{\eta_1}^{\eta_2} Ai(\xi) d\xi \right], \quad (6.2.41)$$

where $\eta_1 = (D/\lambda_1)^{2/3} \frac{1}{D} \lambda_{2c}$, $\eta_2 = (D/\lambda_1)^{2/3} \frac{1}{D} (\lambda_{2c} + \lambda_1 h)$.

Substitution of Eq. (6.2.39) into the boundary conditions Eq. (6.2.37) yields a relation between the constants C_1 and C_2

$$C_2 = \gamma C_1, \quad (6.2.42)$$

where

$$\gamma = \left[\frac{1}{\delta} P_2(\eta_1) - \left(\frac{A_1}{D} \right)^{1/3} P_2'(\eta_1) \right]^{-1} \left[\left(\frac{A_1}{D} \right)^{1/3} P_1'(\eta_1) - \frac{1}{\delta} P_1(\eta_1) \right]. \quad (6.2.43)$$

Here $P_1'(\eta_1)$, $P_2'(\eta_1)$ are derivatives with respect to η at $\eta = \eta_1$.

By using the orthogonal condition Eq. (6.2.36), we obtain

$$\gamma_1 = 0, \quad (6.2.44)$$

$$\begin{aligned}
 \gamma_2 = & -\frac{1}{A} \frac{\int_{-h/2}^{h/2} P_1^4 dz}{\int_{-h/2}^{h/2} P_1^2 dz} (B + 4GQ^2) \\
 & - \frac{1}{A \int_{-h/2}^{h/2} P_1^2 dz} \left[\int_0^h (4g_1 - 2ag_3 + 2g_3z) P_1^2 dz \int_0^h P_1^2 dz \right. \\
 & \left. + \int_0^h (2g_3 - 4ag_2 + 4g_2z) P_1^2 dz \int_0^h P_1^2 (z-a) dz \right].
 \end{aligned} \quad (6.2.45)$$

Substituting into Eq. (6.2.33) and neglecting the higher order terms, we obtain

$$\varepsilon \simeq \pm \sqrt{\frac{T - T_c}{\gamma_2}}. \quad (6.2.46)$$

From Eq. (6.2.33), for P to be real and physical, γ_2 and $T - T_c$ must have the same sign. Thus, in the ferroelectric regime ($T < T_c$), a continuous solution exists only if $\gamma_2 < 0$. Equivalently, if $\gamma_2 > 0$, a real and physical solution cannot be continuous in the ferroelectric regime.

In the case $\gamma_2 = 0$, one has to consider the higher order terms of the expansions, and thus the critical exponents will differ from their bulk counterpart. Indeed, the critical system behavior obtained from the foregoing analysis can also be understood in another way. When $\gamma_2 < 0$, a real and continuous solution exists only if $T < T_c$, and bifurcation analysis (Chow and Hale, 1982) shows that both branches of the solutions are asymptotically stable. The corresponding phase diagram is shown schematically in Fig. 5.11. When $\gamma_2 > 0$, real and continuous solutions exist only in the $T > T_c$ regime, but both branches of the solution are found to be unstable and are therefore physically unimportant. Nevertheless, it is reasonable to argue, from physical considerations, that away from the neighborhood of the critical point, there should be a maximum value of $T_{\max}$ above which only one solution (the trivial $P = 0$ solution) is admissible. Therefore, the subcritical branches should turn in the direction of decreasing temperature at some value $T_{\max}$ (Fig. 5.11), and join the discontinuous solution in the $T < T_c$ regime.

The foregoing shows that the critical system behavior of a thin film is determined by the sign of γ_2 , and is therefore not necessarily the same as that of the bulk material, which is completely governed by the sign of B . The sign of γ_2 depends, in a complicated way, on the bulk material parameters B , G , Q , film thickness h , substrate thickness and the surface conditions, through Eq. (6.2.18). Some general observations can be made. The condition $B > 0$ also implies $\gamma_2 < 0$, and hence the transition of a thin film is always second order with a critical

exponent of $1/2$, if the bulk material is second-order. This is independent of the film thickness, substrate thickness and surface conditions considered here.

The interesting case is when $B < 0$. γ_2 may then take either sign for different film and substrate thicknesses and surface conditions. Thus, if γ_2 is negative, the ferroelectric transition of the film is second-order and continuous, despite being first-order and discontinuous in the bulk.

An example can be found in a well-known first-order ferroelectric material PbTiO_3 . Plotted in Fig. 6.9a is γ_2 / C_1^2 vs. the ratio of the film thickness over the substrate thickness. It can be seen that if one can adjust the ratio h/H to less than a certain value, one can change the first-order transition into second-order. To graphically show the effect of the substrate, a phase diagram is constructed (Fig. 6.9b), where the vertical axis is the elastic modulus G_0 of the substrate, the horizontal axis is h/H and the properties of the film remain the same.

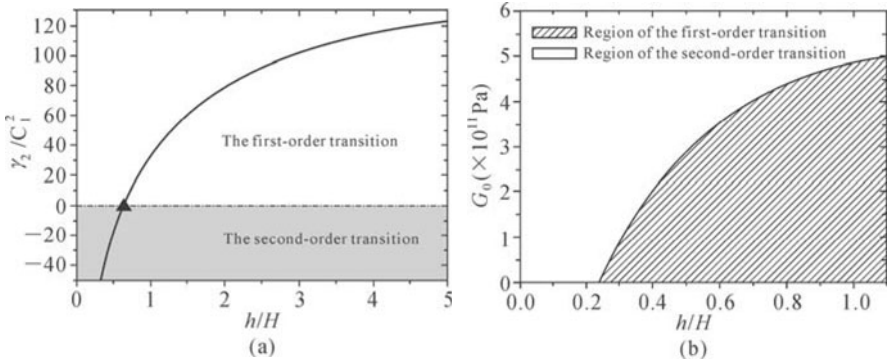


Fig. 6.9 Phase diagram of a ferroelectric film on a compliant substrate

6.3 Piezoelectric Bending Response and Switching Behavior of Ferroelectric Thin Film with Compliant Paraelectric Substrate

The last two sections have shown that the substrate significantly affects phase transition properties of ferroelectric films in general, changing the transition temperature, and causing a first-order ferroelectric film to undergo second-order phase transition in thin film form.

In this section, we consider the effects of film/substrate thicknesses on the polarization distribution, piezoelectric bending response and switching behavior of a ferroelectric/paraelectric (FE/PE) bilayer system. We are particularly interested in the relative vertical displacement of the bending bilayer as a function of the geometry and the actuating field.

6.3.1 Model of Ferroelectric Thin Film with Compliant Paraelectric Substrate and the Energy Expressions

Consider the FE/PE bilayer system as shown in Fig. 6.10a, where h and H are the thicknesses of the respective layers, and L is the common lateral dimension. The coordinate system is defined such that the interface is at $z = 0$, and the free surfaces at $z = h$ and $z = -H$, respectively.

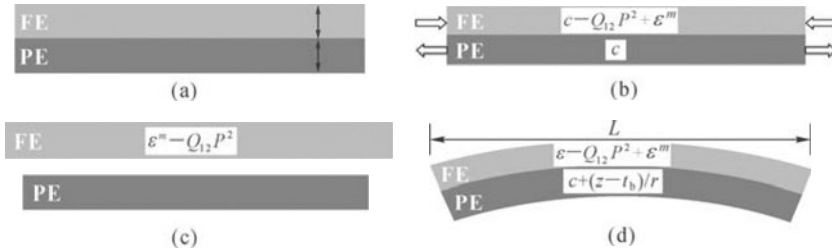


Fig. 6.10 Schematic diagrams showing (a) the FE/PE bilayer system, (b) unconstrained strains due to the misfit strain and self-phase-transformation strains, (c) constrained strains due to main displacement compatibility, and (d) bending induced by asymmetric stresses

As in our previous formulation, the ferroelectric system is considered as an assembly of polar molecules (crystallographic unit cells) occupying a volume V in space.

To simplify discussions, we consider the one-dimensional problem in which the polarization is directed along the plate thickness such that $P_1 = P_2 = 0$ and $P_3 = P(z)$, i.e., independent of x and y . The total free energy per unit area of the bilayer can then be written as

$$U = U_0 + U_{\text{Ela}} + U_p = U_0 + \int_{-H}^h (u_{\text{Ela}}(z) + u_p(z)) dz, \quad (6.3.1)$$

where U_p is the intrinsic free energy of the reference state. U_{Ela} is the contribution due to actions on the reference state by mechanical stresses in the finite crystal. u_p and u_{Ela} are the corresponding energy densities. U_0 is the free energy component that is independent of the polarization. To concentrate on the mechanical behavior of the system in this paper, we consider cases where surface effects such as depolarization and surface relaxation (Wang and Woo, 2005) are negligible. At the same time, we will only discuss mechanical behavior of the FE/PE system far away from the phase transition temperature T_c , which is determined mainly by the misfit between film and substrate. For simplicity, the effect of the induced polarization on the mechanical behavior is also neglected (Kim et al., 2005).

We first consider the elastic energy contribution U_{Ela} due to the lattice misfit of the materials in the bilayer. Let a_s be the lattice constant of PE layer and a_f the equivalent cubic cell constant of the free-standing FE layer. Assuming a coherent interface, the biaxial in-plane misfit strain in the film ε^0 can be written as $\varepsilon_{11}^0 = \varepsilon_{22}^0 = \varepsilon^0 = (a_s - a_f)/a_f$.

In addition to the misfit, we also have a transformation strain associated with the self-polarization generated by the lattice instability. In the present case, the polarization P is along the z direction and the transformation strain can be expressed as $\varepsilon_{11}^T = \varepsilon_{22}^T = \varepsilon^T = QP^2$, where Q is the electrostrictive coefficient. We consider both the FE and PE layers are cubic elastic media with moduli C_{11} , C_{12} and $\bar{C}_{11}$, $\bar{C}_{22}$, respectively. The effective elastic constants are given by $G = C_{11} + C_{12} - 2C_{12}^2/C_{11}$ and $\bar{G} = \bar{C}_{11} + \bar{C}_{12} - 2\bar{C}_{12}^2/\bar{C}_{11}$.

Stresses in the bilayer are thus caused by the combined effects of the misfit and transformation strains, which can be calculated following the treatment of Hsueh and Evans (Hsueh, 1985; 2003). Thus, start with uncoupled FE and PE layers that experience unconstrained and incompatible in-plane strains $\varepsilon_{11} = \varepsilon_{22} = \varepsilon^0 - QP^2$ and $\bar{\varepsilon}_{11} = \bar{\varepsilon}_{22}$, respectively (Fig. 6.10b). Uniform stresses are then imposed on each layer to achieve displacement compatibility and equilibrium, i.e., uniform strain and zero net force (Fig. 6.10c). Finally, relaxing the constraint of strain uniformity, bending of the system occurs because of the development of asymmetric stresses as shown in Fig. 6.10d. In the final configuration, the strain depends on z and can be written as

$$\varepsilon = c + (z - t_b)/r, \text{ for } -H \leq z \leq +h, \quad (6.3.2)$$

where $z = t_b$ is the location of the bending axis, which is different from the conventional neutral axis. r is the radius of curvature of the system. According to Eq. (6.3.2) the strain is proportional to the distance from the bending axis and inversely proportional to the radius of curvature.

Taking into account both the misfit and transformation strains, the stresses σ_s and σ_f in the substrate and film, respectively, are given by

$$\begin{aligned} \sigma_s &= \bar{G}\varepsilon \text{ for } -H \leq z \leq 0, \\ \sigma_f &= G(\varepsilon - QP^2 + \varepsilon^0) \text{ for } 0 \leq z \leq h. \end{aligned} \quad (6.3.3)$$

The strain distributions in the FE/PE bilayer system are contingent upon solutions of the three parameters, c , t_b and r , which can be determined sequentially from the following three equilibrium conditions.

Firstly, the zero resultant force in Fig. 6.10c gives

$$\int_{-H}^0 \bar{G} c dz + \int_0^h G \left(c + \varepsilon^0 - Q P^2 \right) dz = 0. \quad (6.3.4)$$

The uniform strain component c from Eq. (6.3.4) is then given by

$$c = - \frac{Gh \left(\varepsilon^0 - \frac{Q}{h} \int_0^h P^2 dz \right)}{\bar{G}H + Gh}. \quad (6.3.5)$$

Secondly, the force equilibrium in Fig. 6.10d requires

$$\int_{-H}^0 \bar{G} \frac{z - t_b}{r} dz + \int_0^h G \frac{z - t_b}{r} dz = 0, \quad (6.3.6)$$

from which the position of the bending axis t_b can be determined:

$$t_b = \frac{Gh^2 - \bar{G}H^2}{2(Gh + \bar{G}H)}. \quad (6.3.7)$$

Thirdly, the bending equilibrium in Fig. 6.10d requires

$$\int_{-H}^0 \sigma_s z dz + \int_0^h \sigma_f z dz = \int_{-H}^0 \varepsilon(z) z dz + \int_0^h \left[\varepsilon(z) + \varepsilon^0 - Q \langle P^2 \rangle \right] z dz = 0, \quad (6.3.8)$$

from which the radius of the curvature r can be determined from Eqs. (6.3.4)–(6.3.8) to give

$$r = \frac{\bar{G}h_s^2 (2H + 3t_b) + Gh^2 (2h - 3t_b)}{3 \left[\bar{G}cH^2 - Gh^2 \left(c + \varepsilon^0 - \frac{Q}{h} \int_0^h P^2 dz \right) \right]}. \quad (6.3.9)$$

Finally, the total elastic energy U_{Ela} of the bilayer corresponding to Eq. (6.3.3) is given by

$$\begin{aligned} U_{\text{Ela}} &= \frac{1}{2} \int_{-H}^h \left(\sigma_s(z) \varepsilon_1(z) + \sigma_f(z) \varepsilon_2(z) \right) dz \\ &= \int_{-H}^0 \bar{G} \varepsilon(z) \varepsilon(z) dz + \int_0^h G \left[-2 \left[\varepsilon(z) + \varepsilon^0 \right] Q P^2(z) + Q^2 P^4(z) \right. \\ &\quad \left. + \left(\varepsilon^0 \right)^2 + \varepsilon^2(z) + 2\varepsilon(z) \varepsilon^0 \right] dz, \end{aligned} \quad (6.3.10)$$

where $\varepsilon(z)$ can be determined by Eqs. (6.3.2), (6.3.5), (6.3.7) and (6.3.9). It must be pointed out that U_{Ela} here has already included the interaction energy as well.

One can compare the energy equation in Section 6.2.1, and it is clear that $U_{\text{Ela}} = U_s + U_e$.

The intrinsic free energy U_p in the first term of Eq. (6.3.1) is now considered. U_p can be written as the sum of the non-linear Landau free-energy functional for a field-free and uniform infinite crystal and the Ginzburg contribution to account for the presence of non-uniformity in P .

$$U_p = \int_0^h \left[\frac{1}{2} A(T - T_{c0}) P^2 + \frac{1}{4} B P^4 + \frac{1}{6} C P^6 + \frac{1}{2} D \left(\frac{dP}{dz} \right)^2 - \frac{1}{2} E_d P - E_{\text{ext}} P \right] dz, \quad (6.3.11)$$

where A , B and C are the expansion coefficients of the Landau free energy, D can be approximated as $\xi^2 \cdot |A(T - T_{c0})|$, and ξ is a characteristic length along which the polarization varies.

The last two terms in Eq. (6.3.11) are the contributions from the depolarization and external electric fields E_d and E_{ext} , respectively. As a result of possible charge compensation, such as that due to oxygen vacancies in perovskite ferroelectrics, we may assume for simplicity that E_d is negligible. The contribution U_p can then be written as

$$U_p = \int_0^h \left[\frac{1}{2} A(T - T_{c0}) P^2 + \frac{1}{4} B P^4 + \frac{1}{6} C P^6 + \frac{1}{2} D \left(\frac{dP}{dz} \right)^2 - E_{\text{ext}} P \right] dz. \quad (6.3.12)$$

We note that this section mainly focuses on the mechanical behavior of an FE/PE bilayer system. We simplify the boundary conditions on the upper and lower surfaces to $\frac{\partial P}{\partial z} = 0$ at $z = 0$ and $z = h$, with $\delta_0 = \delta_h \rightarrow \infty$. δ_0 and δ_h are the extrapolation lengths at the top and bottom surfaces, which measure the effect of surface relaxation on distribution of the polarization near surface.

6.3.2 Solution of the Evolution Equation

The evolution equation then can be written as

$$\begin{aligned} \frac{\partial P(z, h, H, t)}{\partial t} &= -M \frac{\delta U}{\delta P} \\ &= -M \left[A^*(z, h, H, t) P + B^*(z, h, H, t) P^3 + C P^5 - D \frac{d^2 P}{dz^2} - E_{\text{ext}} \right], \end{aligned} \quad (6.3.13)$$

and the renormalized coefficients are

$$A^*(z, h, H, t) = A(T - T_{c0}) - 4GQ \left[c(h, H, t) + \frac{z - t_b(h, H)}{r(h, H, t)} + \varepsilon^0 \right],$$

$$B^* = B + 4GQ^2,$$

where M is the kinetic coefficient related to the domain wall mobility. From Eqs. (6.3.3) to (6.3.10), we can give the evolution of the polarization, the uniform strain component, the position of the bending axis, and the radius of the curvature as functions of ferroelectric layer thickness h , paraelectric layer thickness H and evolution time t

$$c(h, H, t) = -\frac{Gh}{\bar{G}H + Gh} \left(\varepsilon^0 - Q \langle P^2(t) \rangle \right),$$

$$t_b(h, H) = \frac{Gh^2 - \bar{G}H^2}{2(Gh + \bar{G}H)}, \quad (6.3.14)$$

$$r(h, H, t) = \frac{\bar{G}h_s^2(2H + 3t_b) + Gh^2(2h - 3t_b)}{3[\bar{G}cH^2 - Gh^2(c + \varepsilon^0 - Q \langle P^2(t) \rangle)]}.$$

A finite difference method for spatial derivatives and the Runge-Kutta method for temporal derivatives are employed. Eq. (6.3.13) can be solved numerically to yield the equilibrium polarization. Then the parameters, c , t_b and r can be obtained from Eq. (6.3.14). At $t = 0$, the initial value of the polarization is considered as a small perturbation $P(z, t = 0) = \Delta \approx 0$. The initial values of the uniform strain component $c(h, H, t = 0)$, position of the bending axis $t_b(h, H)$ and radius of the curvature $r(h, H, t = 0)$ can be calculated by solving Eq. (6.3.14) with $\langle P^2(t = 0) \rangle \approx 0$. From Eq. (6.3.13), the time-dependent polarization $P(z, t)$ and average polarization $P_a(t) = \int_0^h P(t) dz / h$ can be obtained from the solution of Eq. (6.3.13). From Eqs. (6.3.13) and (6.3.14), the time-dependent uniform strain component $c(h, H, t)$, the position of the bending axis t_b , and the time-dependent radius of the curvature $r(h, H, t)$ can be obtained, respectively.

6.3.3 The Stationary and Relative Bending Displacements of the Bilayer

By solving Eqs. (6.3.13) and (6.3.14), the stationary polarization P_∞ , uniform

strain component c_∞ , and radius of the curvature r_∞ can be obtained at $t_M > t_s$, where the evolution time of TDGL is $t_M = tM = N\Delta tM$, Δt is the time step, N is the total number of the time steps, and t_s is the time for obtaining stability of P , c and r from their initial state.

According to Fig. 6.11, the stationary displacement caused by the bending due to E_{ext} is given by (Lu et al., 2009)

$$d_\infty(E_{\text{ext}}) = r_\infty - \sqrt{r_\infty^2 - (L/2)^2}. \quad (6.3.15)$$

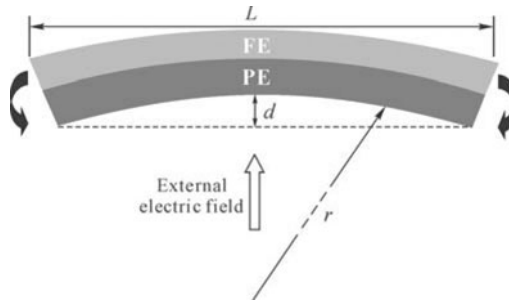


Fig. 6.11 Schematic diagram showing bending of the FE/PE bilayer due to applied external electric fields. Inhomogeneous stress distribution in the FE/PE bilayer results in a radius of curvature r and a dome-shaped bending with displacement d

We define the relative displacement in terms of the stationary ones by

$$\Delta D(E_{\text{ext}}) = \frac{d_\infty(E_{\text{ext}}) - d_\infty(E_{\text{ext}} = 0)}{\bar{H}(h, H)}, \quad (6.3.16)$$

where $\bar{H} = h + H$ is the thickness of the ferro/paraelectric bilayer. The relative displacement measures the piezoelectric bending response as an effect of the applied electric field and vanishes in its absence.

6.3.4 Dynamic Piezoelectric and Bending Response of the Bilayer Under a Cyclic Electric Field

Under a cyclic sinusoidal external field E_{ext} along the z direction as in (Zheng et al., 2006; 2007):

$$E_{\text{ext}} = E_0 \sin\left(\frac{2\pi t}{T'}\right) = E_0 \sin(2\pi t f'), \quad (6.3.17)$$

where E_0 , T' and f' are the amplitude, period and frequency, respectively. The time evolution of the system is governed by the time-dependent Ginzburg-Landau (TDGL) equation

$$\begin{aligned} \frac{\partial P(z, h, H, t)}{\partial t} &= -M \frac{\delta F}{\delta P} \\ &= -M \left[A^*(z, h, H, t) P + B^*(z, h, H, t) P^3 + CP^5 \right. \\ &\quad \left. - D \frac{d^2 P}{dz^2} - E_0 \sin\left(\frac{2\pi t}{T'}\right) \right]. \end{aligned} \quad (6.3.18)$$

The time-dependent mechanical responses of the system, the evolution of the polarization $P(z, h, H, t)$, the uniform strain component $c(h, H, t)$, the position of the bending axis $t_b(h, H)$, and the radius of the curvature $r(h, H, t)$ can be obtained numerically by solving Eqs. (6.3.14) and (6.3.18).

The dynamic and relative displacements of our system can then be calculated according to Eqs. (6.3.15) and (6.3.16):

$$d[E_{\text{ext}}(E_0, t)] = r(E_0, t) - \sqrt{r(E_0, t)^2 - \left(\frac{L(E_0, t)}{2}\right)^2}, \quad (6.3.19)$$

and

$$\Delta D[E_{\text{ext}}(E_0, t)] = \frac{d[E_{\text{ext}}(E_0, t)] - d(E_{\text{ext}} = 0)}{\bar{H}(h, H)}. \quad (6.3.20)$$

6.3.5 Examples and Discussions

The following results and discussions refer to the specific example of the BaTiO₃/SrTiO₃ bilayer system. The material constants for the Landau free energy, the electrostrictive coefficients and the elastic properties are from (Kwak et al., 1994; Li et al., 1997). The thickness h of the FE layer is 0.1 mm, and the lateral dimension L is $50h$. The time step used is $t = 0.1$ ns, and the number of time steps N in a period is 2000. In Fig. 6.12a, we show the spontaneous growth from zero to a final stationary value of the average polarization in the FE layer under static external electric fields $E_{\text{ext}} = 0$ and $E_{\text{ext}} = 200 \text{ kV} \cdot \text{cm}^{-1}$. Figs. 6.12b and 6.12c show the respective stationary values of the corresponding uniform strain component and radius of the curvature. The time t_M in these figures is a unit of 0.2 μs .

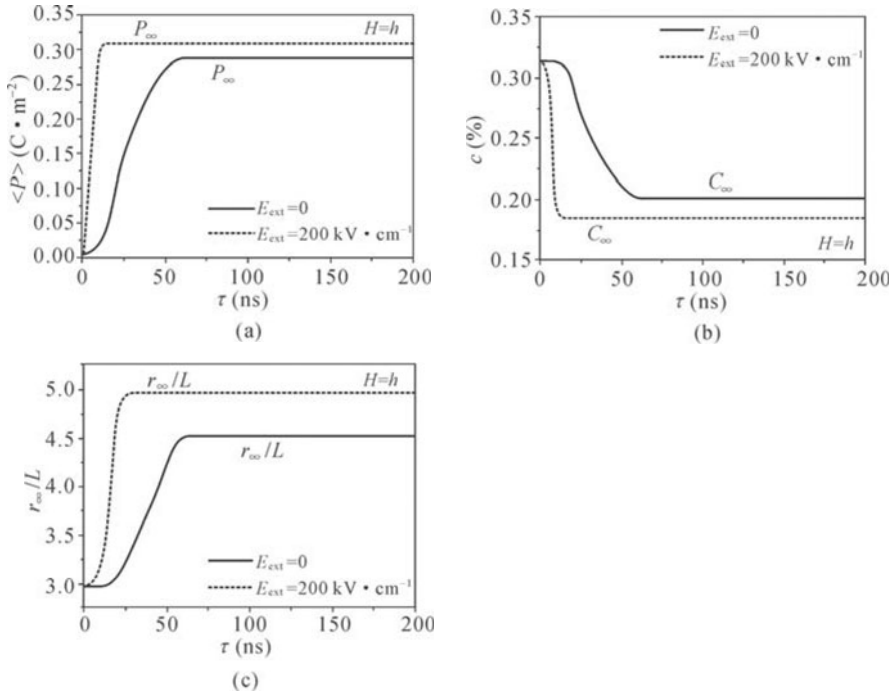


Fig. 6.12 (a) The polarization; (b) The uniform strain component; (c) The radius of the curvature from the solution of the TDGL equation (Eq. (6.3.18)). The calculation is done for $H = h = 0.1 \text{ mm}$. Evolution from the unpolarized unstable state to the polarized stationary state can be clearly seen

- *Stationary polarization, curvature and bending displacement in a static external electric field*

In Fig. 6.13a we plot the stationary average polarization $\langle P_{\infty} \rangle$ calculated using the foregoing procedure as a function of H , for various thicknesses of the FE film h . This result shows that the average polarization is a strong function of h for all values of H except the very small ones. $\langle P_{\infty} \rangle$ can be seen to vary between two limits of H : one in which H is large and the FE layer is totally constrained by a thick PE substrate, and the other in which H is small and the FE layer is free standing and unconstrained. The small peak of $\langle P_{\infty} \rangle$ at $H \approx 0.3h$ is the result of the bending of the bilayer. Figs. 6.13b and 6.13c plot the corresponding radius of curvature and static displacements, respectively, both of which show corresponding minima and maxima at $H \approx 0.3h$. It is interesting to note the very large displacements that can be obtained in the region $H \leq h$, where effective strains as large as 100% can be achieved.

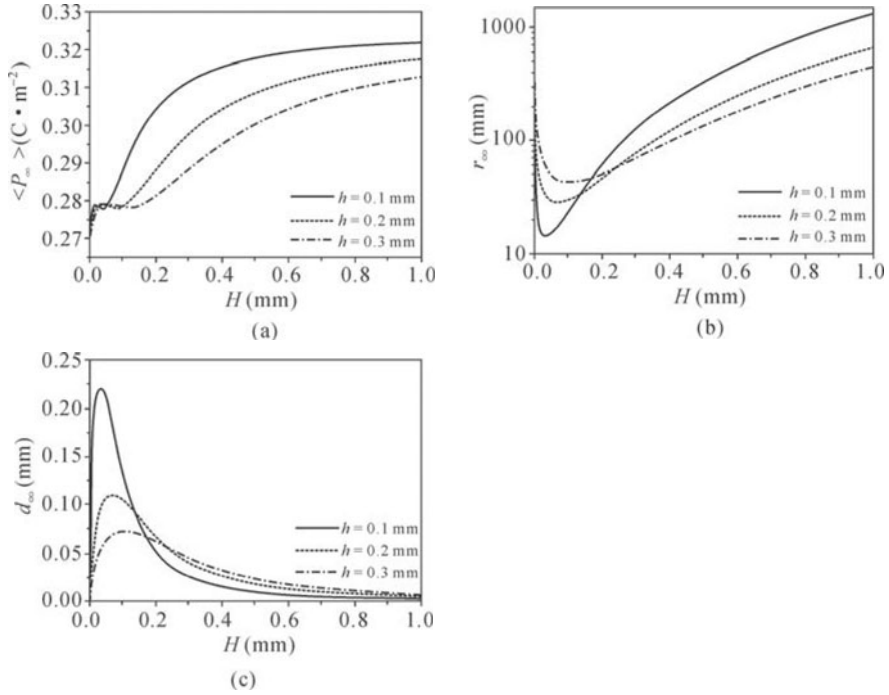


Fig. 6.13 (a) The average polarization; (b) Normalized static radius of curvature; (c) Normalized static displacement, as functions of the thickness of the PE substrate H for various FE film thickness h

• Non-zero external electric field

Fig. 6.14a shows the stationary average polarization $\langle P_{\infty} \rangle$ as a function of the external electric field, for various values of the PE substrate thickness. $\langle P_{\infty} \rangle$ can be seen to increase with an increasing external electric field and increasing substrate thickness. Based on results in Fig. 6.14a, we also calculate the stationary curvature r_{∞} and displacement d_{∞} . Fig. 6.14b shows that the stationary curvature increases and Fig. 6.14c shows that the displacement decreases, with an increasing external electric field and substrate thickness. From Eq. (6.3.16), we can also calculate the relative displacement as a function of the external electric field as shown in Fig. 6.14d. These results confirm the large piezoelectric bending response under an applied electric field. In the case of $H = 0.05h$, the relative displacement $|\Delta D|$ can be as high as 18% for an applied electric field $E_{\text{ext}} = 500 \text{ kV} \cdot \text{cm}^{-1}$. For a substrate thickness of $H = 0.5h$, $|\Delta D|$ may increase to 27%. Lower values of $|\Delta D|$, however, are obtained as the ratio H/h further increases.

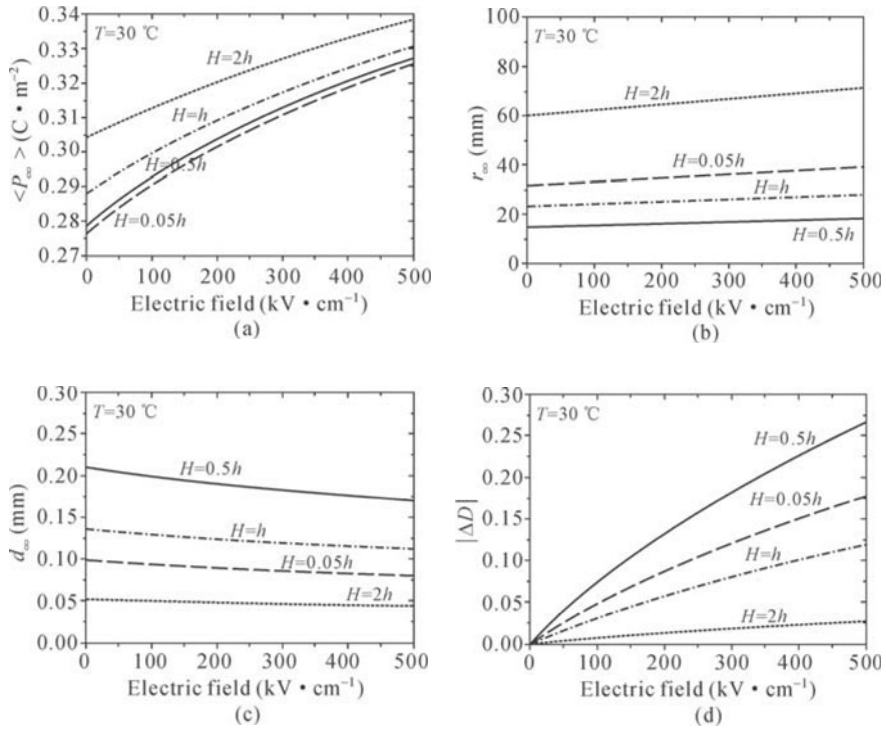


Fig. 6.14 Functions of the applied external electric field for various substrate thicknesses. (a) The average polarization; (b) Normalized radius of curvature; (c) Normalized static displacement; (d) The relative displacement

• Hysteresis loop

We compare the cases where $H = h = 0.1$ mm and $H = 2h = 0.2$ mm. The average polarization $\langle P \rangle$ is calculated as a function of a sinusoidal applied electric field by solving Eq. (6.3.18). In Eq. (6.3.17), the amplitude of the external field used is $E_0 = 1500$ kV · cm⁻¹. To alleviate the effect of the frequency of the external electric field on the switching behavior of the bilayer system, we use a long period $T' = 200$ ns. The time step is $\Delta t = T'/N$. The number of time steps N in a period is 200.

Fig. 6.15 shows the corresponding hysteresis loops of the BaTiO₃/SrTiO₃ bilayer. The remnant polarization P_r and coercive field E_c increase as H increases.

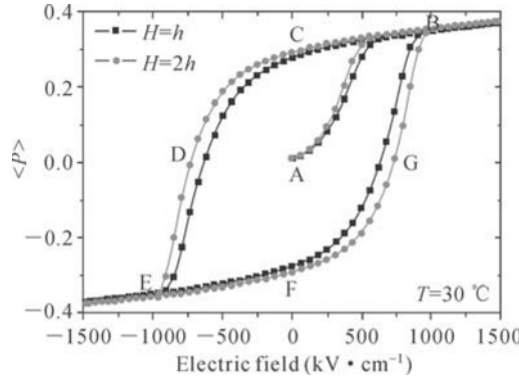


Fig. 6.15 Hysteresis loops in BTO/STO for $H=h$, and $H=2h$

• Dynamic loading

The dynamic curvature and displacement due to the sinusoidal external electric field of Eq. (6.3.17) are calculated and shown in Fig. 6.16 for a substrate thickness $H = h$ (square line). The radius of curvature r is shown in Fig. 6.16a. The dynamic displacement d can be obtained by solving Eq. (6.3.19) (Fig. 6.16b). It can be seen that the displacement can reach about 200% of FE film thickness at external electric fields $E_{\text{ext}} = -730$ (D) and 730 (G) ($\text{kV} \cdot \text{cm}^{-1}$). When the external electric field is large, such as $E_{\text{ext}} = -1100$ (E) or 1100 (B) ($\text{kV} \cdot \text{cm}^{-1}$), the displacement reduces to only about $0.7h$. From Figs. 6.16a and 6.16b, the relative dynamic displacement $|\Delta D|$ can also be obtained by solving Eq. (6.3.20). The results are plotted in Fig. 6.16c. It can be seen that $|\Delta D|$ can be as high as 58% for $E_{\text{ext}} = -1100$ (E) or 1100 (B) and $H = h$.

To examine the effect of the substrate thickness on the piezoelectric bending response, we also calculate the dynamic curvature, displacement d and relative displacement $|\Delta D|$ of a $\text{BaTiO}_3/\text{SrTiO}_3$ bilayer for $H=2h$ and plot the results in Fig. 6.17. It can be seen that the piezoelectric and bending responses are weakened due to the increase in the substrate thickness. Thus, the relative displacement $|\Delta D|$ in Fig. 6.17c is only 1/3 of that in Fig. 6.16c where $H = h$.

Results of Fig. 6.16 and Fig. 6.17 again demonstrate that the FE/PE bilayer system has a large piezoelectric bending response. The results also indicate that the piezoelectric and bending response can be adjusted via the thicknesses of the PE substrate and the FE film.

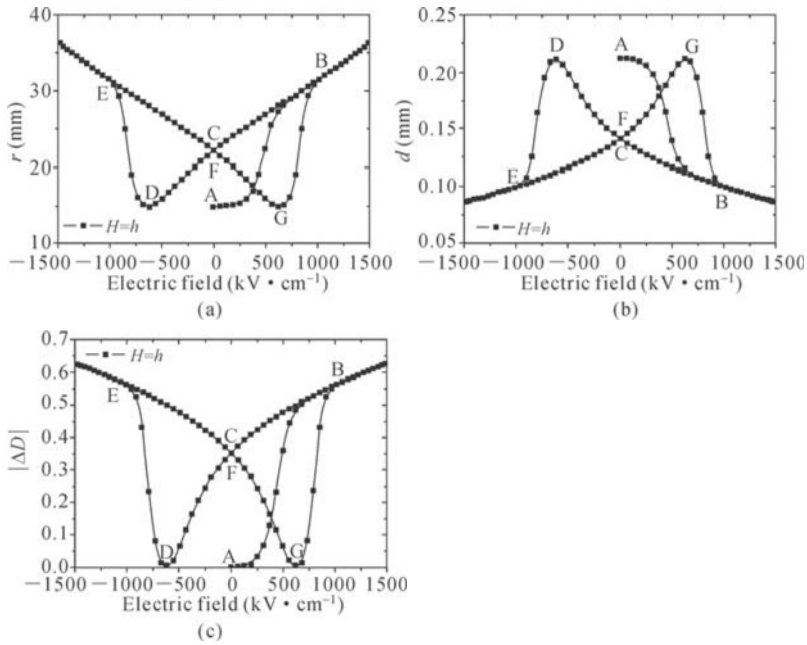


Fig. 6.16 (a) Dynamics of the radius of curvature, (b) displacement of the bonding bilayer, and (c) relative displacement of the bonding bilayer subjected to a sinusoidal applied electric field ($H=h$)

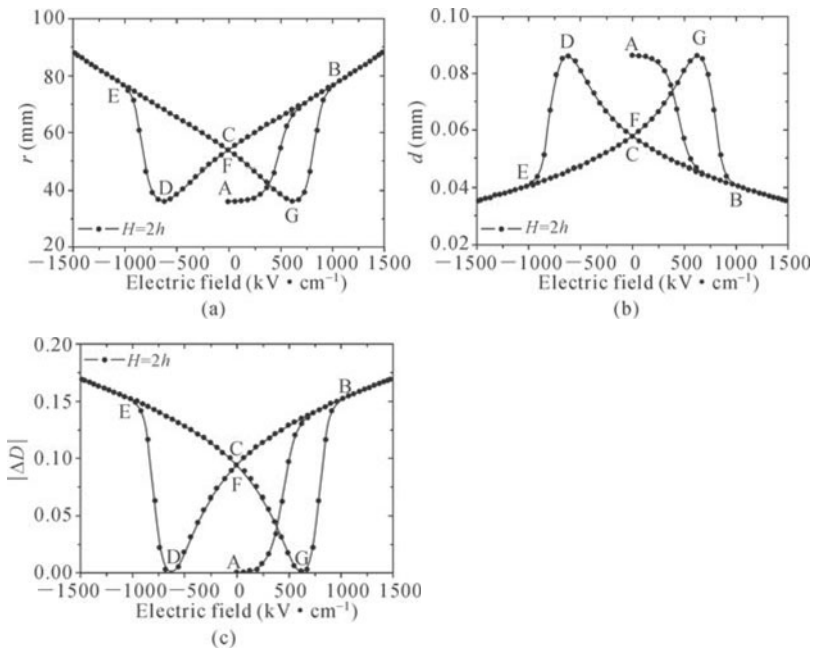


Fig. 6.17 same as Fig. 6.16 but for $H=2h$

6.4 Critical Thickness for Dislocation Generation in Piezoelectric Thin Films on Substrate

Thin films play a pivotal role in modern engineering and technology. Their performance is closely related to the material properties of the film and, in particular, the crystal orientation and the dislocation density and distribution. As a result of the crystalline misfit with the substrate, a newly deposited film is under stress, independent of whether it is polycrystalline or heteroepitaxial. As the film thickens, the stress builds up and, eventually, the conditions favor the generation of dislocations to relax the stress.

The existence of a critical thickness for dislocation generation in elastic thin films was first proposed by Frank and Merwe (1949) and it was subsequently confirmed experimentally and investigated theoretically (Zubia et al., 2002; People and Bean, 1985; Freund and Nix, 1996; Wang et al., 2003). They extended the analysis to piezoelectric thin films and discussed the effect of the piezoelectric properties on the critical thickness. Indeed, if the misfit strain between the substrate material and the film material is sufficiently small, or when the film is very thin, it can be accommodated elastically. However, as the film thickens, the energy required for elastic accommodation increases and eventually becomes excessive, and the dislocations are generated as an alternative means to accommodate the misfit.

In this section, mainly based on Wang et al. (2003), the formulation of coupled elastic and electric fields of a dislocation in a semi-infinite piezoelectric medium is derived. Then, the formation energy of the dislocation and the critical thickness for dislocation generation in a piezoelectric thin film grown on an infinite substrate is established. We also discussed the effects of the polarization and the misfit stresses on the critical thickness.

6.4.1 Elastic and Electric Fields in a Piezoelectric Semi-Infinite Space with a Dislocation

Consider the half-space $x_2 > 0$ occupied by a piezoelectric material, and a straight dislocation with Burger vector $\mathbf{b}$. The dislocation and the associated linecharge q per unit length trapped by the dislocation core (Im et al., 2001) are assumed to be at $(0, h)$, $h > 0$ (Fig. 6.18). In terms of the electrostatic potential φ , the electric field E_i is given by

$$E_i = -\frac{\partial \varphi}{\partial x_i} = -\varphi_{,i} . \quad (6.4.1)$$

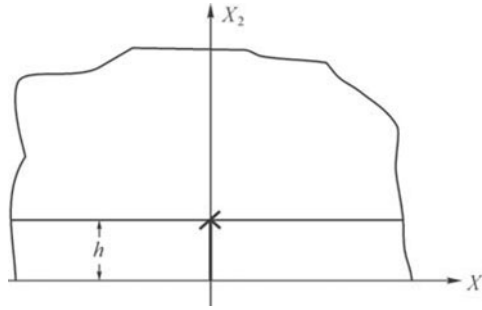


Fig. 6.18 Schematic diagram of a misfit dislocation in the interface between an epilayer and its substrate

The constitutive relation of the stress tensor σ_{ij} with the elastic displacement gradients $u_{i,j}$, and that of the electric displacement vector D_i with the electric field E_i respectively are given by

$$\begin{aligned}\sigma_{ij} &= C_{ijk m} u_{k,m} + e_{mij} \varphi_{,m}, \\ D_i &= e_{ikm} u_{k,m} - \alpha_{im} \varphi_{,m}.\end{aligned}\quad (6.4.2)$$

Here, repeated indices imply summation, and a comma stands for differentiation. $C_{ijk m}$ indicate the elastic stiffness tensor, e_{mij} are the piezoelectric stress constants and α_{im} are the permittivity constants. In the absence of body forces and free charges, the balance laws require that

$$\sigma_{ij,j} = 0, \quad D_{i,i} = 0. \quad (6.4.3)$$

Assuming free surface boundary conditions for the film, the traction and the electrical displacement both vanish at $x_2 = 0$. Following Barnett and Lothe (1975), we define the $3 \times 4 \times 4 \times 3$ matrix $\mathbf{G}$, the 3×4 matrix Σ , and the quadruple $\mathbf{u}$ by

$$G_{iJKm} = \begin{cases} C_{ijk m} & (J, K = 1, 2, 3), \\ e_{mij} & (J = 1, 2, 3; K = 4), \\ e_{ikm} & (J = 4; K = 1, 2, 3), \\ -\alpha_{im} & (J = K = 4), \end{cases} \quad (6.4.4)$$

$$\Sigma_{iJ} = \begin{cases} \sigma_{ij} & (J = 1, 2, 3), \\ D_j & (J = 4), \end{cases} \quad (6.4.5)$$

$$u_J = \begin{cases} u_j & (J = 1, 2, 3), \\ \varphi & (J = 4), \end{cases} \quad (6.4.6)$$

so that Eqs. (6.4.2) and (6.4.3) can be combined to give

$$G_{i/Km} u_{K,mi} = 0. \quad (6.4.7)$$

Here, $i, m=1, 2, 3, J, K=1, 2, 3, 4$. Indices from 1 to 3 refer to the elastic variables, and the index 4 refers to the electric variable. For two-dimensional problems, in which u_k depends on only x_1 and x_2 , the general solution of Eq. (6.4.7) is a function of a linear combination of x_1 and x_2 . Thus, we can write, without loss of generality,

$$u_J = a_J f(z), \quad z = x_1 + p x_2, \quad (6.4.8)$$

where f is an arbitrary function of z , and p and a_J are to be determined in the following.

Substituting Eq. (6.4.8) into Eq. (6.4.7), and noting that for f to be an arbitrary function, we must have

$$[\mathbf{Q} + p(\mathbf{R} + \mathbf{R}^T) + p^2 \mathbf{T}] \mathbf{a} = 0, \quad (6.4.9)$$

where the bold lower-case letters mean quadruples, and the bold capital letters mean 4×4 matrices. $\mathbf{Q}$, $\mathbf{R}$ and $\mathbf{T}$ are defined as

$$\begin{aligned} \mathbf{Q} &= \begin{bmatrix} C_{ilk1} & (e_{11})_i \\ (e_{11}^T)_k & -\alpha_{11} \end{bmatrix}, \\ \mathbf{R} &= \begin{bmatrix} C_{ilk2} & (e_{12})_i \\ (e_{12}^T)_k & -\alpha_{12} \end{bmatrix}, \\ \mathbf{T} &= \begin{bmatrix} C_{i2k2} & (e_{22})_i \\ (e_{22}^T)_k & -\alpha_{22} \end{bmatrix}, \\ (e_{ij})_s &= e_{ijs}. \end{aligned} \quad (6.4.10)$$

Introducing the generalized stress function ψ_J

$$\psi_J = b_J f(z), \quad \mathbf{b} = (\mathbf{R}^T + p\mathbf{T}) \mathbf{a} = -\frac{1}{p} (\mathbf{Q} + p\mathbf{R}) \mathbf{a}, \quad (6.4.11)$$

one obtains

$$\sigma_{i1} = -\psi_{i,2}, \quad \sigma_{i2} = -\psi_{i,1}, \quad D_1 = -\psi_{4,2}, \quad D_2 = -\psi_{4,1}. \quad (6.4.12)$$

It has been shown that p has eight eigenvalues, consisting of four pairs of complex conjugates (Barnett and Lothe, 1975; Ting, 1996). The four eigenvalues with $\text{Im}(p) > 0$ are denoted by p_α with $\alpha = 1, 2, 3, 4$, and $\bar{p}_\alpha$ with $\text{Im}(p) < 0$ are the complex conjugates of p_α . The corresponding eigenvectors in Eq. (6.4.9) are denoted by $\mathbf{a}_\alpha$.

By using the solution of the elastic counterpart (Ting, 1992), the general solution of Eq. (6.4.7) is given by

$$\begin{aligned}\mathbf{u} &= \frac{1}{\pi} \text{Im}[\mathbf{A}\Delta(z_* - p_*h)\mathbf{q}^\infty] + \frac{1}{\pi} \text{Im}\left(\sum_{\beta=1}^4 [\mathbf{A}\Delta(z_* - \bar{p}_\beta h)\mathbf{q}_\beta]\right), \\ \boldsymbol{\psi} &= \frac{1}{\pi} \text{Im}[\mathbf{B}\Delta(z_* - p_*h)\mathbf{q}^\infty] + \frac{1}{\pi} \text{Im}\left(\sum_{\beta=1}^4 [\mathbf{B}\Delta(z_* - \bar{p}_\beta h)\mathbf{q}_\beta]\right),\end{aligned}\quad (6.4.13)$$

where $\mathbf{A}$ and $\mathbf{B}$ are 4×4 complex matrices defined by using $\mathbf{b}_\alpha$ given by Eq. (6.4.13) for various eigenvalues p_α and eigenvectors $\mathbf{a}_\alpha$:

$$\mathbf{A}=[\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \mathbf{a}_4], \quad \mathbf{B}=[\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3, \mathbf{b}_4], \quad (6.4.14)$$

and

$$\begin{aligned}\Delta(z_* - p_*h) &= \text{diag}\{\log(z_1 - p_1h), \log(z_2 - p_2h), \log(z_3 - p_3h), \log(z_4 - p_4h)\}, \\ \Delta(z_* - \bar{p}_\beta h) &= \text{diag}\{\log(z_1 - \bar{p}_\beta h), \log(z_2 - \bar{p}_\beta h), \log(z_3 - \bar{p}_\beta h), \log(z_4 - \bar{p}_\beta h)\}.\end{aligned}\quad (6.4.15)$$

In Eq. (6.4.13), $\mathbf{q}^\infty$ and $\mathbf{q}_\beta$ are complex quantities to be determined.

The boundary conditions can be written in the following form:

$$\begin{aligned}\boldsymbol{\psi} &= 0, \quad x_2 = 0, \\ \int_{\mathcal{Q}} d\mathbf{u} &= \mathbf{b}, \text{ for any closed curve } \mathcal{Q} \text{ enclosing } (0, h), \\ \int_{\mathcal{Q}} d\boldsymbol{\psi} &= \mathbf{r}, \text{ for any closed curve } \mathcal{Q} \text{ enclosing } (0, h), \\ \sigma_{ij} &\rightarrow 0; \quad D_i \rightarrow 0, \text{ when } |Z| = \infty.\end{aligned}\quad (6.4.16)$$

In Eq. (6.4.16), $\mathbf{b} = (\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3, \mathbf{b}_4)$, $\mathbf{r} = (0, 0, 0, -q)$, b_i , $i = 1, 2, 3$, are the components of the Burgers vector. The quantity $b_4 \neq 0$ corresponds to an electric dipole layer along the cut plane. Note the distinction between $\mathbf{b}$ and the eigenvectors $\mathbf{b}_1, \mathbf{b}_2$,

etc., in Eq. (6.4.14). q is the line charge per unit length. Eq. (6.4.16) is a general form of the boundary conditions for the dislocation problems in epilayers. Substitution of Eq. (6.4.14) into Eq. (6.4.16) gives

$$\begin{aligned}\mathbf{q}^\infty &= \mathbf{A}^\top \mathbf{r} + \mathbf{B}^\top \mathbf{b}, \\ \mathbf{q}_\beta &= \mathbf{B}^{-1} \bar{\mathbf{B}} I_\beta \bar{\mathbf{q}}_\beta,\end{aligned}\quad (6.4.17)$$

where

$$\begin{aligned}\mathbf{I}_1 &= \text{diag}\{1, 0, 0, 0\}, \quad \mathbf{I}_2 = \text{diag}\{0, 1, 0, 0\}, \\ \mathbf{I}_3 &= \text{diag}\{0, 0, 1, 0\}, \quad \mathbf{I}_4 = \text{diag}\{0, 0, 0, 1\}.\end{aligned}\quad (6.4.18)$$

The solution for the half-space in Eq. (6.4.13) is made up of the two parts. The first terms on the right-hand side are the solution for the infinite space with a line charge of q per unit length along the dislocation line at $(x_1, x_2) = (0, h)$. The second terms represent 16 one-component images for the infinite space whose singularities are located outside the half-space. The first terms can be expressed in real form using the following relations (Ting, 1996):

$$\begin{aligned}\text{Im}[\mathbf{A}\Delta(z_* - p_*h)\mathbf{q}^\infty] &= \text{Im}[\mathbf{A}\Delta(z_* - p_*h)(\mathbf{A}^\top \mathbf{r} + \mathbf{B}^\top \mathbf{b})] \\ &= -\frac{1}{2} \log(\tilde{r}) \mathbf{h} - \frac{\pi}{2} \mathbf{S}(\hat{\theta}) \cdot \mathbf{h} - \frac{\pi}{2} \mathbf{H}(\hat{\theta}) \cdot \mathbf{g}, \\ \text{Im}[\mathbf{B}\Delta(z_* - p_*h)\mathbf{q}^\infty] &= \text{Im}[\mathbf{B}\Delta(z_* - p_*h)(\mathbf{A}^\top \mathbf{r} + \mathbf{B}^\top \mathbf{b})] \\ &= -\frac{1}{2} \log(\tilde{r}) \mathbf{g} - \frac{\pi}{2} \mathbf{L}(\hat{\theta}) \cdot \mathbf{h} - \frac{\pi}{2} \mathbf{S}^\top(\hat{\theta}) \cdot \mathbf{g},\end{aligned}\quad (6.4.19)$$

where $\mathbf{h} = \mathbf{S} \cdot \mathbf{b} + \mathbf{H} \cdot \mathbf{r}$, and $\mathbf{g} = \mathbf{S}^\top \cdot \mathbf{r} - \mathbf{L} \cdot \mathbf{b}$. The generalized Barnett-Lothe tensors $\mathbf{S}(\theta)$, $\mathbf{H}(\theta)$ and $\mathbf{L}(\theta)$ are defined as

$$\mathbf{S}(\theta) = \frac{1}{\pi} \int_0^\theta \mathbf{N}_1(\omega) d\omega, \quad \mathbf{H}(\theta) = \frac{1}{\pi} \int_0^\theta \mathbf{N}_2(\omega) d\omega, \quad \mathbf{L}(\theta) = \frac{1}{\pi} \int_0^\theta \mathbf{N}_3(\omega) d\omega, \quad (6.4.20)$$

with

$$\begin{aligned}\mathbf{N}_1(\theta) &= -\mathbf{T}^{-1}(\theta) \mathbf{R}^\top(\theta), \\ \mathbf{N}_2(\theta) &= \mathbf{T}^{-1}(\theta), \\ \mathbf{N}_3(\theta) &= \mathbf{R}(\theta) \mathbf{T}^{-1}(\theta) \mathbf{R}^\top(\theta) - \mathbf{Q}(\theta), \\ \underline{Q}_{JK}(\theta) &= n_i G_{iJKm} n_m, \quad R_{JK}(\theta) = n_i G_{iJKm} m_m, \quad T_{JK}(\theta) = m_i G_{iJKm} m_m.\end{aligned}\quad (6.4.21)$$

$\mathbf{n}^\top = [\cos \theta, \sin \theta, 0]$ and $\mathbf{m}^\top = [-\sin \theta, \cos \theta, 0]$ are the respective unit vectors, normal and tangential, to a circle with its centre at $r = 0$, and $\mathbf{S}$, $\mathbf{H}$ and $\mathbf{L}$ are the corresponding values of $\mathbf{S}(\theta)$, $\mathbf{H}(\theta)$ and $\mathbf{L}(\theta)$ at $\theta = \pi$. In Eq. (6.4.19), $(\hat{r}, \hat{\theta})$ is the polar coordinate system with origin at $(0, h)$. The second terms in Eq. (6.4.13) can be written as

$$\begin{aligned} \frac{1}{\pi} \operatorname{Im} \left(\sum_{\beta=1}^4 \{ \mathbf{A} \Delta(z_* - \bar{p}_\beta h) \mathbf{q}_\beta \} \right) &= \sum_{\alpha=1}^4 \sum_{\beta=1}^4 U^{\alpha\beta}, \\ \frac{1}{\pi} \operatorname{Im} \left(\sum_{\beta=1}^4 \{ \mathbf{B} \Delta(z_* - \bar{p}_\beta h) \mathbf{q}_\beta \} \right) &= \sum_{\alpha=1}^4 \sum_{\beta=1}^4 \psi^{\alpha\beta}, \end{aligned} \quad (6.4.22)$$

with

$$\begin{aligned} U^{\alpha\beta} &= \frac{1}{\pi} \operatorname{Im} [\log(z_\alpha - z^{\alpha\beta}) q_{\alpha\beta} \mathbf{a}_\alpha], \\ \psi^{\alpha\beta} &= \frac{1}{\pi} \operatorname{Im} [\log(z_\alpha - z^{\alpha\beta}) q_{\alpha\beta} \mathbf{b}_\alpha], \end{aligned} \quad (6.4.23)$$

$$z^{\alpha\beta} = \bar{p}_\beta h = x_1^{\alpha\beta} + p_\alpha x_2^{\alpha\beta}, \quad (6.4.24)$$

where $q_{\alpha\beta}$ is the α -th component of $\mathbf{q}_\beta$, that is $q_{\alpha\beta} = (\mathbf{q}_\beta)_\alpha$, and $\mathbf{a}_\alpha$ and $\mathbf{b}_\alpha$ are the eigenvectors of Eqs. (6.4.9) and (6.4.11) defined earlier.

6.4.2 Critical Thickness for Dislocation Generation

Dislocation formation in an epilayer is energetically favourable when the strain energy that can be relieved by the dislocation is larger than its self-energy. The thickness for the generation of a dislocation in an epilayer is thus at least equal to that marked by the balance point of the misfit energy and the dislocation self-energy, which is also the point of zero Gibbs free-energy change when the dislocation forms in the piezoelectric film.

Neglecting the line charge, the formation energy of a dislocation in an originally stress-free epilayer-substrate system can be written as

$$\begin{aligned} E_f &= \frac{1}{2} \iiint_V [(\sigma_{ij}^m + \sigma_{ij}^d)(\epsilon_{ij}^m + \epsilon_{ij}^d) - (D_j^m + D_j^d)(\phi_{,j}^m + \phi_{,j}^d)] dV \\ &\quad - \frac{1}{2} \iiint_V (\sigma_{ij}^m \epsilon_{ij}^m - D_j^m \phi_{,j}^m) dV. \end{aligned} \quad (6.4.25)$$

Here quantities with the superscripts m and d represent the fields produced by the mismatch and the dislocation, respectively. The integration domain encloses the entire solid of unit thickness except the dislocation core. For convenience, the plane obtained from a cut along the x_2 axis is taken to be the glide plane (Fig. 6.18). Applying the divergence theorem to Eq. (6.4.22) yields

$$\begin{aligned}
 E_{\text{f}} &= \frac{1}{2} \iint_S [(\sigma_{ij}^{\text{m}} + \sigma_{ij}^{\text{d}})(u_j^{\text{m}} + u_j^{\text{d}}) - (D_j^{\text{m}} + D_j^{\text{d}})(\varphi^{\text{m}} + \varphi^{\text{d}})] n_j \, \text{d}s \\
 &\quad - \frac{1}{2} \iint_S (\sigma_{ij}^{\text{m}} \varepsilon_{ij}^{\text{m}} - D_j^{\text{m}} \varphi^{\text{m}}) n_j \, \text{d}s \\
 &= \frac{1}{2} \iint_S [(\sigma_{ij}^{\text{d}} u_j^{\text{d}} - D_j^{\text{d}} \varphi^{\text{d}}) n_j] \, \text{d}s + \iint_S (C_{ijkl} \varepsilon_{ij}^{\text{m}} u_{kl}^{\text{d}} - \alpha_{ij} \varphi_i^{\text{m}} \varphi_j^{\text{d}}) n_j \, \text{d}s \\
 &= E_{\text{dis}} + E_{\text{int}},
 \end{aligned} \tag{6.4.26}$$

where E_{dis} and E_{int} denote the self-energy of the dislocation and the interaction energy between the dislocation and the mismatch strain, respectively. The integration boundary S includes the upper and lower sections with side planes ($x_1 \rightarrow \infty$; $x_2 = 0$; $x_2 \rightarrow \infty$), and two sides of the cut. In deriving Eq. (6.4.26), we have used the constitutive relation Eq. (6.4.2). The integration of the lower and upper planes cancels each other on $x_2 = 0$, $x_1 \rightarrow \pm\infty$ or $x_2 \rightarrow \infty$, $\sigma_{ij} = 0$ and $D_j = 0$. Therefore, the only non-zero contributions come from the two sides A^+ and A^- of the cut plane, and the contour of the dislocation core.

In the following, we also omit the contribution from the dislocation core. Therefore

$$\begin{aligned}
 E_{\text{dis}} &= \frac{1}{2} \int_0^{h-r_0} \sigma_{ij}^{\text{a}} n_j [u_i] \, \text{d}x_2 - \frac{1}{2} \int_0^{h-r_0} D_j^{\text{a}} n_j [\varphi] \, \text{d}x_2 \\
 &= \frac{1}{2} \int_0^{h-r_0} \sigma_{ij}^{\text{a}} n_j b_i \, \text{d}x_2 - \frac{1}{2} \int_0^{h-r_0} D_j^{\text{a}} n_j b_4 \, \text{d}x_2,
 \end{aligned} \tag{6.4.27}$$

where

$$[u_i] = u_i(A^-) - u_i(A^+), \quad [\varphi] = \varphi(A^-) - \varphi(A^+), \tag{6.4.28}$$

$$\begin{aligned}
 E_{\text{int}} &= - \int_0^h C_{ijkl} \varepsilon_{ij}^{\text{m}} n_j [u_i] \, \text{d}x_2 + \int_0^h \alpha_{ij} \varphi_i^{\text{m}} n_j [\varphi] \, \text{d}x_2 \\
 &= - \int_0^h C_{ijkl} \varepsilon_{ij}^{\text{m}} n_j b_i \, \text{d}x_2 + \int_0^h \alpha_{ij} \varphi_i^{\text{m}} n_j b_4 \, \text{d}x_2.
 \end{aligned} \tag{6.4.29}$$

Under the boundary condition $D_j = 0$ on the surface of the epilayer, one obtains from the second equation of the constitutive relation Eq. (6.4.2)

$$\alpha_{ij} \varphi_j^{\text{m}} = e_{ikl} \varepsilon_{kl}^{\text{m}}, \tag{6.4.30}$$

from which the interaction energy can be obtained:

$$E_{\text{int}} = -(C_{ijkl}\epsilon_{ij}^m b_i - e_{1kl}\epsilon_{kl}^m b_4)h. \quad (6.4.31)$$

Here we have used $\mathbf{n} = (1, 0, 0)$. The self-energy can be obtained from Eq. (6.4.27):

$$\begin{aligned} E_{\text{dis}} = & -\frac{1}{4\pi}(b_1 g_1 + b_2 g_2 + b_3 g_3 - b_4 g_4) \log\left(\frac{h}{r_0}\right) \\ & - \frac{1}{2\pi} \sum_{\alpha=1}^4 \sum_{\beta=1}^4 \text{Im} \left\{ \log \left[\left(1 - \frac{p_\alpha}{\bar{p}_\beta}\right) + \frac{p_\alpha}{\bar{p}_\beta} \frac{r_0}{h} \right] q_{\alpha\beta} \mathbf{b}_\alpha \right\} \cdot \mathbf{b} \\ & + \frac{1}{2\pi} \sum_{\alpha=1}^4 \sum_{\beta=1}^4 \text{Im} \left\{ \log \left[\left(1 - \frac{p_\alpha}{\bar{p}_\beta}\right) + \frac{p_\alpha}{\bar{p}_\beta} \frac{r_0}{h} \right] q_{\alpha\beta} \mathbf{b}_{\alpha 4} \right\} \cdot \mathbf{b}_4, \end{aligned} \quad (6.4.32)$$

where $b_{\alpha 4}$ is the fourth component of the eigenvector $\mathbf{b}_\alpha$, g_i is the i th component of $\mathbf{g}$, which is defined after Eq. (6.4.19), and $\mathbf{b}$ is the Burgers vector as defined before. If there are no electric dipoles along the cut, $b_4 = 0$. Eq. (6.4.32) can be simplified accordingly. Finally, substitution of Eqs. (6.4.31) and (6.4.32) into Eq. (6.4.26) yields the following equation for the critical thickness h_c of the piezoelectric epilayer for dislocation generation:

$$\begin{aligned} & \frac{1}{4\pi}(b_1 g_1 + b_2 g_2 + b_3 g_3 - b_4 g_4) \log\left(\frac{h}{r_0}\right) \\ & + \frac{1}{2\pi} \sum_{\alpha=1}^4 \sum_{\beta=1}^4 \text{Im} \left\{ \log \left[\left(1 - \frac{p_\alpha}{\bar{p}_\beta}\right) + \frac{p_\alpha}{\bar{p}_\beta} \frac{h}{r_0} \right] q_{\alpha\beta} \mathbf{b}_\alpha \right\} \cdot \mathbf{b} \\ & - \frac{1}{2\pi} \sum_{\alpha=1}^4 \sum_{\beta=1}^4 \text{Im} \left\{ \log \left[\left(1 - \frac{p_\alpha}{\bar{p}_\beta}\right) + \frac{p_\alpha}{\bar{p}_\beta} \frac{h}{r_0} \right] q_{\alpha\beta} \mathbf{b}_{\alpha 4} \right\} \cdot \mathbf{b}_4 = (C_{ijkl}\epsilon_{ij}^m b_i - e_{1kl}\epsilon_{kl}^m b_4)h_c. \end{aligned} \quad (6.4.33)$$

6.4.3 Effect of Piezoelectric Behavior of the Materials on the Critical Thickness for Dislocation Formation

In the following, we use the theory developed in the foregoing to calculate numerically the critical thickness for dislocation formation in this system and explore the dependence of the critical thickness on the piezoelectric properties of the epilayer.

Consider $\text{Al}_x\text{Ga}_{1-x}\text{N}$ with a wurtzite crystal structure to grow in the (0001)

orientation on GaN, with the same crystal structure. We use the lattice constants of the $\text{Al}_x\text{Ga}_{1-x}\text{N}$ film obtained using the linear interpolation (Ambacher et al., 2000):

$$\begin{aligned} a_0(x) &= (-0.077x + 3.189) \times 10^{-10} \text{ m}, \\ c_0(x) &= (-0.203x + 5.189) \times 10^{-10} \text{ m}. \end{aligned} \quad (6.4.34)$$

When $x=0$, one can obtain the lattice constants for GaN. The elastic constants are as follows (Shimada et al., 1998):

$$\begin{aligned} C_{11} = C_{33} &= 350 \text{ GPa}, C_{22} = 376 \text{ GPa}, C_{13} = 140 \text{ GPa}, \\ C_{12} &= 104 \text{ GPa}, C_{44} = C_{66} = 101 \text{ GPa}. \end{aligned} \quad (6.4.35)$$

The piezoelectric and dielectric constants are (Ambacher et al., 2000)

$$\begin{aligned} e_{22} &= 1 \text{ C} \cdot \text{m}^{-2}, e_{21} = e_{23} = -0.36 \text{ C} \cdot \text{m}^{-2}, e_{16} = e_{34} = -0.3 \text{ C} \cdot \text{m}^{-2}, \\ \alpha_{11} = \alpha_{33} &= 9.5\epsilon_0, \alpha_{22} = 10.4\epsilon_0. \end{aligned} \quad (6.4.36)$$

Here, we note that in our coordinate system the axis of symmetry of transversely isotropic (or hexagonal) materials is the x_2 axis, instead of the x_3 axis, as is usually the case. For single-crystal $\text{Al}_x\text{Ga}_{1-x}\text{N}$, with a wurtzite structure, grown on the basal plane of sapphire or GaN, a high density of threading dislocations parallel to the c axis crossed the film from the interface to the film surface (Ning et al., 1996; Shen et al., 2000). They have a predominantly edge character with a $\frac{1}{3}\langle 11\bar{2}0 \rangle$ Burgers vector. In our coordinate system, the components of the Burgers vector can be expressed in the form

$$\mathbf{b} = [a_0, 0, 0]. \quad (6.4.37)$$

To obtain the solution, one needs to derive the eigenvalues and eigenvectors through Eqs. (6.4.9) and (6.4.11). The wide range of material constants (from about 10^{-10} to about 10^{10}) presents considerable difficulties in the numerical calculation. To solve this problem, we rewrite the constitutive Eq. (6.4.2) in the form

$$\begin{aligned} \sigma_{ij} &= \frac{1}{\alpha_0} (\alpha_0 C_{ijkl} + e_{mij} \hat{\phi}_{,m}), \\ D_i &= e_{ikm} u_{k,m} - \alpha'_{im} \hat{\phi}_{,m}, \end{aligned} \quad (6.4.38)$$

where $\hat{\phi}_{,m} = \alpha_0 \phi_{,m}$ and α_0, α'_{im} are the vacuum dielectric constant and the relative dielectric constant, respectively. By substituting Eq. (6.4.38) into Eq. (6.4.7), one can dispose of the parameter $1/\alpha_0$ before the parentheses. Then, using the material matrixes Eq. (6.4.10), one can establish the matrix $\mathbf{N}$ and derive the

eigenvalues and eigenvectors as follows:

$$\begin{bmatrix} \mathbf{N}_1 & \mathbf{N}_2 \\ \mathbf{N}_3 & \mathbf{N}_1^T \end{bmatrix} \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix} = p \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix}, \quad (6.4.39)$$

where

$$\mathbf{N}_1 = -\mathbf{T}^{-1} \mathbf{R}^T, \quad \mathbf{N}_2 = -\mathbf{T}^{-1}, \quad \mathbf{N}_3 = \mathbf{R} \mathbf{T}^{-1} \mathbf{R}^T - \mathbf{Q}, \quad (6.4.40)$$

and $\mathbf{T}$, $\mathbf{Q}$ and $\mathbf{R}$ are given by Eq. (6.4.10), replacing C_{ijkl} with $\alpha_0 C_{ijkl}$, and α_{im} with the relative dielectric constants.

Assuming $b_4 = 0$, the mismatch strain then becomes

$$\varepsilon_{11}^m = \varepsilon_{33}^m = \frac{a}{a_0} - 1 \approx 0.00531. \quad (6.4.41)$$

We are now ready to consider the effect of the piezoelectric properties of the epilayer on the critical thickness, by considering the $\text{Al}_x\text{Ga}_{1-x}\text{N}$ material system, whose material constants are given by Eqs. (6.4.35) and (6.4.36). There are three independent piezoelectric constants, and their effects on the critical thickness are shown respectively in Figs. 6.19 – 6.21. The piezoelectric constants e_{21} and e_{16} assume negative values. The relations between the critical thickness and the mismatch strain, with and without the piezoelectric effect respectively, are compared in Fig. 6.22. It is interesting to note that the piezoelectric properties affect the critical thickness significantly, by an increase of about 10% – 50%. The case where the piezoelectric constant equals zero corresponds to the non-piezoelectric anisotropic elastic materials.

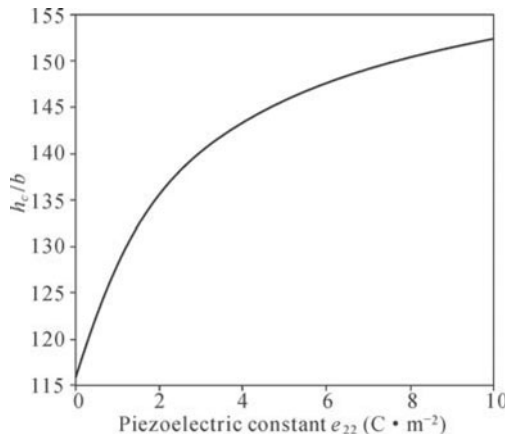


Fig. 6.19 The critical thickness h_c/b versus the piezoelectric constant e_{22} of the epilayer, where the index 2 corresponds to the direction of the c axis

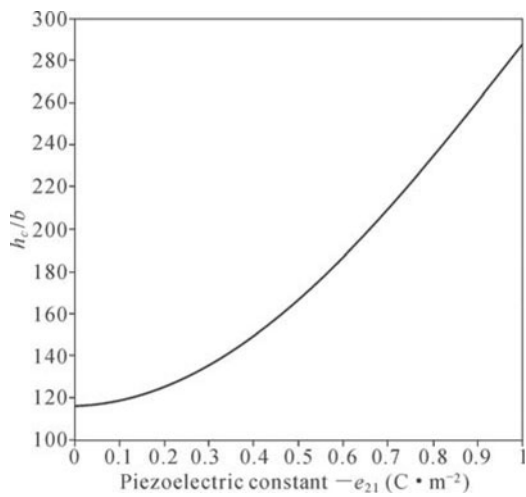


Fig. 6.20 The critical thickness versus the piezoelectric constant $-e_{21}$

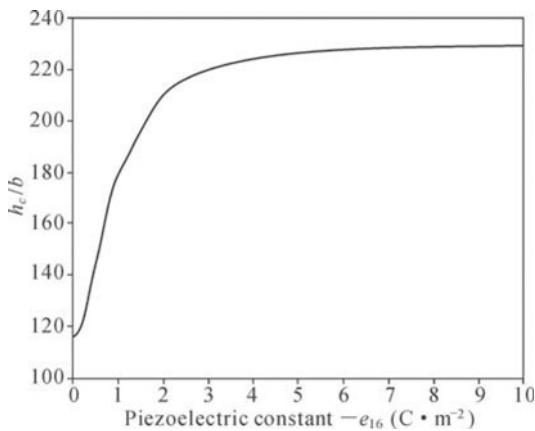


Fig. 6.21 The critical thickness versus the piezoelectric constant $-e_{16}$ of the epilayer

6.5 Critical Thickness of Dislocation Generation in Ferroelectric Thin Film on a Compliant Substrate

In this section, based on the Landau-Devonshire formalism and Timosheko’s method for thermal stresses, the critical thickness for misfit dislocation generation in ferroelectric thin film on a compliant substrate is established. And we also discussed the effects of the polarization and the misfit stresses on the critical thickness.

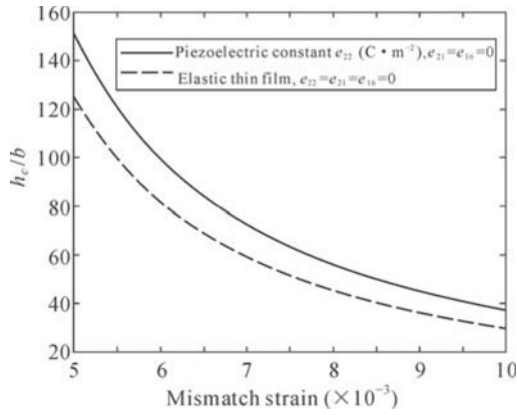


Fig. 6.22 The critical thickness versus the mismatch strain

6.5.1 Mechanical Properties of the Problem

A ferroelectric thin film deposited on a compliant substrate incorporated on a thick viscous borophosphosilicate glass layer (BPSG) is considered (Fig. 6.23). h and H are the thicknesses of the film and substrate, respectively. The substrate on BPSG can either shrink or expand to minimize the strain energy. For this setup, it has been shown that the observed final strains agree well with that predicted by stress balance, and that the bending of the film/substrate system can be neglected (Kastner and Gosele, 2000; Yin et al., 2003). The origin of our coordinate system is put at the lower-surface of the ferroelectric thin film. We also consider the x plane and y plane to be infinite, so that all the associated fields are functions of z only. We suppose the polarization P is only along the z -direction.

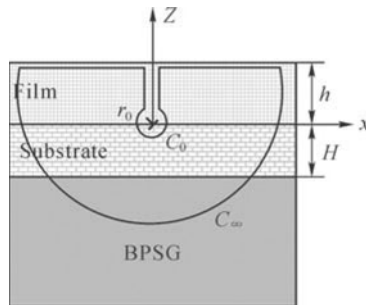


Fig. 6.23 Schematics diagram of a ferroelectric thin film on a compliant substrate

The total free energy of a ferroelectric thin film on compliant substrate is the sum of the bulk free energy, the gradient energy, the elastic energy, the electrostatic energy and the surface energy of the ferroelectric thin film. Without

the elastic interactions, the free energy can be expressed in terms of the Ginzburg-Landau functional as

$$U_p = \int_0^h \left\{ \frac{1}{2} A (T - T_{c0}) P^2 + \frac{1}{4} B P^4 + \frac{1}{6} C P^6 + \frac{1}{2} D \left(\frac{dP}{dz} \right)^2 - \frac{1}{2} E_d P \right\} dz + \frac{1}{2} \left(\frac{D P_0^2}{\delta} + \frac{D P_h^2}{\delta} \right), \quad (6.5.1)$$

where A , B , C and D are phenomenological coefficients. T_{c0} is the cooling phase-transition temperature of the bulk crystal. P_0 and P_h are the polarization at the surface and the interface of the thin film, respectively. δ is the extrapolation length. E_d is the depolarization field (Zheng et al., 2006; Choudhury et al., 2005).

For a ferroelectric thin film grown on a compliant substrate of finite-thickness, the total elastic energy can be derived following Timoshenko's method for calculating thermal stresses (Timoshenko and Goodier, 1970). For a coherent interface between the film and substrate, we define the biaxial in-plane misfit strains in the film as $\varepsilon_{11}^0 = \varepsilon_{22}^0 = \varepsilon^0 = (a_s - a_f)/a_f$, in which a_s and a_f are the lattice constants of the substrate and film, respectively. We note that the critical thickness refers to the point just before dislocation formation, at which the interface is still coherent. In addition to the misfit strain, there is also an in-plane electrostrictive strain given by $\varepsilon_{11}^T = \varepsilon_{22}^T = \varepsilon^T = QP^2$, where Q is the electrostrictive coefficient. Both the film and the substrate are treated as a cubic elastic body with elastic constants C_{11} , C_{12} and $\bar{C}_{11}$, $\bar{C}_{22}$, respectively. By applying the misfit stresses and the electrostrictive stresses on the ferroelectric thin film $\sigma_{\text{appl}} = G(\varepsilon^0 - QP^2)$, where $G = C_{11} + C_{12} - 2C_{12}^2/C_{11}$, and keeping the lattice constant of the film equal to that of the unstressed substrate, the resultant force (per unit length) due to σ_{appl} is given by

$$N_1 = N_2 = N = \int_0^h \sigma_{\text{appl}} dz = \int_0^h G(\varepsilon^0 - QP^2) dz. \quad (6.5.2)$$

Equilibrium of the film/substrate system requires an equal and opposite force $-N$ to balance the effect of σ_{appl} . The uniform strain ε^b produced by the force is given by

$$\sigma_{11}^b = \sigma_{22}^b = \sigma^b = G\varepsilon^b, \bar{\sigma}_{11}^b = \bar{\sigma}_{22}^b = \bar{\sigma}^b = \bar{G}\varepsilon^b, \\ h\sigma_{11}^b + H\bar{\sigma}_{11}^b = h\sigma_{22}^b + H\bar{\sigma}_{22}^b = -N, \quad (6.5.3)$$

where $\bar{G} = \bar{C}_{11} + \bar{C}_{12} - 2\bar{C}_{12}^2/\bar{C}_{11}$ and the barred quantities refer to the substrate.

Eqs. (6.5.2) and (6.5.3) can be solved to give ε^b as $\varepsilon^b = -N / (hG + H\bar{G})$.

The total elastic energy of the film/substrate system is then given by

$$U_{\text{elastic}} = \int_0^h G(\varepsilon^0 - QP^2 + \varepsilon^b)(\varepsilon^0 - QP^2 + \varepsilon^b) dz + \int_h^{h+H} \bar{G}\varepsilon^b \varepsilon^b dz. \quad (6.5.4)$$

The stationary polarization state corresponds to the minimum total free energy $U = U_p + U_{\text{elastic}}$, which can be obtained by solving the variational equation

$$\frac{\delta U}{\delta P} = 0. \text{ To focus on the critical thickness of the misfit dislocation generation, we}$$

simplify the boundary conditions on the upper and lower surfaces to $\frac{\partial P}{\partial z} = 0$ at z

$= 0$ and $z = h$, with extrapolation lengths $\delta_0 = \delta_h \rightarrow \infty$. The corresponding polarization state is uniform along z direction. If P_∞ is the stationary value of the polarization, the strains in the film and substrate are given by

$$\begin{aligned} \varepsilon_{11} = \varepsilon_{22} = e = \varepsilon^0 - QP_\infty^2 + \varepsilon^b, \\ \bar{\varepsilon}_{11} = \bar{\varepsilon}_{22} = \bar{e} = \varepsilon^b. \end{aligned} \quad (6.5.5)$$

6.5.2 The Formation Energy and the Critical Thickness of Spontaneous Formation of Misfit Dislocation

The effect of elastic constants on the critical thickness of an epitaxial layer grown on a compliant substrate was first formulated by Freund and Nix (1996). And as introduced in Section 6.4, in (Wang et al., 2003), the formulation of coupled elastic and electric fields of a dislocation in a semi-infinite piezoelectric medium is first derived. Then the formation energy of the dislocation and the critical thickness for dislocation generation in a piezoelectric thin film grown on an infinite substrate is established. Based on the Landau-Devonshire formalism and Timoshenko's method for thermal stresses, Zheng et al. (2006) extended their work and established the critical thickness for dislocation generation in a ferroelectric thin film on compliant substrate, which will be introduced as follows.

We now introduce a dislocation with Burgers vector $\mathbf{b} = [\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3]$ at the coherent interface. The formation energy E_f of the dislocation at the coherent interface is given by (Kastner and Gosele, 2000; Yin et al., 2003)

$$E_f = \frac{1}{2} \iint_{\Sigma} (\sigma_{\text{dis}} + \sigma_c)(\varepsilon_{\text{dis}} + \varepsilon_c) dx dz - \frac{1}{2} \iint_{\Sigma} \sigma_c \varepsilon_c dx dz = E_{\text{dis}} + E_{\text{int}}, \quad (6.5.6a)$$

where $E_{\text{dis}} = \frac{1}{2} \iint_{\Sigma} \sigma_{\text{dis}} \varepsilon_{\text{dis}} dx dz$ and $E_{\text{int}} = \frac{1}{2} \iint_{\Sigma} (\sigma_{\text{c}} \varepsilon_{\text{dis}} + \sigma_{\text{dis}} \varepsilon_{\text{c}}) dx dz$. σ_{dis} and ε_{dis} are the stress and strain fields of the dislocation respectively. σ_{c} and ε_{c} are the stress and strain in the film/substrate system from Eq. (6.5.5), i.e., without the dislocation. E_{dis} and E_{int} are the self-energy of the dislocation and the elastic interaction between the strain fields of the dislocation, the misfit and the spontaneous polarization. Σ is the integration domain that encloses the whole system, the film, the substrate and the BPSG. By using contour integration under the traction-free condition along the surface $z = h$, E_{dis} in Eq. (6.5.6a) can be rewritten as

$$E_{\text{dis}} = E_0 + E_{\text{c}} = -\frac{1}{2} \int_{r_0}^L \sigma_{ij}^{\text{dis}} n_j b_i ds + \frac{1}{2} \int_{C_0} \sigma_{ij}^{\text{dis}} n_j u_{\text{dis},i} ds, \quad (6.5.6b)$$

where r_0 is the atomic-scale core cutoff radius, and L is the length of the cut. At the same time, $E_{\text{int}} = -\int_0^L \sigma_{ij}^{\text{c}} n_j b_i ds$. For simplicity, we replace in the present calculation $M = 2\mu(1+\nu)/(1-\nu)$ and $\bar{M} = 2\bar{\mu}(1+\bar{\nu})/(1-\bar{\nu})$, where μ and $\bar{\mu}$ are shear modulus and ν and $\bar{\nu}$ are Poisson's ratios of the film and the substrate, respectively. E_0 can be evaluated as

$$\begin{aligned} E_0 = & \frac{\mu}{4\pi(1-\nu)} \left[b_1^2 + b_2^2 + (1-\nu)b_3^2 \right] \ln \left[\frac{4hH}{(h+2H)r_0} \right] \\ & - \frac{\mu(b_1^2 + b_2^2)}{4\pi(1-\nu)} \left[1 - \frac{2H}{h+2H} + \frac{2H^2}{(h+2H)^2} \right] \\ & + E_{0,4s} + E_{0,4c}, \end{aligned} \quad (6.5.7)$$

where $E_{0,4c}$ and $E_{0,4s}$ are caused by the stress fields of the edge and screw components of the dislocation, respectively. The interaction energy can be calculated by integrating from the film surface to the dislocation core. For the ferroelectric thin film, E_{int} can be written as $E_{\text{int}} = -Mb_1h(\varepsilon^0 - QP_{\infty}^2 + \varepsilon^b)$.

The spontaneous formation of the dislocation is energetically viable when the formation energy satisfies $E_{\text{f}} < 0$. Solutions of Eq. (6.5.6) for $E_{\text{f}} = 0$ then give the minimum film thickness for misfit dislocation formation. For an infinite substrate, the condition for zero formation energy can be rewritten as

$$\begin{aligned} & \frac{\mu}{4\pi(1-\nu)} \left[b_1^2 + b_2^2 + (1-\nu)b_3^2 \right] \ln \left[\frac{2h}{r_0} \right] - \frac{\mu(b_1^2 + b_2^2)}{8\pi(1-\nu)} + E_{\text{c}} \\ & = Mb_1h(\varepsilon^0 - QP_{\infty}^2 + \varepsilon^b). \end{aligned} \quad (6.5.8)$$

We now apply the foregoing to consider the case of a BaTiO_3 thin film. Two kinds of substrates are considered: compressive substrates (e.g., SrTiO_3 , LaAlO_3) and tensile substrates (e.g., KTaO_3 , MgO), with a BPSG substrate as shown in Fig. 6.19. The material constants are taken from (Pertsev et al., 1988). We assume here that the dislocation lies on the $\{1,1,1\}$ plane, with Burgers vectors $\mathbf{b} = \pm b[\frac{1}{2}, \frac{1}{\sqrt{2}}, \frac{1}{2}]$ for the tensile and compressive substrate, respectively. We also assume an atomic-scale core cutoff radius of $r_0 = b/4$. We note that the Burgers vector may change with the polarization (Wang et al., 2003).

For a $\text{BaTiO}_3/\text{SrTiO}_3$ system, the stationary values P_∞ and the total strain e in the film are calculated and shown in Figs. 6.24a and 6.24b, respectively, as a function of the film and substrate thickness h and H . In Fig. 6.24b, we compare the total strain e in the film in the paraelectric phase ($T > T_c$) and ferroelectric phase ($T = 300$ K). Using Eq. (6.5.7), the dislocation formation energy E_f for films in the paraelectric and ferroelectric phases ($T = 300$ K) can be expressed as a function of the film and substrate thicknesses, as shown in Fig. 6.24c. The corresponding critical thickness as a function of the substrate thickness is shown in Fig. 6.20 where it is seen to be larger in the ferroelectric phase ($T = 0, 300, 600$ K) than in the paraelectric phase (i.e., $T > T_c$), showing the effect of the electrostriction on the dislocation formation energy E_f .

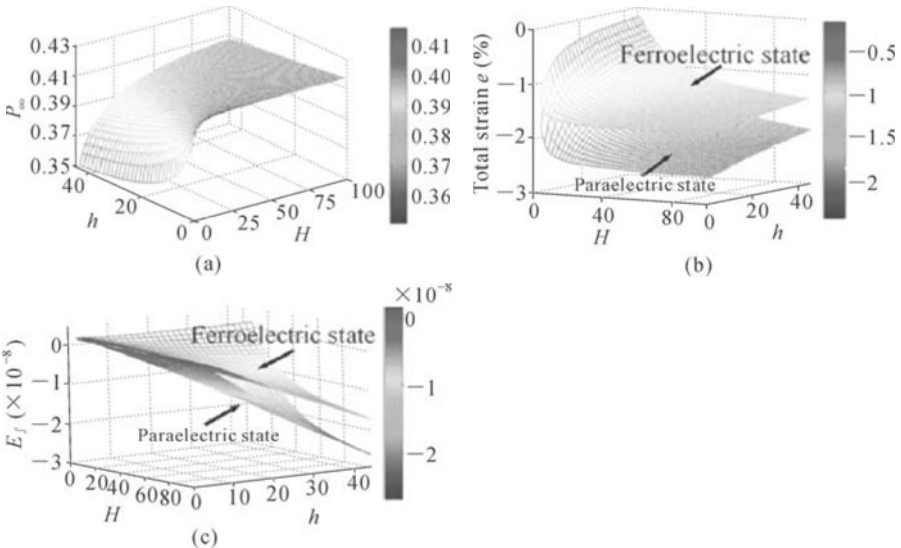


Fig. 6.24 (a) The polarization distribution ($T = 300$ K); (b) The distribution of the total strain; (c) The dislocation formation energy as a function of the film and substrate thicknesses in the paraelectric ($T > T_c$) and ferroelectric phases ($T = 300$ K)

In Fig. 6.25, we consider the critical thickness as a function of misfit strain for various values of P_∞ ($P_\infty = 0, 0.2, 0.4, 0.6 \text{ C}\cdot\text{m}^{-2}$) in the important case of thick

substrates. Both cases with compressive misfit strains ($\epsilon^0 < 0$), such as in the BaTiO₃/SrTiO₃ system) and tensile ones ($\epsilon^0 > 0$) (such as in the BaTiO₃/KTaO₃ system) were considered. In this regard, we note that the polarization depends on variables such as the ambient temperature and the external field, while the misfit strain varies with the substrate. Fig. 6.26a shows the critical thicknesses in the compressive misfit case. In general, the critical thickness increases as the magnitude of the misfit strain decreases, as expected. This is exactly the case for $P_\infty = 0$. For non-zero values of P_∞ the transition from the cubic to tetragonal phase causes a reduction in the lattice constant of the film, and polarization leads to a reduction in the effective misfit between the film and the substrate, causing an increase in the critical thickness. As the polarization increases, the cancellation between the misfit strain by electrostriction occurs, producing an infinite critical thickness when the cancellation is complete. This can be seen in the cases of $P_\infty = 0.4$ and $P_\infty = 0.6$, in which the electrostriction exactly cancels the misfit strain at about -0.7% and -1.6% respectively. If the misfit strain is smaller than that value, the misfit becomes effectively tensile under the polarization instead of compressive, and the critical thickness will decrease, as the magnitude of the misfit strain decreases. In Fig. 6.26a, interception of the curves with $P_\infty = 0$ and $P_\infty \neq 0$ can also be observed. It shows that even when there is polarization, with the same misfit strain, the critical thickness can be equal. Fig. 6.26b shows results for positive misfit strains (tensile). It can be seen that when the polarization increases, the critical thickness decreases as expected, showing much less complexity compared to the case where the epitaxial stress is in the same direction as the electrostriction.

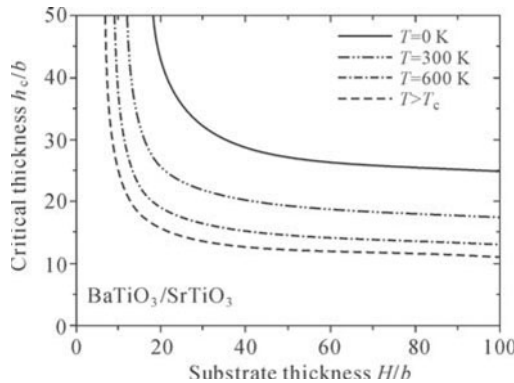


Fig. 6.25 Critical thickness for misfit dislocation generation h_c at different temperatures plotted as a function of substrate thickness

To summarize this section, from an energy perspective within the Landau-Devonshire formalism, we calculated the critical thickness of a spontaneous formation of misfit dislocations in a ferroelectric thin film on a compliant

substrate. It is found that the critical thickness depends on the substrate thickness, the polarization state and the misfit strain, particularly its sign.

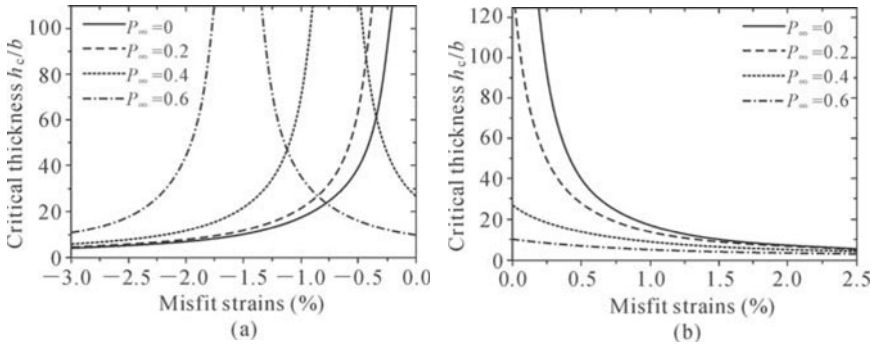


Fig. 6.26 The critical thickness h_c of misfit dislocation generation for film grown on thick, (a) compressive substrates and (b) tensile substrates with different polarization

6.6 Remarks

In this chapter, we mainly discussed how to manipulate the transition and critical characteristics of ferroelectric thin films by using the mechanical strains. We can see that if the ferroelectric thin film is on a compliant substrate, ferroelectric thin films generally exhibit characteristics that are different from the free standing ones. Through mechanics and Landau phase transition theory, we can find the total free energy of the system, which consists of elastic energy, interaction energy and intrinsic free energy. Then we can establish the evolution equation or the time-dependent Ginzburg-Landau equation. By treating the ferroelectric transition as a point of instability in the time-dependent Ginzburg-Landau equation, we can finally reveal how to tune the transition and critical characteristics of ferroelectric thin films.

In Section 6.2, we demonstrate that the characteristics of the phase transition, such as the transition temperature, the order of the phase transition, can be tuned simply by changing the relative thickness of the film and substrate. For a PbTiO_3 film/ SrTiO_3 substrate system, the transition temperature of the film can be adjusted in the range from 0 K to 1000 K, and the ferroelectric phase transition can also be tuned from the first-order as the bulk counterpart to the second-order.

In Section 6.3, by taking into account the effects of the misfit and transformation strains, we had shown that both the static and dynamic piezoelectric and bending responses due to the inhomogeneous stress in the bilayer can produce large vertical displacements. At the same time, this displacement, together with the associated polarization, radius of curvature and vertical displacement, can be adjusted by varying the thicknesses of the two layers.

In Section 6.4, the formulation of coupled elastic and electric fields of a dislocation

in a semi-infinite piezoelectric medium was derived. Then the formation energy of the dislocation and the critical thickness for dislocation generation in a piezoelectric thin film grown on an infinite substrate was established.

From an energy perspective within the Landau-Devonshire formalism, the critical thickness of the spontaneous formation of misfit dislocations in a ferroelectric thin film on a compliant substrate was calculated in Section 6.5. It was found that the critical thickness depends on the substrate thickness, the polarization state and the misfit strain, particularly its sign. Of particular importance, the generation of misfit dislocations can be totally suppressed or dragged on significantly through a careful design of the misfit strain and transformation strain in structures.

References

- [1] Alexander, L.R., Julia, S.: Thermodynamics of polydomain ferroelectric bilayers and graded multilayers. *Appl. Phys. Lett.* **89**, 042907 (2006)
- [2] Alpay, S.P., Misirlioglu, I.B., Nagarajan, V., Ramesh, R.: Can interface dislocations degrade ferroelectric properties? *Appl. Phys. Lett.* **85**, 2044-2046 (2004)
- [3] Ambacher, O., Foutz, B., Smart, J., Shealy, J.R., Weimann, N.G., Chu, K., Murphy, M., Sierakowski, A.J., Schaff, W.J., Eastman, L.F., Dimitrov, R., Mitchell, A., Stutzmann, M.: Two dimensional electron gases induced by spontaneous and piezoelectric polarization in undoped and doped AlGaIn/GaN heterostructures. *J. Appl. Phys.* **87**, 334-344 (2000)
- [4] Barnett, D.M., Lothe, J.: Dislocations and line charges in anisotropic piezoelectric insulators. *Phys. Status Solid B* **67**, 105-111 (1975)
- [5] Bhattacharya, K., James, R.D.: A theory of thin films of martensitic materials with applications to microactuators^{fn2}. *J. Mech. Phys. Solids* **47**, 531-576 (1999)
- [6] Choudhury, S., Li, Y.L., Krill Iii, C.E., Chen, L.Q.: Phase-field simulation of polarization switching and domain evolution in ferroelectric polycrystals. *Acta Mater.* **53**, 5313-5321 (2005)
- [7] Chow, S.N., Hale, J.K.: *Methods of bifurcation theory*. Springer, New York (1982)
- [8] Fong, D.D., Stephenson, G.B., Streiffner, S.K., Eastman, J.A., Auciello, O., Fuoss, P.H., Thompson, C.: Ferroelectricity in ultrathin perovskite films. *Science* **304**, 1650-1653 (2004)
- [9] Frank, F.C., van der Merwe, J.H.: One-Dimensional Dislocations. I. Static Theory. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* **198**, 205-216 (1949)
- [10] Freund, L.B., Nix, W.D.: A critical thickness condition for a strained compliant substrate/epitaxial film system. *Appl. Phys. Lett.* **69**, 173-175 (1996)

- [11] Gosling, T.J., Willis, J.R.: The energy of arrays of dislocations in an anisotropic half-space. *Phil. Mag. A* **69**, 65-90 (1994)
- [12] Hsueh, C.H.: Stress distribution and curvature in graded semiconductor layers. *J. Cryst. Growth* **258**, 302-309 (2003)
- [13] Hsueh, C.H., Evans, A.G.: Residual stresses in meta/ceramic bonded strips. *J. Am. Ceram. Soc.* **68**, 241-248 (1985)
- [14] Hsueh, C.H., Lee, S., Chuang, T.J.: An Alternative Method of Solving Multilayer Bending Problems. *J. Appl. Mech.* **70**, 151-154 (2003)
- [15] Hu, S.Y., Li, Y.L., Chen, L.Q.: Effect of interfacial dislocations on ferroelectric phase stability and domain morphology in a thin film—a phase-field model. *J. Appl. Phys.* **94**, 2542-2547 (2003)
- [16] Im, H.J., Ding, Y., Pelz, J.P., Heying, B., Speck, J.S.: Characterization of individual threading dislocations in GaN using ballistic electron emission microscopy. *Phys. Rev. Lett.* **87**, 106802 (2001)
- [17] Kastner, G., Gosele, U.: Principles of strain relaxation in heteroepitaxial films growing on compliant substrates. *J. Appl. Phys.* **88**, 4048-4055 (2000)
- [18] Kim, D.J., Jo, J.Y., Kim, Y.S., Chang, Y.J., Lee, J.S., Yoon, J.G., Song, T.K., Noh, T.W.: Polarization relaxation induced by a depolarization field in ultrathin ferroelectric BaTiO₃ capacitors. *Phys. Rev. Lett.* **95**, 237602 (2005)
- [19] Kwak, B.S., Erbil, A., Budai, J.D., Chisholm, M.F., Boatner, L.A., Wilkens, B.J.: Domain formation and strain relaxation in epitaxial ferroelectric heterostructures. *Phys. Rev. B* **49**, 14865-14879 (1994)
- [20] Lee, J.H., Fang, L., Vlahos, E., Ke, X., Jung, Y.W., Kourkoutis, L.F., Kim, J.W., Ryan, P.J., Heeg, T., Roeckerath, M., Goian, V., Bernhagen, M., Uecker, R., Hammel, P.C., Rabe, K.M., Kamba, S., Schubert, J., Freeland, J.W., Muller, D.A., Fennie, C.J., Schiffer, P., Gopalan, V., Johnston-Halperin, E., Schlom, D.G.: A strong ferroelectric ferromagnet created by means of spin-lattice coupling. *Nature* **466**, 954-958 (2010)
- [21] Li, S., Eastman, J.A., Vetrone, J.M., Foster, C.M., Newnham, R.E., Cross, L.E.: Dimension and size effects in ferroelectrics. *Jpn. J. Appl. Phys.* **36**, 5169-5174 (1997)
- [22] Lines, M.E., Glass, A.M.: *Principles and Applications of Ferroelectrics and Related Materials*. Oxford University Press, USA (1977)
- [23] Matthews, J.W., Blakeslee, A.E.: Defects in epitaxial multilayers. *J. Cryst. Growth* **27**, 118-125 (1974)
- [24] Matthews, J.W., Blakeslee, A.E.: Defects in epitaxial multilayers: I. Misfit dislocations. *J. Cryst. Growth* **27**, 118-125 (1974)
- [25] Mehta, R.R., Silverman, B.D., Jacobs, J.T.: Depolarization fields in thin ferroelectric films. *J. Appl. Phys.* **44**, 3379-3385 (1973)
- [26] Nicolis, G., Prigogine, I.: *Self-organization in Nonequilibrium Systems*. John Wiley & Sons, New York (1977)
- [27] Ning, X.J., Chien, F.R., Pirouz, P., Yang, J.W., Asif Khan, M.: Growth defects in GaN films on sapphire: The probable origin of threading dislocations. *J. Mater. Res.* **11**, 580-592 (1996)
- [28] Nix, W.: Mechanical properties of thin films. *Metal. Mater. Trans. A* **20**,

- 2217-2245 (1989)
- [29] People, R., Bean, J.C.: Calculation of critical layer thickness versus lattice mismatch for $\text{Ge}_x\text{Si}_{1-x}/\text{Si}$ strained-layer heterostructures. *Appl. Phys. Lett.* **47**, 322-324 (1985)
 - [30] Pertsev, N.A., Zembilgotov, A.G., Tagantsev, A.K.: Effect of mechanical boundary conditions on phase diagrams of epitaxial ferroelectric thin films. *Phys. Rev. Lett.* **80**, 1988-1991 (1998)
 - [31] Schlom, D.G., Chen, L.Q., Eom, C.B., Rabe, K.M., Streiffer, S.K., Triscone, J.M.: Strain tuning of ferroelectric thin films. *Mater. Res.* **37**, 589-626 (2007)
 - [32] Shen, B., Someya, T., Arakawa, Y.: Influence of strain relaxation of the $\text{Al}_x\text{Ga}_{1-x}\text{N}$ barrier on transport properties of the two-dimensional electron gas in modulation-doped $\text{Al}_x\text{Ga}_{1-x}\text{N}/\text{GaN}$ heterostructures. *Appl. Phys. Lett.* **76**, 2746-2748 (2000)
 - [33] Shimada, K., Sota, T., Suzuki, K.: First-principles study on electronic and elastic properties of BN, AlN, and GaN. *J. Appl. Phys.* **84**, 4951-4958 (1998)
 - [34] Timoshenko, S.P., Goodier, J.N.: *Theory of Elasticity*, 3th ed. McGraw-Hill, New York (1970)
 - [35] Ting, T.C.T.: Image singularities of Green's functions for anisotropic elastic half-spaces and bimetals. *Q. J. Mech. Appl. Math.* **45**, 119-139 (1992)
 - [36] Ting, T.C.T.: *Anisotropic Elasticity*. Oxford University Press, London (1996)
 - [37] Tybell, T., Ahn, C.H., Triscone, J.M.: Ferroelectricity in thin perovskite films. *Appl. Phys. Lett.* **75**, 856-858 (1999)
 - [38] van Vleck, J.H.: *The Theory of Electric and Magnetic Susceptibilities*. Clarendon, Oxford (1932)
 - [39] Wang, B., Woo, C.H., Zheng, Y.: Tunable Ferroelectric Phase Transition. IUTAM Symposium on Size Effects on Material and Structural Behavior at Micron- and Nano-Scales, 1-12 (2006)
 - [40] Wang, B., Woo, C.H.: The order of transition of a ferroelectric thin film on a compliant substrate. *Acta Mater.* **52**, 5639-5644 (2004)
 - [41] Wang, B., Woo, C.H.: Curie temperature and critical thickness of ferroelectric thin films. *J. Appl. Phys.* **97**, 084109 (2005)
 - [42] Wang, B., Woo, C.H., Sun, Q., Yu, T.X.: Critical thickness for dislocation generation in epitaxial piezoelectric thin films. *Phil. Mag.* **83**, 3753-3764 (2003)
 - [43] Wang, J., Zhang, T.Y.: Phase field simulations of polarization switching-induced toughening in ferroelectric ceramics. *Acta Mater.* **55**, 2465-2478 (2007)
 - [44] Willis, J.R., Jain, S.C., Bullough, R.: Work hardening and strain relaxation in strained-layer buffers. *Appl. Phys. Lett.* **59**, 920-922 (1991)
 - [45] Yanase, N., Abe, K., Fukushima, N., Kawakubo, T.: Thickness dependence of ferroelectricity in heteroepitaxial BaTiO_3 thin film capacitors. *Jpn. J. Appl. Phys.* **38**, 5305-5308 (1999)
 - [46] Yin, H., Hobart, K.D., Kub, F.J., Shieh, S.R., Duffy, T.S., Sturm, J.C.: Strain partition of Si/SiGe and SiO_2/SiGe on compliant substrates. *Appl. Phys. Lett.* **82**, 3853-3855 (2003)

- [47] Zembilgotov, A.G., Pertsev, N.A., Kohlstedt, H., Waser, R.: Ultrathin epitaxial ferroelectric films grown on compressive substrates: competition between the surface and strain effects. *J. Appl. Phys.* **91**, 2247-2254 (2002)
- [48] Zheng, Y., Wang, B., Woo, C.H.: Critical thickness for dislocation generation during ferroelectric transition in thin film on a compliant substrate. *Appl. Phys. Lett.* **89**, 083115 (2006)
- [49] Zheng, Y., Wang, B., Woo, C.H.: Effects of strain gradient on charge offsets and pyroelectric properties of ferroelectric thin films. *Appl. Phys. Lett.* **89**, 062904 (2006)
- [50] Zheng, Y., Wang, B., Woo, C.H.: Simulation of interface dislocations effect on polarization distribution of ferroelectric thin films. *Appl. Phys. Lett.* **88**, 092903 (2006)
- [51] Zheng, Y., Wang, B., Woo, C.H.: Piezoelectric bending response and switching behavior of ferroelectric/paraelectric bilayers. *Acta Mater.* **56**, 479-488 (2008)
- [52] Zheng, Y., Woo, C.H., Wang, B., Zhu, Z.Y.: Ferroelectric rods with adjustable dielectric tunability. *Appl. Phys. Lett.* **90**, 092905 (2007)
- [53] Zubia, D., Hersee, S.D., Khraishi, T.: Strain partitioning in coherent compliant heterostructures. *Appl. Phys. Lett.* **80**, 740-742 (2002)

Derivation of the Landau-Ginzburg Expansion Coefficients

7.1 Introduction

Ferroelectric materials have broad applications in multifunctional electronic devices, such as transducers, actuators, sensors and capacitors. Recently, particular interest was raised in finite size ferroelectrics because of the increasingly important role they play in the development of ultra-high density non-volatile ferroelectric random access memories (FRAMs). Meanwhile, the properties of a ferroelectric material, such as spontaneous polarization and Curie temperature, are well known to depend on many factors, including sample dimensions, ambient temperature, external mechanical loads, surface effects and electrostatic boundary conditions (Rabe et al., 2007; Landau et al., 1960; Junquera and Ghosez, 2003; Tybell et al., 1999; Zhong et al., 1994a; 1994b; Zembilgotov et al., 2002; Wang et al., 1995; Li et al., 1997; Lupascu, 2004; Wang and Woo, 2005; Lines and Glass, 1977). Since ferroelectricity is a collective phenomenon, it was believed for a long time that the effects of finite size would play an important role in the properties of ferroelectrics, such as polarization, dielectric constants and Curie temperature. Many theoretical and experimental studies have been conducted in the last few decades. Recently, research interest in this area has found a renaissance due to the surge of technological interest in ultra-miniaturized electronics, and many previously unsolved problems have been settled with some new techniques. For example, it was found that the critical thickness of ferroelectrics can be tuned through external strain engineering (Choi et al., 2004; Pertsev et al., 2007; Fong et al., 2006).

Over recent decades, theoretical simulations such as first-principle calculations, atomistic simulations and thermodynamic models have contributed significantly to the understanding of nanoscale ferroelectric systems (Junquera and Ghosez, 2003; Pertsev et al., 1998; Dawber et al., 2005a; 2005b; 2007; Tenne et al., 2006). Using proper order parameters, Landau theory can provide a reliable and reasonable

description of a system's equilibrium behavior near the phase transition. Thermodynamic models of ferroelectrics based on the Landau free-energy expansion are popular and well suited to study the properties of ferroelectrics. This phenomenological approach is also a conceptual bridge between microscopic models and macroscopic phenomena of materials (Kretschmer and Binder, 1979; Mehta et al., 1973; Nicolis and Prigogine, 1977; Streiffer et al., 2002; Zheng and Woo, 2008; Wang and Woo, 2006).

The attractions contain simplicity and easy experimental comparability of the models. Indeed, it is the general validity of thermodynamics from which the models are derived. For example, Landau's symmetry-based treatment of phase transitions was applied in the case of ferroelectrics by Devonshire (1949) and ferroelectrics of finite size by Ginzburg (Rabe et al., 2007; Landau et al., 1960). Then, many related models based on Landau theory (i.e., Landau-Devonshire (LD), Ginzburg-Landau (GL), Landau-Ginzburg-Devonshire (LGD), the time dependent Ginzburg-Landau (TDGL), phase field model (PFM), etc.) have been widely used for investigating the properties of ferroelectrics on different dimension scales. With concise regular energy expansion in the Landau-Devonshire-Ginzburg theory, it can describe many experimental observations associated with the phase transition. Despite inaccuracies due to the neglect of stochastic fluctuations of a thermodynamic system, these approaches have been proven powerful and effective, and is widely used to investigate the properties of ferroelectrics.

Landau free-energy can be expanded near the phase-transition instability in terms of order parameters in the form of the Taylor series, with coefficients that can be fitted to experimental data or gleaned from microscopic calculations. According to the work of Devonshire (1949), when truncating this series with only sixth-order terms, the Landau theory could also describe the first-order transition in ferroelectrics. The secret of the powerful Landau theory is that the order parameter itself is a function of expansion coefficients that are temperature dependent. Therefore, the input parameters are of paramount importance in the Landau-Ginzburg-Devonshire theory, and the quality of the coefficients will determine whether they can give a reliable description of the measurable quantities.

Recent modeling efforts of nanoscale ferroelectrics using the Landau approach have faced challenges in the formulation when contributions of the gradient energy, which include the effects of surface, inhomogeneous polarization distribution and depolarization field, have to be incorporated in the derivation of the free-energy functional. In this regard, it is of paramount importance that the thermodynamics, mechanics and electromagnetics involved are considered self-consistency (Zheng and Woo, 2008; 2009; Tagantsev, 2008).

When surface effects from polarization gradients and the near-surface effect are added, the accuracy of the Landau approach depends much on the accuracy of the input expansion coefficients, and the gradient terms are particularly significant in nanoscale ferroelectrics and domain structures. Therefore, the determination of the expansion coefficients of the energy expression is extremely important and

should receive more attention. As the Landau theory is a phenomenological theory, the parameters are usually obtained from the experimental data. In order to find the physical mechanism behind this theory, we must involve the microscopic theory, such as the lattice dynamic and microscopic statistical model, to explain the suitable condition for the expansion and the parameters derived from the fitting.

In this chapter, we will begin from the derived thermodynamic functions in Section 7.2. The phase transition temperature, polarization and susceptibility of ferroelectrics are calculated and discussed from the thermodynamic model. In Section 7.3 the experimental observations and the related approaches used to determine the coefficient of the Landau free-energy expansion are summarized. Moreover, in order to investigate the properties of nanoscale ferroelectric thin film, we introduce the gradient energy in the energy expression, and the determination of the gradient coefficient from a lattice dynamic model is reviewed in Section 7.4. Based on statistical mechanics, the transverse Ising model is introduced with its application in the phase transition in ferroelectrics. Then the gradient coefficients are obtained from the transverse Ising model. At last, some general characteristics of the gradient coefficients are derived based on a renormalization theory in Section 7.5.

7.2 Fundamental of the Landau-Devonshire Theory

7.2.1 The History of the Landau Free Energy Theory

In physics, phase transitions were first investigated by phenomenological theory, which used an internal mean field to describe the multi-body interaction of macroscopic systems and its related physical phenomena. The first equation of state to describe phase transition was given by Van der Waals in 1873. In 1937, Landau established his famous theory to describe the general second order phase transitions, in which thermodynamic order parameters of the system change continuously but its symmetry switches abruptly. To the best of our knowledge, the application of phenomenological theory to ferroelectric material began in the 1940s. The first work was done by Mueller (1940a; 1940b; 1940c), who applied it to Rochelle salt with considerable success. The basic idea of phenomenological theory is to expand the free energy of the system in a power series in the vicinity of the transition with the spontaneous polarization as its order parameter, to build the corresponding relationship between the expansion coefficient and the properties of the crystals. With a few parameters obtained from certain experiments, they could predict some properties of the ferroelectrics, including those with temperature dependence, such as phase transition. The Landau phase transition theory is a phenomenological theory about structural transition. It relates to the presence of the order parameter and the broken symmetry to solve the problem of continuous phase transition.

Due to the neglect of the fluctuation of the order parameter, Landau's theory is valid only when the fluctuation of the order parameter is small compared to the order parameter itself. Based on the criterion of Levanyuk (1959) and Ginzburg (1960), it was found that Landau's theory worked well when the coordinate number of the system was high. This could be achieved either by the long range interactions or in a large dimension space. Since the electric dipole-dipole interaction is a long range interaction, the Landau approach is well suited to describe the phenomenon in ferroelectrics, and it is invalid only in a very small region when the temperature is extremely close to the critical point. Since the interaction of the electric dipoles and the coupling between the polarizations and lattices are much stronger than their magnetic counterparts, Landau's phenomenological theories play an especially important role in ferroelectrics, more than in magnetism. It is noted that the Landau theory of structural phase transition, which highlights the relationship of broken symmetry and phase transition, is a mean (self-consistent) field theory, and in simplified cases it will reduce to the Van der Waals and Weiss mean field that has been known for a long time. In the case of long range interaction, the more polarized dipoles are included to contribute to the local field, and the more precise it will approach to the average mean field. Thus, the Landau phenomenological and other mean field theories are exact in the limit of the long range interactions.

After the discovery of Barium Titanate (BaTiO_3) as a ferroelectric material (Wul et al., 1945; 1946; Harwood et al., 1947), Ginzburg (1945; 1946) and Devonshire (1949; 1951; 1954) applied the Landau theory to describe its behavior. Later, the theory was applied to antiferroelectric material by Kittel (1951). After a series of investigations, the phenomenological theory of ferroelectrics was generally established with clarity. Since the work of Devonshire gave an integrated form of the theory, it is also called the Landau-Devonshire theory.

The merit of the Landau-Devonshire theory is that it is a straightforward theory linking the measurable thermodynamic properties to the order parameter that minimizes the free energy. However, the reliable description of the system depends on its input parameters, the coefficients of the expanded terms in the free energy, which are determined either by comparison with experiment or from microscopic approaches. Since the Landau theory is strictly a macroscopic theory, it cannot deal with problems related to microscopic physics.

7.2.2 The Thermodynamic Functions of the Dielectrics and Phase Transition

According to thermodynamic theory, the thermodynamic state of equilibrium of any system can be completely determined by the independent variables. In uniform polarized bulk ferroelectrics, for example, the appropriate independent variable can be chosen from the temperature (T), the entropy (S), the polarization

(P) or electric displacement (D), the electric field (E), the strain (η), and the stress (σ). According to the first law and second law of thermodynamics, the change in internal energy density in an elastic dielectric material can be written as

$$dU \leq TdS + \sum \sigma_{ij} d\eta_{ij} + \sum E_i dD_i. \quad (7.2.1)$$

The equality holds for a reversible process, and inequality holds for an irreversible process. In order to construct the thermodynamic functions, it is generally possible to choose the independent variables from one of the pairs of variables (T, S), (P or D, E), and (η, σ). There are several possible combinations, and some of the possible thermodynamic functions are given in Table 7.1. The first derivatives of the free energy give immediately the other dependent variables that can describe the properties of a macroscopic system in thermal equilibrium.

Table 7.1 The thermodynamic function of the dielectrics

Name	Independent variables	Thermodynamic functions	Thermodynamic functions in differential form
Helmholtz free energy	η, D, T	$F=U-TS$	$dF \leq -SdT + \sum \sigma_{ij} d\eta_{ij} + \sum E_i dD_i$
Enthalpy	σ, E, S	$H=U - \sum \sigma_{ij} \eta_{ij} - \sum E_i D_i$	$dH \leq TdS - \sum \eta_{ij} d\sigma_{ij} - \sum D_i dE_i$
Elastic enthalpy	σ, D, S	$H_1=U - \sum \sigma_{ij} \eta_{ij}$	$dH_1 \leq TdS - \sum \eta_{ij} d\sigma_{ij} + \sum E_i dD_i$
Electric enthalpy	η, E, S	$H_2=U - \sum E_i D_i$	$dH_2 \leq TdS + \sum \sigma_{ij} d\eta_{ij} - \sum D_i dE_i$
Gibbs free energy	σ, E, T	$G=U-TS - \sum \sigma_{ij} \eta_{ij} - \sum E_i D_i$	$dG \leq -SdT - \sum \eta_{ij} d\sigma_{ij} - \sum D_i dE_i$
Elastic Gibbs free energy	σ, D, T	$G_1=U-TS - \sum \sigma_{ij} \eta_{ij}$	$dG_1 \leq -SdT - \sum \eta_{ij} d\sigma_{ij} + \sum E_i dD_i$
Electric Gibbs free energy	η, E, T	$G_2=U-TS - \sum E_i D_i$	$dG_2 \leq -SdT + \sum \sigma_{ij} d\eta_{ij} - \sum D_i dE_i$

To be more specific, let us take σ , E and T as independent variables and consider the case of Gibbs free energy. Since, in most experiments, the control of temperature T , electric field E and stress σ is more convenient, in this case, by using the definition of Gibbs free energy, we have

$$dG \leq 0. \quad (7.2.2)$$

This means any increase in Gibbs free energy is impossible if the temperature T , stress σ and electric field E are fixed. Let us take a simple example, in which the stress σ and electric field E are zero in ferroelectrics. The paraelectric phase and ferroelectric phase are possible at a certain temperature. At a specific temperature, if there is a relationship $G_{\text{ferroelectric}} - G_{\text{paraelectric}} < 0$ between the two

phases, the change from paraelectric phase to ferroelectric phase is possible, and the change from ferroelectric phase to paraelectric phase is impossible. Therefore the ferroelectric phase is the stable phase.

7.2.3 The Expansion of the Free Energy and Phase Transition

For a uniform and infinite crystal (surfaceless and thus with no depolarization charge or depolarized field), we can use the energy of a free polarized, unstrained crystal as original reference and set it to zero. Due to the center symmetry, the g is an even function and all the components of tensor coefficients of odd rank are zero. Then the free energy can be written as

$$g(T, E) = \frac{A}{2}P^2 + \frac{B}{4}P^4 + \frac{C}{6}P^6 - PE_{\text{ext}} - \frac{1}{2}\epsilon_b E_{\text{ext}}^2. \quad (7.2.3)$$

The equilibrium configuration is determined by finding the minimum free energy density, that is

$$\frac{\partial g(T, E)}{\partial P} = 0. \quad (7.2.4)$$

If the electric field is parallel to the polarization, then we can derive

$$E_{\text{ext}} = AP + BP^3 + CP^5. \quad (7.2.5)$$

Particularly when $E_{\text{ext}}=0$, we obtain the solutions

$$P=0, \quad (7.2.6)$$

$$P^2 = \frac{-B \pm \sqrt{B^2 - 4AC}}{2C}. \quad (7.2.7)$$

7.2.3.1 The Continuous Ferroelectric Phase Transition

The zero solution corresponds to the paraelectric state. In the above expression, the coefficient C should be equal to or larger than zero. If $C < 0$, the polarization state would be always the stable state for a field-free, uniform and infinite ferroelectric crystal, because in that case, the larger the polarization, the more favorable will the free energy be.

For solution Eq. (7.2.7), if the coefficient $B > 0$, then $A < 0$ gives the necessary condition for the appearance of the spontaneous polarization, and the value is given by

$$P^2 = \frac{-B + \sqrt{B^2 - 4AC}}{2C}. \quad (7.2.8)$$

According to the definition of Ehrenfest (Smoluctowski et al., 1951), the n th order transition is a transition where the $(n-1)$ th derivative of Gibbs free energy G is continuous while the n th derivative is discontinuous at the Curie critical temperature. Usually, a phase transition with n th ($n \geq 2$) order is also called the continuous phase transition. A continuous (second-order) transition occurs at $T = T_c$ if $B > 0$ and $A < 0$, where the free energy and order parameter continuously decrease to zero when the temperature approaches the transition temperature, as seen in Fig. 7.1a. Once above the transition temperature, $A > 0$ and the polarization becomes unstable. Therefore, in the Landau-Devonshire theory, it is assumed that around the critical point the coefficient $A = (T - T_c) / (\epsilon_0 C_0)$, where C_0 is the Curie-Weiss constant, ϵ_0 the permittivity of vacuum, and T_c is the Curie temperature.

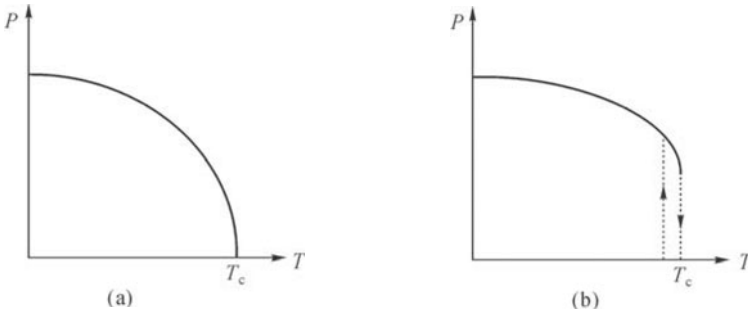


Fig. 7.1 Typical behavior of the spontaneous polarization P_s as a function of temperature, (a) the continuous (second order) phase and (b) the first order phase transition

The susceptibility χ^T of a ferroelectric material measures the response of the total observable polarization in an electric field, and is defined by

$$\chi^T = \left. \frac{\partial P^T}{\partial E} \right|_{E \rightarrow 0} = \chi_b + \left. \frac{\partial P}{\partial E} \right|_{E \rightarrow 0}. \quad (7.2.9)$$

The total polarization $\mathbf{P}^T$ of ferroelectric material is composed of two components: one is due to the permanent electric moment of the molecule itself and the other is induced by the electric fields. The former is the spontaneous polarization field P , which is used as the order parameter in our Gibbs free energy. Since the spontaneous polarization P is very small in the vicinity of the Curie

point, above the Curie transition point the quartic and sixth-order terms can be neglected, thus $\chi = 1/A$. When just below the Curie point, as $|A|$ is very small, it is reasonable to assume that $|A|C \ll B^2$ and Eq. (7.2.7) becomes

$$\begin{aligned} P^2 &= \frac{-B + \sqrt{B^2 - 4AC}}{2C} \\ &\approx \frac{-B + \sqrt{B^2 - 4AC + (2AC/B)^2}}{2C} = -\frac{A}{C}. \end{aligned} \quad (7.2.10)$$

It also gives the correct limit of the total susceptibility far away from its phase transition temperature T_c , i.e., $\chi^T \Big|_{T \rightarrow \infty} = \left[\frac{\varepsilon_0 C_0}{T - T_c} + \chi_b \right] \Big|_{T \rightarrow \infty} = \chi_b$, which should be experimentally determined. For example, the background susceptibility χ_b or dielectric constants ε_b of PbTiO_3 , BaTiO_3 , SrTiO_3 , CaTiO_3 are of the order of $50 \varepsilon_0$ in the limit of infinite temperature (Woo and Zheng, 2008; Pintilie et al., 2008; Rupprecht and Bell, 1964; Kittel, 2005). For $T > T_c$, the dielectric constant ε is given by

$$\left. \frac{\partial^2 g}{\partial E^2} \right|_{E \rightarrow 0} = \left. \frac{\partial D}{\partial E} \right|_{E \rightarrow 0} = \varepsilon_0 + \chi^T = \varepsilon_b + \chi = \varepsilon_b + \frac{\varepsilon_0 C_0}{|T - T_c|}, \quad (7.2.11)$$

which is consistent with the Curie-Weiss law observed in the vicinity of the Curie point, i.e.,

$$\varepsilon = \varepsilon_r + \frac{C}{T - T_c}, \quad (7.2.12)$$

where ε_r is the linear-part of the dielectric constant independent of temperature and C is the Currie-Weiss constant (Fig. 7.2).

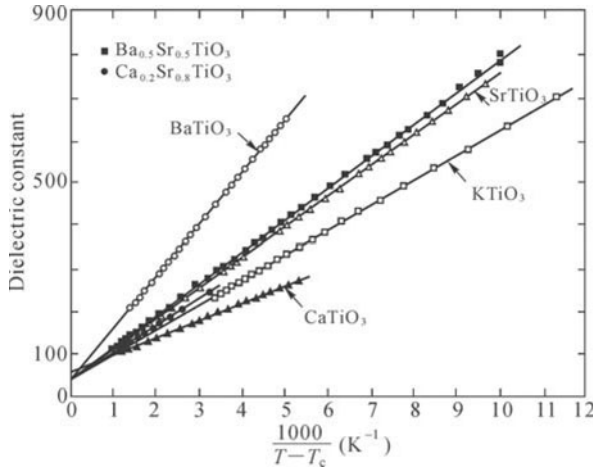


Fig. 7.2 The dielectric constants versus $1/(T-T_c)$ in the paraelectric state ($T > T_c$) of perovskites. Reprinted with permission from Rupprecht and Bell (1964). Copyright (1964), American Institute of Physics

7.2.3.2 The First-Order Ferroelectric Phase Transition

With the coefficient $B < 0$, Eq. (7.2.7) describes a first-order transition. In the first-order transition, the entropy S and the spontaneous polarization P will change discontinuously at the Curie point (see the polarization jump in Fig. 7.1b). Its Gibbs free energy evolutions under different temperatures are schematically illustrated in Fig. 7.3. As shown in Fig. 7.3, it is clear that there are three characteristic temperatures that can describe the different states of the ferroelectrics. When $T > T_1$, there is only one energy minimum in g , which corresponds to the paraelectric state without spontaneous polarization. When $T = T_1$, two inflexions appear in g , which can induce some ferroelectric polarizations under the external electric field. As the temperature further decreases, during $T_1 > T > T_c$, there are three minima in g , but the energy of the unpolarized state is lower than that of the polarized state, so the paraelectric state is the stable structure and the ferroelectric states are metastable. It is noted that at $T = T_c$, the Curie temperature, there are three minima in g which are energetically degenerate, and the polarization state of the system will depend on whether the temperature is approaching T_c from a lower or higher temperature. At any temperature between T_c and T_0 , the two subsidiary minima at symmetrical nonzero P will drop in energy below that of the unpolarized state; thus the ferroelectric states are thermodynamically favorable. At the same time, the paraelectric state may exist as a metastable state due to its local minimum in g . Once the temperature is below T_0 , there are two minima in g , corresponding to the energetically equivalent spontaneous polarizations of different spatial orientations.

If $B^2 > 4AC$ and $B > 0$, the equilibrium configurations that have the minimum

energy should satisfy the conditions $\frac{\partial g(T, E)}{\partial P} = 0$ and $\frac{\partial^2 g(T, E)}{\partial P^2} > 0$, which will give rise to the solution

$$P^2 = \frac{|B| + \sqrt{B^2 - 4AC}}{2C} > 0. \quad (7.2.13)$$

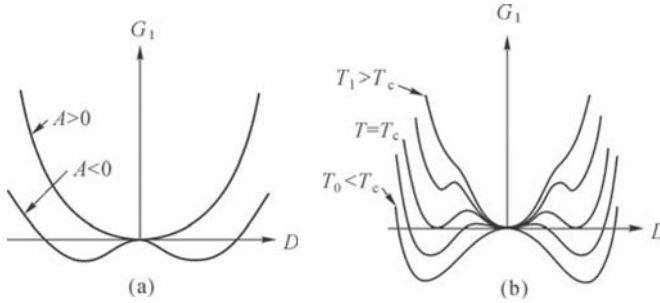


Fig. 7.3 Dependence of the Landau free energy on the order parameter for (a) second order phase transition, and (b) first order phase transitions (Lines and Glass, 1977)

When the system is at the Curie temperature T_c , the free energy of the paraelectric and ferroelectric state is equivalent. Since we use the energy of the free polarized, unstrained crystal as reference, its energy is set to be zero

$$\frac{A_c}{2} P_{sc}^2 + \frac{B}{4} P_{sc}^4 + \frac{C}{6} P_{sc}^6 = 0. \quad (7.2.14)$$

Here P_{sc} is the spontaneous polarization at the lower temperature close to the Curie point. When substituting Eq. (7.2.13) into Eq. (7.2.14), we obtain

$$A_c = \frac{3B^2}{16C}. \quad (7.2.15)$$

With A_c and Eq. (7.2.13), we have

$$P_{sc}^2 = \frac{3|B|}{4C}. \quad (7.2.16)$$

This spontaneous polarization jumps discontinuously from zero to this value at the Curie temperature, which is the most important feature of the first order phase transition.

From the Curie-Weiss law and following the work of Devonshire, we can obtain

$$A = (T - T_0) / (\varepsilon_0 C_0), \quad (7.2.17)$$

where C_0 is the Curie-Weiss constant, ε_0 is the permittivity of a vacuum with a value of $8.85 \times 10^{-12} \text{ F}\cdot\text{m}^{-1}$, and T_0 is the temperature where only the ferroelectric state can exist, which is different from the Curie temperature in the first-order transition.

Combining Eqs. (7.2.16) and (7.2.17), we can obtain the expression of the Curie temperature

$$T_c = T_0 + \frac{3B^2 \varepsilon_0 C_0}{16C}. \quad (7.2.18)$$

Another important feature of the first-order transition is that the entropy S will also change discontinuously at the transition point. Here, from the original idea of Devonshire, we assume that the coefficients of the quartic and sixth-order terms are independent of temperature. We will discuss this assumption in more detail in a later section.

Then, under the free electric and external stress forces, we have

$$dS = \frac{\partial g}{\partial T} = \frac{1}{2} P_{sc}^2 \frac{\partial A}{\partial T} = \frac{1}{2\varepsilon_0 C_0} P_{sc}^2 = \frac{3|B|}{8\varepsilon_0 C_0 C}. \quad (7.2.19)$$

The entropy S has a discontinuous jump at the Curie point with an amount of dS .

We will discuss this in the case of Barium Titanate (BaTiO_3) in more detail in the next section.

7.3 Determination of Landau Free Energy Expansion Coefficients Based on Experimental Methods

In this section, we first give a short introduction to the experimental observation of the ferroelectric characteristic phenomenon. Then we discuss how to describe the first-order phase transition of BaTiO_3 and PbTiO_3 by using the phenomenological theory after Devonshire. Finally, we focus on the experimental techniques to determine the free energy expansion coefficients.

7.3.1 The Experimental Observation of the Phase Transition Characteristics in Ferroelectrics

Experimental data have shown that the ferroelectricity of BaTiO_3 exists only in the perovskite type structure. Above the transition temperature, the BaTiO_3 has a

cubic structure, as shown in Fig. 7.4, where the barium ions lie in the center of the cubic lattice, the titanium ions at the corners of the cubic lattice, and the oxygen in the center of each Ti pair.

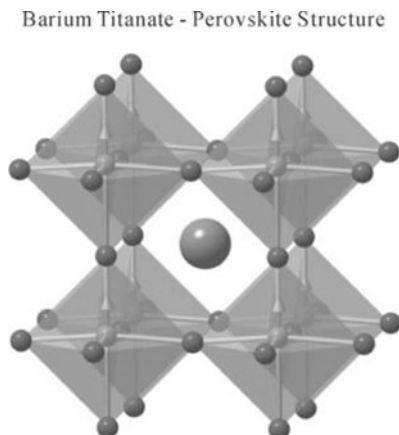


Fig. 7.4 The cubic structure of cubic BaTiO_3 perovskite. The large ball is the Ba cations, the small gray ball represents the Ti cations located in the center of the oxygen octahedron, and the small black ball is the oxygen

At high temperature, the properties of BaTiO_3 are similar to the normal dielectrics besides relatively large dielectric constants. As the temperature falls, the dielectric constants increase with the fall in temperature and reach a peak at the transition temperature. As shown in Fig. 7.5, the lattice becomes tetragonal ($P4mm$) below the transition temperature (about 393 K), and the lattice constant changes discontinuously with one of the axes (usually taken as the c axis) lengthened and the other two shortened. The ratio of c/a is 1.005 at the Curie point (Harwood et al., 1947). At the same time, evidence showed that the domains in the crystals had macroscopic polarization along the c axis $\langle 001 \rangle$. When the temperature is further decreased, there is another phase transition at about 273 K (Kay et al., 1949), with a crystal change to a lower orthorhombic ($Bmm2$) symmetry. In this phase, the c -axis shortens slightly and the a axis increases so that the two become equivalent and the polar direction is along the $\langle 011 \rangle$ direction in the ca plane. There is a further transition at about 193 K. Below this temperature, the ground state of BaTiO_3 is the rhombohedral ($R3m$) structure with the polarization axis along the $\langle 111 \rangle$ direction. The three phase transitions are all first-order with the spontaneous polarization changing discontinuously at the Curie point, which is measured as $P_{sc} = 18 \mu\text{C} \cdot \text{cm}^{-2}$ at the tetragonal to cubic transition point (Merz, 1953). Besides that, there was evidence that a hump in the specific heat curve appeared at the transition temperature, and latent heat was estimated to be $47 \text{ cal} \cdot \text{mol}^{-1}$ at the 393 K transition and $16 \text{ cal} \cdot \text{mol}^{-1}$ at 273 K.

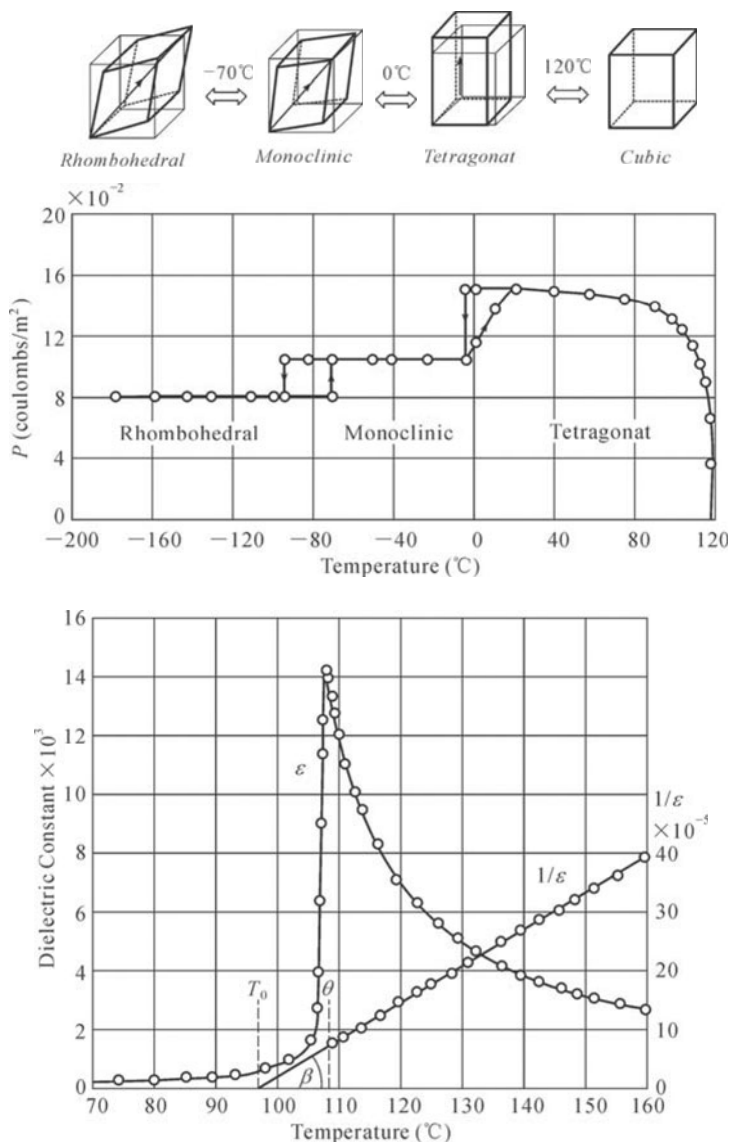


Fig. 7.5 The temperature dependence of spontaneous polarization and dielectric constants of BaTiO₃ (Merz, 1953). Copyright (1953), American Institute of Physics

A number of measurements of the dielectric constants in small fields have been made, and we cited the results of the measurement of the dielectric constants by Merz (1953) in Fig. 7.5. From these figures we can see that the dielectric constant obeys the Curie-Weiss law above the highest transition temperature, which was described by Eq. (7.2.12). At the transition temperature, the dielectric constant was 1.7×10^5 , and below this temperature the dielectric constant dropped

to about 1000–2000. As the temperature decreased, there was a slow decrease to a certain constant. There were small peaks at further lower temperatures, which corresponded to the lower crystal transitions. The discontinuities of the dielectric constants at the lower transition temperatures were very apparent.

For PbTiO_3 , there were also some experimental data accumulated in recent years. Using the X-ray diffraction technique, the evolution of the lattice constants of PbTiO_3 with the temperature was plotted in Fig. 7.6, where the extrapolation of the pseudo-cubic cell length a'_c was the cubed root of the tetragonal volume ($a_T^2 c_T$)^{1/3} in a tetragonal state (Haun et al., 1987; 1989).

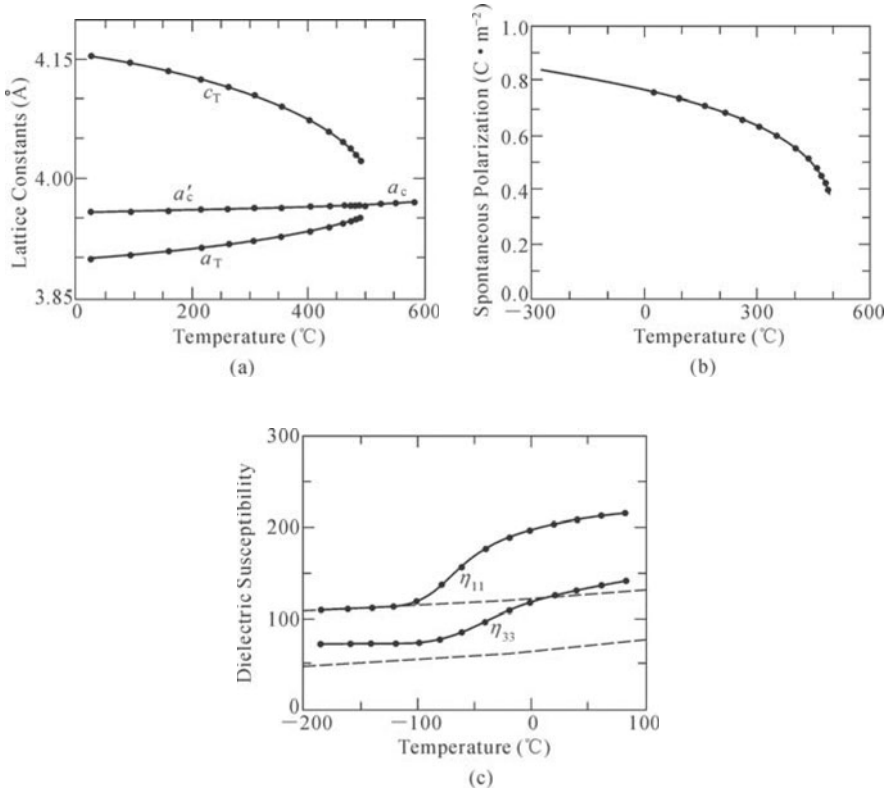


Fig. 7.6 (a) Lattice constants of PbTiO_3 as a function of temperature; (b) The temperature dependence of spontaneous polarization; (c) The dielectric susceptibility of PbTiO_3 (Haun et al., 1987). Copyright (1987), American Institute of Physics

In order to obtain more precise structural information of ferroelectric materials, the neutron-scattering technique can be used. Traditional X-ray scattering has several limitations in the structural analysis of ferroelectric materials. In X-ray scattering, the scattering is an interaction between the electromagnetic radiation field and the electrons of the atom. Therefore the amplitude of the signal is

proportional to the number of electrons. In light, with atoms with few electrons, such as the hydrogen atom, it will be difficult to detect their thermal motion with X-ray analysis. Furthermore, the position detected by X-rays is the gravity of the electron cloud, rather than the position of an atomic nucleus, which will introduce difficulty in the determination of the ferroelectric polarization contributed by the electrons at low temperature. The introduction of elastic neutron scattering has its advantages in locating the more precise positions of nuclei in the determination of the structure.

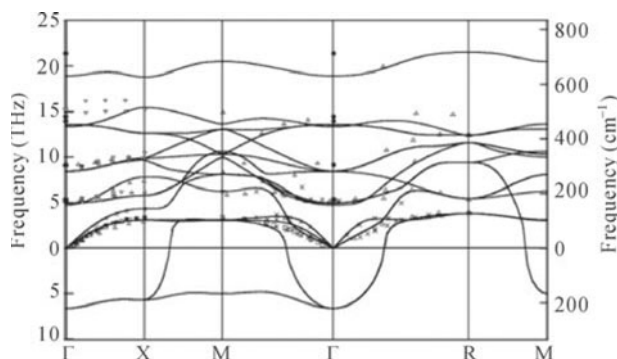


Fig. 7.7 Phonon dispersion curves of cubic BaTiO₃. The solid line is the calculated result from first-principles, with the experimental data: (●) (Anderson, 1960), (○) (Axe, 1967), (+) (Bachelet et al., 1982), (□) (Bader et al., 1981), (×) (Bader, 1990), (▽) (Baldereschi and Resta, 1983), (△) (Baroni et al., 1987). After Ph.D thesis of Ghosez (1997)

On the other hand, in the dynamic phase transition, inelastic neutron scattering has a great advantage over light scattering and X-ray scattering. It can measure both the change in the wavelength and time dependence of dynamic correlations. The typical energy of a thermal neutron is around 30 meV, which corresponds to wavelengths of 3.8 Å. Thus it can measure the changes in the energy of the phonon scattering, which is particularly useful in lattice dynamics to scan the dispersive curve in energy-momentum space. It is noted that in the ferroelectric phase transition, the energy of the soft phonon transverse mode $\hbar\omega$ will decrease to be smaller than 1 meV, and then the measurement of the soft-mode will become difficult once again, so the exact value can be obtained only by extrapolation.

7.3.2 The Phenomenological Treatment of Devonshire Theory

7.3.2.1 The Gibbs Free Energy Expansion and Phase Transition

According to considerations of symmetry, the free energy form of a ferroelectric material must satisfy the symmetry requirements of both the paraelectric and ferroelectric states. Since the paraelectric phase usually has a higher symmetric

structure than the ferroelectric phase, we need only to establish a free energy expansion to satisfy the symmetry of the paraelectric state. When the substance is a strained cubic crystal, it has a center of symmetry; all the odd ranks of the components of the free energy expansion are zero. Thus, the expression for Gibbs free energy can be expanded into a power series of the polarization, assuming isothermal conditions:

$$g = g_0 + \frac{1}{2} s_{ij} \sigma_i \sigma_j + \frac{1}{2} \alpha_{ij} P_i P_j + \frac{1}{4} \beta_{ijkl} P_i P_j P_k P_l + \frac{1}{6} \zeta_{ijklmn} P_i P_j P_k P_l P_m P_n + \frac{1}{2} Q_{kij} \sigma_k P_i P_j, \quad (7.3.1)$$

where g_0 is the energy of the paraelectric state, P_i and σ_i are the polarization and stresses; α_{ij} , β_{ijkl} , and ζ_{ijklmn} are the square-, quartic-, and sixth-order expansion coefficients at constant stress; s_{ij} is the elastic compliance coefficient at constant polarization; and Q_{kij} represents the cubic electrostrictive constants. Since the phase transition of BaTiO₃ is the first order, terms of order P^6 must be included in the energy expression.

We can start by expressing the Gibbs free energy as a function of the polarization and stresses with the stresses equivalent to zero. We have

$$g = g_0 + \frac{1}{2} \alpha_1 (P_1^2 + P_2^2 + P_3^2) + \frac{1}{4} \beta_{11} (P_1^4 + P_2^4 + P_3^4) + \frac{1}{2} \beta_{12} (P_1^2 P_2^2 + P_2^2 P_3^2 + P_1^2 P_3^2) + \frac{1}{6} \zeta_{111} (P_1^6 + P_2^6 + P_3^6) + \frac{1}{2} \zeta_{112} [P_1^2 (P_2^4 + P_3^4) + P_2^2 (P_3^4 + P_1^4) + P_3^2 (P_1^4 + P_2^4)] + \frac{1}{2} \zeta_{123} P_1^2 P_2^2 P_3^2. \quad (7.3.2)$$

The reduced notations are adopted in the above expression, i.e., by using the rules:

$$11 \rightarrow 1, 22 \rightarrow 2, 33 \rightarrow 3, 23=32 \rightarrow 4, 13=31 \rightarrow 5, 12=21 \rightarrow 6.$$

In the work of Devonshire (1949), only the term $\frac{1}{6} \zeta_{111} (P_1^6 + P_2^6 + P_3^6)$ is taken into account in the sixth-order expansion due to the minor influence of the other sixth-order terms. That is,

$$g = g_0 + \frac{1}{2}\alpha_1(P_1^2 + P_2^2 + P_3^2) + \frac{1}{4}\beta_{11}(P_1^4 + P_2^4 + P_3^4) + \frac{1}{2}\beta_{12}(P_1^2P_2^2 + P_2^2P_3^2 + P_1^2P_3^2) + \frac{1}{6}\zeta_{111}(P_1^6 + P_2^6 + P_3^6). \quad (7.3.3)$$

The derivatives of g with respect to P_1 give rise to the x -component of the electric field in the unstressed crystals.

$$E_1 = \alpha_1 P_1 + \beta_{11} P_1^3 + \beta_{12} P_1(P_2^2 + P_3^2) + \zeta_{111} P_1^5. \quad (7.3.4)$$

Similar equations for z - and y -components of the electric field can be determined in the same way. When the external electric field is absent, the right hand side of Eq. (7.3.4) should equal zero. Then the equations become

$$\begin{aligned} P_1=0, \text{ or } \alpha_1 + \beta_{11}P_1^2 + \beta_{12}(P_2^2 + P_3^2) + \zeta_{111}P_1^4 &= 0, \\ P_2=0, \text{ or } \alpha_1 + \beta_{11}P_2^2 + \beta_{12}(P_1^2 + P_3^2) + \zeta_{111}P_2^4 &= 0, \\ P_3=0, \text{ or } \alpha_1 + \beta_{11}P_3^2 + \beta_{12}(P_2^2 + P_1^2) + \zeta_{111}P_3^4 &= 0. \end{aligned} \quad (7.3.5)$$

In order to describe the first-order phase transition, we assume that ζ_{111} and β_{12} are positive, β_{11} is negative and α_1 is a function of temperature with the sign change when passing through the transition temperature. When α_1 is negative, the minimum of the free energy will give rise to polarizations of a finite value. The quartic-order terms have their minimum along certain axis and the sixth-order term has its minimum along the diagonal directions.

To minimize the Gibbs free energy in a zero electric field, Eq. (7.3.5) has the following four sets of distinguishing solutions:

$$\text{Cubic:} \quad P_1 = P_2 = P_3 = 0, \quad (7.3.6a)$$

$$\text{Tetragonal:} \quad P_1 = P_2 = 0, \alpha_1 + \beta_{11}P_3^2 + \zeta_{111}P_3^4 = 0, \quad (7.3.6b)$$

$$\text{Orthorhombic:} \quad P_1 = 0, P_2 = P_3, \alpha_1 + (\beta_{11} + \beta_{12})P_3^2 + \zeta_{111}P_3^4 = 0, \quad (7.3.6c)$$

$$\text{Rhombohedral:} \quad P_1 = P_2 = P_3, \alpha_1 + (\beta_{11} + 2\beta_{12})P_3^2 + \zeta_{111}P_3^4 = 0. \quad (7.3.6d)$$

The corresponding minima of the Gibbs free energy for the above solutions are given by:

$$g = 0, \quad (7.3.7a)$$

$$g = \frac{1}{2}\alpha_1 P_3^2 + \frac{1}{4}\beta_{11}P_3^4 + \frac{1}{6}\zeta_{111}P_3^6, \quad (7.3.7b)$$

$$g = \alpha_1 P_3^2 + \frac{1}{2}(\beta_{11} + \beta_{12})P_3^4 + \frac{1}{3}\zeta_{111}P_3^6, \quad (7.3.7c)$$

$$g = \frac{3}{2} \alpha_1 P_3^2 + \frac{3}{4} (\beta_{11} + 2\beta_{12}) P_3^4 + \frac{1}{2} \zeta_{111} P_3^6. \quad (7.3.7d)$$

We will show below how to obtain the values of ζ_{111} , β_{12} , β_{11} , and α_1 from the experimental measurements near the Curie point. Once we assume that ζ_{111} , β_{12} , β_{11} are constants, then we can find that if α_1 is positive and large enough, the least minimum of g is a nonpolarized state, that is $P=0$. When α_1 decreases and becomes negative, the minimum of g will successively shift to the state described by Eqs. (7.3.7b, 7.3.7c, and 7.3.7d). In particular, at Curie temperature, the ferroelectric state has the same energy as that of the paraelectric state. Therefore we have

$$\frac{1}{2} \alpha_{1c} P_{3c}^2 + \frac{1}{4} \beta_{11} P_{3c}^4 + \frac{1}{6} \zeta_{111} P_{3c}^6 = 0, \quad (7.3.8)$$

where α_{1c} is the value of α_1 at the Curie point. Together with Eq. (7.3.6b), we have

$$\beta_{11} = \frac{-4\alpha_{1c}}{P_{3c}^2}, \quad (7.3.9)$$

$$\zeta_{111} = \frac{3\alpha_{1c}}{P_{3c}^4}. \quad (7.3.10)$$

Since α_1 is closely related to the reciprocal susceptibility or reciprocal dielectric constants, we can determine its value from the dielectric measurement above the highest transition temperature. For BaTiO_3 with the first-order phase transition, T_0 is lower than the Curie temperature T_c . As mentioned before, α_1 is temperature dependent and has the form $\alpha_1 = (T - T_0)/(\varepsilon_0 C_0)$. Devonshire (1949) rewrote it as

$$\alpha_1 = \frac{1}{\chi} = \frac{T - T_0}{T_c - T_0} \alpha_{1c}, \quad (7.3.11)$$

where T_0 was the temperature at which α_1 vanished and could be determined from experimental observation.

From the measurement of Merz (1953), it was estimated that $T_c - T_0 = 10$ K. We can use the experimental Curie-Weiss law and Eq. (7.2.17) to determine the value of α_{1c} . From Eq. (7.3.11), the value of $\frac{\alpha_{1c}}{T_c - T_0}$ can be determined from the observed non-linear reciprocal susceptibility near the transition temperature region. Therefore, it is obtained as

$$\frac{\alpha_{1c}}{T_c - T_0} = 1.0 \times 10^{-4} \text{ K}^{-1}, \quad (7.3.12)$$

$$\alpha_{1c} = 1.0 \times 10^{-3}. \quad (7.3.13)$$

Combined with the experimental value of P_{sc} measured at the given Curie temperature, using Eqs. (7.3.9) and (7.3.10), Devonshire gave the results

$$\beta_{11} = -4.4 \times 10^{-12}, \quad (7.3.14)$$

$$\zeta_{111} = 3.7 \times 10^{-21}. \quad (7.3.15)$$

It is assumed that the β_{12} is proportional to β_{11} , that is $\beta_{12} = -a\beta_{11}$. Here a is determined by making the calculated lower transition temperatures agree with the observed value, and a value of 1.2 is obtained. Therefore, it is found that

$$\beta_{12} = 5.3 \times 10^{-21}. \quad (7.3.16)$$

With the coefficients of α_1 , ζ_{111} , β_{11} , and β_{12} , it is shown that as the temperature decreases from a high temperature, the minimum of the Gibbs free energy shifts from Eq. (7.3.7). We can get more insight from Eq. (7.3.3), since α_1 is highly positive above the Curie temperature, and the spontaneous polarization is prohibited by the positive values of α_1 due to its large contribution to the free energy, that is corresponding to the state of (a) no polarization. When the temperature is decreased, α_1 becomes small and the negative value of β_{11} will favor spontaneously polarized crystals. While the positive β_{12} indicates the interaction between the components of the spontaneous polarization, as a result of their interaction the polarization can have only one nonzero component, which is (b) spontaneous polarization along one axis direction such as $\langle 001 \rangle$. When the temperature is further decreased, below the T_0 , α_1 becomes negative and the spontaneous polarization becomes larger and more stable. When the temperature is below a certain value, α_1 becomes more negative, the combined negative effects of β_{11} and α_1 will overcome that of β_{12} , and more components of the polarization will appear. That is (c) spontaneous polarization along the plane diagonal $\langle 011 \rangle$ below 273 K, and (d) spontaneous polarization along the body diagonal $\langle 111 \rangle$ below 193 K. The competition of the free energy determined by the temperature dependent coefficients drives the phase transition from one phase to another in BaTiO_3 .

7.3.2.2 The Strain Coupling and the Helmholtz Free Energy Expansion

In ferroelectrics, the strain is always accompanied by spontaneous polarization. Taking the cubic crystal as the zero free energy reference, Devonshire gave the

Helmholtz free energy of BaTiO₃ as a function of polarization and strain after considering the crystal symmetry:

$$\begin{aligned}
 F = & \frac{1}{2}c_{11}(\eta_1^2 + \eta_2^2 + \eta_3^2) + c_{12}(\eta_2\eta_3 + \eta_3\eta_1 + \eta_1\eta_2) + \frac{1}{2}c_{44}(\eta_4^2 + \eta_5^2 + \eta_6^2) \\
 & + \frac{1}{2}\alpha'_1(P_1^2 + P_2^2 + P_3^2) + \frac{1}{4}\beta'_{11}(P_1^4 + P_2^4 + P_3^4) \\
 & + \frac{1}{2}\beta'_{12}(P_1^2P_2^2 + P_2^2P_3^2 + P_1^2P_3^2) + q_{11}(\eta_1P_1^2 + \eta_2P_2^2 + \eta_3P_3^2) \\
 & + q_{12}\{\eta_1(P_2^2 + P_3^2) + \eta_2(P_1^2 + P_3^2) + \eta_3(P_1^2 + P_2^2)\} \\
 & + q_{44}(\eta_4P_2P_3 + \eta_5P_1P_3 + \eta_6P_1P_2).
 \end{aligned} \tag{7.3.17}$$

Only the quartic-order terms are given and it is assumed that the strain is the second order. The components of the electric field and stress can be obtained by

$$\sigma_i = \frac{\partial F}{\partial \eta_i}, \tag{7.3.18}$$

$$E_i = \frac{\partial F}{\partial P_i}. \tag{7.3.19}$$

In the absence of external force, we can obtain the following relationship by substituting the results solved from Eq. (7.3.18) into Eq. (7.3.17):

$$\alpha_1 = \alpha'_1, \tag{7.3.20a}$$

$$\beta_{11} = \beta'_{11} + 2 \frac{-q_{11}^2(c_{11} + c_{12}) + 4q_{11}q_{12}c_{12} - 2q_{12}^2c_{11}}{(c_{11} - c_{12})(c_{11} + 2c_{12})}, \tag{7.3.20b}$$

$$\beta_{12} = \beta'_{12} + 2 \frac{q_{11}^2c_{12} - 2q_{11}q_{12}c_{11} + q_{12}^2(-c_{11} + 2c_{12})}{(c_{11} - c_{12})(c_{11} + 2c_{12})} - \frac{q_{44}^2}{c_{44}}, \tag{7.3.20c}$$

where β'_{11} and β'_{12} are the coefficients of quartic-order terms in the free energy expansion of clamped crystals, which are more important than free crystals because they reflect the practical situation of ferroelectrics in reality. In particular, the strains are proportional to the square of the polarization. As the temperature decreases, the strains in the cubic, tetragonal, orthorhombic and rhombohedral phases have the following expressions:

$$\eta_1 = \eta_2 = \eta_3 = \eta_4 = \eta_5 = \eta_6 = 0, \tag{7.3.21a}$$

$$\eta_1 = \eta_2 = Q_{12}P_3^2, \eta_3 = Q_{11}P_3^2, \eta_4 = \eta_5 = \eta_6 = 0, \tag{7.3.21b}$$

$$\eta_1 = 2Q_{12}P_3^2, \eta_2 = \eta_3 = (Q_{11} + Q_{12})P_3^2, \eta_4 = Q_{44}P_3^2, \eta_5 = \eta_6 = 0, \tag{7.3.21c}$$

$$\eta_4 = \eta_5 = \eta_6 = Q_{44}P_3^2, \quad \eta_1 = \eta_2 = \eta_3 = (Q_{11} + 2Q_{12})P_3^2, \quad (7.3.21d)$$

where

$$Q_{11} = \frac{-q_{11}(c_{11} + c_{12}) + 2q_{12}c_{12}}{(c_{11} - c_{12})(c_{11} + 2c_{12})}, \quad Q_{12} = \frac{q_{11}c_{12} - q_{12}c_{11}}{(c_{11} - c_{12})(c_{11} + 2c_{12})}, \quad Q_{44} = \frac{-q_{44}^2}{c_{44}}. \quad (7.3.22)$$

In order to determine the electrostrictive parameters, we can compare the information from the lattice constants, such as the difference in strain deformation along the c and a axes, e.g., $\eta_3 - \eta_1 = \frac{q_{12} - q_{11}}{c_{11} - c_{12}}P_3^2$, the volume expansion

$$\eta_1 + \eta_2 + \eta_3 = \frac{2(-q_{11} - 2q_{12})}{c_{11} + 2c_{12}}P_3^2 \text{ for the orthorhombic phase, and the shear angle of}$$

the rhombohedral structure $\eta_6 = -q_{44}^2P_3^2 / c_{44}$. By using a molecular theory model, the elastic constants are obtained in a rough approximation in the original work of Devonshire. As a result, the following values are obtained for BaTiO₃:

$$\begin{aligned} c_{11} &= 600 \text{ GPa}, \quad c_{12} = c_{44} = 300 \text{ GPa}, \\ q_{11} &= -11.2, \quad q_{12} = 0.8, \quad q_{44} = -8.1, \\ \beta'_{11} &= 65.6 \times 10^{-12}, \quad \beta'_{12} = -0.2 \times 10^{-12}. \end{aligned} \quad (7.3.23)$$

Together with the value from Eqs. (7.3.12) – (7.3.16), Devonshire used them to study the first-order phase transition of BaTiO₃ within Landau's mean field theory, which well explained the rich phenomena observed in experiments.

7.3.2.3 The Landau-Devonshire Free Energy Coefficients Obtained from Experiment

By using more precise equipment and advanced techniques, more experimental data were obtained. It was found that the coefficients of quartic order terms were also temperature dependent (Hubregtse and Young, 1956). The parameters given below are after Bell (1984):

$$\begin{aligned} \alpha_1 &= 3.34(T - T_0) \times 2 \times 10^5 \text{ C}^{-2} \cdot \text{m}^2 \cdot \text{N}, \text{ where } T_0 = 381 \text{ K}, \\ \beta_{11} &= (4.69(T - T_2) - 202) \times 4 \times 10^6 \text{ C}^{-4} \cdot \text{m}^6 \cdot \text{N}, \text{ where } T_2 = 393 \text{ K}, \\ \beta_{12} &= 3.23 \times 4 \times 10^8 \text{ C}^{-4} \cdot \text{m}^6 \cdot \text{N} \\ \zeta_{111} &= (-5.52(T - 120) + 276) \times 6 \times 10^7 \text{ C}^{-6} \cdot \text{m}^{10} \cdot \text{N}, \\ \zeta_{112} &= 4.47 \times 2 \times 10^9 \text{ C}^{-6} \cdot \text{m}^{10} \cdot \text{N}, \\ \zeta_{113} &= 4.919 \times 2 \times 10^9 \text{ C}^{-6} \cdot \text{m}^{10} \cdot \text{N}. \end{aligned} \quad (7.3.24)$$

The electrostrictive constants were also given by Ehrhart (1991) using the methods illustrated above:

$$Q_{11}=0.10 \text{ m}^4 \cdot \text{C}^{-2}, \quad Q_{12}=-0.034 \text{ m}^4 \cdot \text{C}^{-2}, \quad Q_{44}=0.029 \text{ m}^4 \cdot \text{C}^{-2}. \quad (7.3.25)$$

From the first-principles calculation and experimental observation, there was still an unknown monoclinic phase existing in epitaxially grown BaTiO₃ ferroelectric thin film (Noheda et al., 1999; Dieguez et al., 2004). In order to describe this new phase, Vanderbilt and Cohen (2001) suggested that the eighth order term should also be included in the free energy expansion. Then a new set of input parameters were given by Li (2005) as shown in Table 7.2:

$$\begin{aligned} g = & g_0 + \alpha_1(P_1^2 + P_2^2 + P_3^2) + \beta_{11}(P_1^4 + P_2^4 + P_3^4) \\ & + \beta_{12}(P_1^2 P_2^2 + P_2^2 P_3^2 + P_1^2 P_3^2) + \zeta_{111}(P_1^6 + P_2^6 + P_3^6) + \zeta_{123}P_1^2 P_2^2 P_3^2 \\ & + \zeta_{112} \left[P_1^2(P_2^4 + P_3^4) + P_2^2(P_3^4 + P_1^4) + P_3^2(P_1^4 + P_2^4) \right] + \xi_{1111}(P_1^8 + P_2^8 + P_3^8) \\ & + \xi_{1112} \left[P_1^6(P_2^2 + P_3^2) + P_2^6(P_3^2 + P_1^2) + P_3^6(P_1^2 + P_2^2) \right] \\ & + \xi_{1122}(P_1^4 P_2^4 + P_2^4 P_3^4 + P_1^4 P_3^4) + \xi_{1123}(P_3^4 P_1^2 P_2^2 + P_1^4 P_2^2 P_3^2 + P_2^4 P_1^2 P_3^2). \end{aligned} \quad (7.3.26)$$

Table 7.2 Landau-Devonshire potential coefficients for BaTiO₃

Coefficient	Value
$\alpha_1 (\times 10^5 \text{ C}^{-2} \cdot \text{m}^2 \cdot \text{N})$	4.124 ($T=388 \text{ K}$)
$\beta_{11} (\times 10^6 \text{ C}^{-4} \cdot \text{m}^6 \cdot \text{N})$	-209.7
$\beta_{12} (\times 10^8 \text{ C}^{-4} \cdot \text{m}^6 \cdot \text{N})$	7.974
$\zeta_{111} (\times 10^7 \text{ C}^{-6} \cdot \text{m}^{10} \cdot \text{N})$	129.4
$\zeta_{112} (\times 10^9 \text{ C}^{-6} \cdot \text{m}^{10} \cdot \text{N})$	-1.950
$\zeta_{123} (\times 10^9 \text{ C}^{-6} \cdot \text{m}^{10} \cdot \text{N})$	-2.5009
$\xi_{1111} (\times 10^{10} \text{ C}^{-8} \cdot \text{m}^{14} \cdot \text{N})$	3.863
$\xi_{1112} (\times 10^{10} \text{ C}^{-8} \cdot \text{m}^{14} \cdot \text{N})$	2.529
$\xi_{1122} (\times 10^{10} \text{ C}^{-8} \cdot \text{m}^{14} \cdot \text{N})$	1.637
$\xi_{1123} (\times 10^{10} \text{ C}^{-8} \cdot \text{m}^{14} \cdot \text{N})$	1.367
$c_{11} (\times 10^{11} \text{ N} \cdot \text{m}^{-2})$	1.78
$c_{12} (\times 10^{11} \text{ N} \cdot \text{m}^{-2})$	0.964
$c_{44} (\times 10^{11} \text{ N} \cdot \text{m}^{-2})$	1.22
$Q_{11} (\times 10^{-2} \text{ m}^4 \cdot \text{C}^{-2})$	11
$Q_{12} (\times 10^{-2} \text{ m}^4 \cdot \text{C}^{-2})$	4.5
$Q_{44} (\times 10^{-2} \text{ m}^4 \cdot \text{C}^{-2})$	5.9

7.3.3 The Elastic Gibbs Free Energy of PbTiO_3 and Its Coefficients

Similar to BaTiO_3 , the Landau thermodynamic theory can also apply to other ferroelectrics (Mason, 1947; Matthias et al., 1956), such as $\text{NaKC}_4\text{H}_4\text{O}_6 \cdot 4\text{H}_2\text{O}$, $(\text{NH}_4)_2\text{Cd}_2(\text{SO}_4)_3$, PbTiO_3 and KH_2PO_4 , or KDP. Due to its large and important technological applications in industry, we give a brief Landau-Devonshire (LD) explanation of the thermodynamic theory of PbTiO_3 in the following. The LD theory is demonstrated to be a particularly useful method to provide the complete description of the intrinsic properties of BaTiO_3 as introduced above. But it is a little bit difficult for the extension to PbTiO_3 , mainly because of the lack of exact experimental data. As there are high electric conductivity and large tetragonal distortion occurring at the paraelectric to ferroelectric transitions (765 K), a purified crystal and sensitive apparatus are required. After Remeika and Glass (1970) had determined some of the coefficients of the elastic Gibbs free-energy functions, many other authors (Gavrilyachenko et al., 1970; Haun et al., 1987; Amin et al., 1985) did more experimental measurements. Haun (1987) obtained more reliable data for PbTiO_3 which we will review below.

The same as for BaTiO_3 , after consideration of the symmetry of the paraelectric phase, the elastic Gibbs free energy was expanded as:

$$\begin{aligned} g = & g_0 + \alpha_1(P_1^2 + P_2^2 + P_3^2) + \beta_{11}(P_1^4 + P_2^4 + P_3^4) + \beta_{12}(P_1^2 P_2^2 + P_2^2 P_3^2 + P_1^2 P_3^2) \\ & + \zeta_{111}(P_1^6 + P_2^6 + P_3^6) + \zeta_{112} \left[P_1^2(P_2^4 + P_3^4) + P_2^2(P_3^4 + P_1^4) + P_3^2(P_1^4 + P_2^4) \right] \\ & + \zeta_{123} P_1^2 P_2^2 P_3^2 - \frac{1}{2} s_{11}(\sigma_1^2 + \sigma_2^2 + \sigma_3^2) - s_{12}(\sigma_2 \sigma_3 + \sigma_3 \sigma_1 + \sigma_1 \sigma_2) \\ & - \frac{1}{2} s_{44}(\sigma_4^2 + \sigma_5^2 + \sigma_6^2) - Q_{12} \{ \sigma_1(P_2^2 + P_3^2) + \sigma_2(P_1^2 + P_3^2) + \sigma_3(P_1^2 + P_2^2) \} \\ & - Q_{44}(\sigma_4 P_2 P_3 + \sigma_5 P_1 P_3 + \sigma_6 P_1 P_2) - Q_{11}(\sigma_1 P_1^2 + \sigma_2 P_2^2 + \sigma_3 P_3^2). \end{aligned} \quad (7.3.27)$$

For PbTiO_3 , it has four solutions similar to Eq. (7.3.7), but the orthorhombic and rhombohedral phases are metastable, and only the cubic and tetragonal phases are stable. When using the cubic phase as the reference state, the spontaneous polarization is given by Eq. (7.3.6), and the spontaneous elastic strains are described by Eq. (7.3.21b). The coefficients needed for calculations are the dielectric stiffness α_1 , higher-order dielectric stiffness β_{11} , β_{12} , ζ_{111} , and ζ_{112} , the elastic compliance coefficients s_{11} , s_{12} , and s_{44} , and the electrostrictive constants Q_{11} , Q_{12} and Q_{44} . These coefficients can be determined from experimental data in a similar way to that shown before in the case of BaTiO_3 .

With these coefficients, other properties can also be obtained, such as the relative dielectric susceptibility coefficients:

$$\begin{aligned}
\chi_{11} = \chi_{22} &= \left[(2\alpha_1 + 2\beta_{12}P_3^2 + 2\zeta_{112}P_3^4) \varepsilon_0 \right]^{-1}, \\
\chi_{33} &= \left[(2\alpha_1 + 12\beta_{11}P_3^2 + 30\zeta_{111}P_3^4) \varepsilon_0 \right]^{-1}, \\
\chi_{12} = \chi_{13} = \chi_{23} &= 0.
\end{aligned} \tag{7.3.28}$$

The piezoelectric voltage coefficients:

$$g_{33} = 2Q_{11}P_3, \quad g_{31} = 2Q_{12}P_3, \quad g_{15} = Q_{44}P_3. \tag{7.3.29}$$

Piezoelectric charge coefficients:

$$d_{33} = \varepsilon_0 \chi_{33} g_{33}, \quad d_{31} = \varepsilon_0 \chi_{33} g_{31}, \quad d_{15} = \varepsilon_0 \chi_{11} g_{15}. \tag{7.3.30}$$

The lattice parameter for PbTiO_3 is $3.8995 \text{ \AA}$ for the in-plane lattice constant, and the c/a ratio is 1.066 in its tetragonal ground state. The dielectric stiffness coefficients α_1 , β_{11} , and ζ_{111} were calculated from Eqs. (7.3.9) – (7.3.11), and the electrostrictive constants Q_{11} , Q_{12} , and Q_{44} can be calculated from Eqs. (7.3.29) – (7.3.30) and Eq. (7.3.21). Using the experimental data of temperature dependent dielectric susceptibility coefficients (χ_{11} , χ_{33}) and by the fitting of theoretical values of Eq. (7.3.28), we can determine the dielectric stiffness coefficients β_{12} and ζ_{112} . The same method is applicable to other ferroelectrics such as $\text{PbZr}_{1-x}\text{Ti}_x\text{O}_3$ (PZT). As a result, some of the expansion coefficients derived from measured data of the phase transition are listed in Table 7.3.

Table 7.3 Constants and coefficients derived from the experimental data (Haun et al., 1987; 1989; Pertsev et al., 1998)

Parameters	PbTiO_3	$\text{PbZr}_{0.3}\text{Ti}_{0.7}\text{O}_3$	$\text{PbZr}_{0.5}\text{Ti}_{0.5}\text{O}_3$	$\text{PbZr}_{0.52}\text{Ti}_{0.48}\text{O}_3$
T_c (°C)	492.2	334.4	394.10	399.27
T_0 (°C)	478.8	334.5	392.72	398.13
C_0 ($\times 10^5$ °C)	1.5	2.15	4.25	4.11
P_{sc} ($\text{C}\cdot\text{m}^{-2}$)	0.373		0.38	0.38
Q_{11} ($\times 10^{-2} \text{ m}^4\cdot\text{C}^{-2}$)	8.9	6.17	9.66	9.66
Q_{12} ($\times 10^{-2} \text{ m}^4\cdot\text{C}^{-2}$)	-2.6	-6.81	1.92	1.92
Q_{44} ($\times 10^{-2} \text{ m}^4\cdot\text{C}^{-2}$)	6.75	2.76	4.095	4.095
α_1 ($\times 10^5 \text{ m}\cdot\text{F}^{-1}$)	$3.76 \times (T - T_0)$	$2.63 \times (T - T_0)$	$1.33 \times (T - T_0)$	$1.37 \times (T - T_0)$
α_1 ($\times 10^5 \text{ m}\cdot\text{F}^{-1}$)	$3.76 \times (T - T_0)$	$2.63 \times (T - T_0)$	$1.33 \times (T - T_0)$	$1.37 \times (T - T_0)$
β_{11} ($\times 10^7 \text{ m}^5\cdot\text{C}^{-2}\cdot\text{F}$)	-7.252	22.304	4.764	4.922
β_{12} ($\times 10^8 \text{ m}^5\cdot\text{C}^{-2}\cdot\text{F}$)	7.5	1.688	1.735	1.793
ζ_{111} ($\times 10^8 \text{ m}^9\cdot\text{C}^{-4}\cdot\text{F}$)	2.606	3.56	1.336	1.38
ζ_{112} ($\times 10^8 \text{ m}^9\cdot\text{C}^{-4}\cdot\text{F}$)	6.1	15.27	6.128	6.33
s_{11} ($\text{C}^{-12}\cdot\text{m}^{-2}\cdot\text{N}^{-1}$)	8.0			
s_{12} ($\text{C}^{-12}\cdot\text{m}^{-2}\cdot\text{N}^{-1}$)	-2.5			
s_{44} ($\text{C}^{-12}\cdot\text{m}^{-2}\cdot\text{N}^{-1}$)	9.0			

In summary, the coefficients of the expanded Gibbs free energy function were determined from the experimental data collected in the literature, and the theory was found to predict the spontaneous polarization and strains very well. Owing to the successful application of the Landau-Devonshire theory on the crystallized ferroelectrics, it is natural to apply the theory to more complex structures with domain or interface, which is still not quite clear in experimental measurements. Besides that, application of the Landau theory to multiferroic materials is a rapid developing area, where the electrical, mechanical and magnetic properties are coupled together.

7.3.4 The Determination of the Expansion Coefficients from the First-Principle Calculation Based on the Effective Hamiltonian Method

It is worth mentioning that in recent years the first-principle based calculation was developed to account for the influence of temperature and provided much useful information on the microscopic scale. In a recent work of the first-principle based effective Hamiltonian method (Iniguez et al., 2001), the Devonshire-Landau potential of BaTiO_3 was studied in the phase transition. And the coefficients of the free energy were determined from an effective Hamiltonian method with parameters determined from first-principle. By using Monte Carlo simulations with the effective Hamiltonian of BaTiO_3 obtained from *ab initio*, Iniguez calculated the temperature dependent properties of the BaTiO_3 , thus linking the phenomenological Devonshire-Landau theory with the microscopic density functional calculations. It was found that there were some nontrivial temperature dependence for all the coefficients in the expansion, including the quadratic and sixth-order terms, as shown in Fig. 7.8. In particular, the sixth-order terms were nonzero only at the transition temperature range.

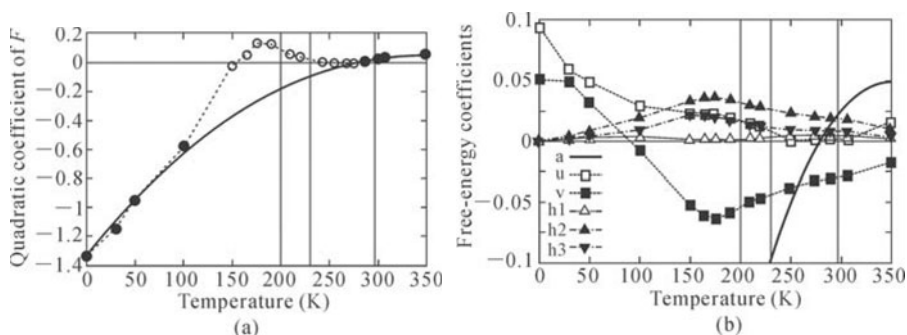


Fig. 7.8 The expanded coefficients as a function of the temperature, for (a) the quadratic coefficients (solid line), (b) the higher order expansion coefficients. Reprinted with permission from Iniguez (2001). Copyright (2000), American Institute of Physics

These findings were consistent with the other authors (Troster et al., 2005; Radescu et al., 1995; Shavrtsman and Kholkin, 2004; Wang et al., 2006), and provided legitimate concern that all Landau coefficients depended on temperature, and needed more investigation in this area.

7.4 Gradient Terms in the Landau-Devonshire-Ginzburg Free Energy Expansion

7.4.1 *The Consideration of Spatial Non-Uniform Distribution of the Order Parameters in Landau Theory*

In the above introductions, the Landau-Devonshire theory is equivalent to the self-consistent mean field theory, and all the non-uniform changes in the order parameters are neglected. So the Landau-Devonshire is reliable only in describing the system where the contribution of the fluctuation is small, or for a polar bulk ferroelectrics near its transition with polarization that is spatially uniform. The Landau free energy expression is expanded near the phase transition and when far away from it the foundation of the expansion is not valid. On the other hand, as the temperature approaches the transition temperature, the fluctuation in the order parameter is so large that it will also lead to the invalidation of Landau theory. In fact, the order parameters are spatially dependent variables. In this regard, Ginzburg incorporated an additional contribution in the free energy density (Levanyuk, 1959; Ginzburg, 1960), as the first term in the expansion involving the gradient terms, the gradient energy ΔF that was promotional to the spatial gradient of the order parameter $|\nabla P|^2$, and it was assumed that the fluctuations in the order parameter at different wave vectors were independent. Therefore, the Landau-Ginzburg-Devonshire theory incorporated small spatial variations of the polarization. Thus it can be used to treat spatially non-uniform polarization in nanoscale ferroelectrics.

After the consideration of the spatial variations of the order parameters, the density of the free energy in a ferroelectric system can be expanded as:

$$g(T, E) = g_0 + \frac{A}{2} P(r)^2 + \frac{B}{4} P(r)^4 + \frac{C}{6} P(r)^6 + D[\nabla P(r)]^2 - EP, \quad (7.4.1)$$

where g_0 is the free energy per unit cell at the high temperature and highly-symmetric paraelectric phase, ∇P is the spatial gradient of the polarization, and the coefficient D measures the contribution of the spatial heterogeneity to the free energy. The change in the total energy with volume V is given by $G = \iiint [g(T, E, \sigma) - g_0] dV$.

7.4.2 The Critical Region and the Applicability of Landau Mean Field Theory

For simplicity, we consider the case of single-component polarization, i.e., the tetragonal phase of the BaTiO₃. In the absence of an external strain field, the change in the total free energy is

$$\Delta G(T, E) = \iiint \left\{ \frac{A}{2} P(r)^2 + D [\nabla P(r)]^2 + \frac{B}{4} P(r)^4 + \frac{C}{6} P(r)^6 - EP \right\} dV. \quad (7.4.2)$$

When the polarization at different positions is non-parallel, the second order term is the major contribution to the free energy expansion. After neglecting the sixth order terms, minimizing the Landau-Ginzburg-Devonshire (LGD) free energy as a function of the spontaneous polarization $\frac{\delta G}{\delta P} = 0$, we obtain

$$(A + BP^2 - 2D\nabla^2)P = E. \quad (7.4.3)$$

If the variation in the spontaneous polarization is small, we can drop the gradient term and the equation is

$$(A + BP^2)P = E. \quad (7.4.4)$$

And in the absence of an external field, the order parameter is given by $P^2 = -A/B$, which yields

$$P = \sqrt{\frac{T_c - T}{B\epsilon_0 C_0}}. \quad (7.4.5)$$

When the fluctuation in the order parameter is large, and the associate fluctuation energy is comparable to the total free energy, the traditional mean field theory will break down. In this case, we introduce the statistical definition of the mean spontaneous polarization, as done by Lines and Glass (1977):

$$\langle P(r) \rangle = \frac{\text{Tr}(P(r) \exp\{-\beta[G_0 - \int P(r')E(r')dr']\})}{\text{Tr}(\exp\{-\beta[G_0 - \int P(r')E(r')dr']\})}, \quad (7.4.6)$$

where $\beta = 1/(k_B T)$, G_0 is the macroscopic free energy in the absence of an applied field.

Therefore, we have the variation in the order parameter using the classical

differentiation:

$$\Delta \langle P(r) \rangle = \beta \int \Gamma(r, r') \Delta E(r') dr' , \quad (7.4.7)$$

where the two-site correlation function $\Gamma(r, r')$ is given by

$$\Gamma(r, r') = \langle \{P(r) - \langle P(r) \rangle\} \{P(r') - \langle P(r') \rangle\} \rangle . \quad (7.4.8)$$

Since the ferroelectric dipole-dipole interaction is a long range interaction, we can still use the classical treatment. Assuming that the polarization and external field have increments $\Delta \langle P(r) \rangle$ and $\Delta E(r)$ in Eq. (7.4.3), we have

$$(A + 3B \langle P(r) \rangle^2 - 2D \nabla^2) \Gamma(r, r') = k_B T \Delta(r - r') . \quad (7.4.9)$$

In the limit of zero applied fields, the correlation function has a solution for spontaneous fluctuation in the form:

$$\Gamma(r, r') = \frac{k_B T}{8\pi D |r - r'|} \exp(-|r - r'| / \xi) , \quad (7.4.10)$$

where the correlation length ξ is given by

$$\xi = \begin{cases} \sqrt{\frac{2\varepsilon_0 C_0 D}{T - T_c}} , & \text{for } T > T_c , \\ \sqrt{\frac{\varepsilon_0 C_0 D}{T_c - T}} , & \text{for } T < T_c . \end{cases} \quad (7.4.11)$$

ξ diverges at $T = T_c$, indicating that fluctuations become very large in a greater space. For $T < T_c$ it measures the length scale over which the polarization fluctuation is correlated with the other polarization fluctuation.

When the fluctuation is very large, the Landau theory is no longer valid. According to the criterion of Ginzburg, the region where Landau theory could be used was the region where the fluctuation of the order parameters was quite small when compared with the order parameter itself, i.e. $\Gamma(r) \ll P^2$. From Eqs. (7.4.5), (7.4.10) and (7.4.11), we choose $|r - r'| = \xi$; thus the Ginzburg criterion in $T < T_c$ is:

$$\frac{T_c - T}{\varepsilon_0 C_0 B} \gg \frac{k_B T}{8\pi D \xi e} , \quad (7.4.12)$$

where B is the coefficient of the square order term in the free energy expansion. We note that the coefficient of the gradient term plays an important role in the correlation length and the Ginzburg criterion within which the Landau expansion is valid.

Considering the jump of the heat capacity at transition temperature T_c , $\Delta c = \frac{T_c}{2(\epsilon_0 C_0)^2 B}$, and rewriting the correlation length as $\xi = \xi(0)|t|^{-1/2}$, where $t = \frac{T - T_c}{T_c}$, we can obtain a more physical expression for the Ginzburg criterion in terms of $\xi(0)$ and Δc :

$$|t|^{-1/2} \gg \frac{k_B}{16\pi e \Delta c} \xi(0)^{-3}. \quad (7.4.13)$$

It is clear that the region of the critical temperature is inversely proportional to the sixth order of the zero temperature correlation length. The larger the zero temperature correlation is, the smaller the critical region will be. Since the dipole-dipole interaction is a long range force, the critical region is very small for most ferroelectrics. For example, the TGS has a critical temperature of 10^{-2} K near the transition temperature (Deguchi and Nakamura, 1972).

7.4.3 Determination of the Gradient Terms from the Lattice Dynamic Theory

In the above phenomenological theory, all the parameters can be derived from experimental measurement. In order to connect the phenomenological theory with its microscopic origin, we must turn to the microscopic theory such as the lattice dynamics or transverse Ising models. We will give a short introduction to the lattice dynamics in this section, leaving the transverse Ising models to the next section.

In general, the Ginzburg gradient terms can be expressed as the following lowest order terms in spatial derivatives of the polarization based on symmetry considerations:

$$\begin{aligned} g_{\text{Ginz}} = & \frac{D_{11}}{2} (P_{1,1}^2 + P_{2,2}^2 + P_{3,3}^2) + D_{12} (P_{1,1}P_{2,2} + P_{1,1}P_{3,3} + P_{2,2}P_{3,3}) \\ & + \frac{D_{44}}{2} [(P_{1,2} + P_{2,1})^2 + (P_{2,3} + P_{3,2})^2 + (P_{1,3} + P_{3,1})^2]. \end{aligned} \quad (7.4.14)$$

where $P_{i,j} = \frac{\partial P_i}{\partial r_j}$. This expression is similar to the expression of the elastic energy density in Eq. (7.3.18) and $D_{11}=D_{1111}$, $D_{12}=D_{1122}$, and $D_{44}=D_{1212}$. Some authors (Hu and Chen, 1998; Nambu and Sagala, 1994; Gao and Suo, 2002) also include a further D'_{44} term in Eq. (7.4.14):

$$\frac{D'_{44}}{2} \left[(P_{1,2} - P_{2,1})^2 + (P_{2,3} - P_{3,2})^2 + (P_{1,3} - P_{3,1})^2 \right]. \quad (7.4.15)$$

As mentioned by Hlinka (2006a), the D'_{44} term can be transformed to a combination of D_{12} and D_{44} due to the equivalence of the volume integral of $P_{1,2}P_{2,1}$ and $P_{1,1}P_{2,2}$. Only the symmetrical part of the P_{ij} matrix makes an important contribution in the Ginzburg gradient functions.

The gradient terms play an important role in nanoscale materials and are also crucial for a quantitative description of the domain walls in ferroelectrics within the phenomenological theory. The determination of the gradient coefficients can make a significant contribution in the understanding of the polarization fluctuation and in the investigation and applications of nano-electronic devices.

In order to derive the coefficients D_{11} , D_{12} and D_{44} of the gradient terms, Cao (1994) developed a lattice dynamics model within the soft mode theory, and gave some relationships between those gradient coefficients and the soft-phonon modes, which we will briefly introduce in the following.

We start from Eq. (7.4.1). When taking the paraelectric phase as the reference, i.e., $g_0=0$, combined with Eq. (7.4.14), to the lowest order the Landau-Ginzburg-Devonshire free energy density is

$$g = \frac{A}{2} (P_1^2 + P_2^2 + P_3^2) + g_{\text{Ginz}}, \quad (7.4.16)$$

where $A \propto (T - T_c)$ and the Ginzburg term takes the inhomogeneous order parameter into account. This expression is consistent with the soft mode theory, where the potential energy of the system can be expanded in terms of the relative displacement field u in a cubic ferroelectric system:

$$\begin{aligned} G = & \frac{A}{2} (u_1^2 + u_2^2 + u_3^2) + \frac{D_{11}}{2} (u_{1,1}^2 + u_{2,2}^2 + u_{3,3}^2) \\ & + D_{12} (u_{1,1}u_{2,2} + u_{1,1}u_{3,3} + u_{2,2}u_{3,3}) \\ & + \frac{D_{44}}{2} \left[(u_{1,2} + u_{2,1})^2 + (u_{2,3} + u_{3,2})^2 + (u_{1,3} + u_{3,1})^2 \right]. \end{aligned} \quad (7.4.17)$$

In fact, the spontaneous polarization is proportional to the eigenvector of the soft mode, i.e., the displacement field u . In a harmonic approximation, we can obtain the equation of motion $\ddot{u}_i = \frac{\delta G}{\delta u_i}$ in the following form:

$$\begin{cases} M\ddot{u}_1 - D_{11}u_{1,11} - D_{12}(u_{2,21} + u_{3,31}) - D_{44}(u_{1,22} + u_{2,12} + u_{1,33} + u_{3,13}) - Au_1 = 0, \\ M\ddot{u}_2 - D_{11}u_{2,22} - D_{12}(u_{1,12} + u_{3,32}) - D_{44}(u_{1,21} + u_{2,11} + u_{2,33} + u_{3,23}) - Au_2 = 0, \\ M\ddot{u}_3 - D_{11}u_{3,33} - D_{12}(u_{1,13} + u_{2,23}) - D_{44}(u_{1,31} + u_{3,11} + u_{2,32} + u_{3,22}) - Au_3 = 0, \end{cases} \quad (7.4.18)$$

where M is the effective mass of the phonon mode, it can relate with the mode-plasma frequency Ω_p using $M^{-1} = \varepsilon_b \Omega_p^2$ (Hlinka et al., 2006b), where ε_b is the background dielectric constant. Eq. (7.4.18) has the following plane-wave solution:

$$u_i = U e^{i(k \cdot x - \omega t)}. \quad (7.4.19)$$

Substituting Eq. (7.4.19) into Eq. (7.4.18) leads to the following eigenvalue problem:

$$\tilde{H}U = M\omega^2 U, \quad (7.4.20)$$

where the dynamical matrix $\mathbf{H}$ is

$$\mathbf{H} = \begin{pmatrix} A + D_{11}k_1^2 + D_{44}(k_2^2 + k_3^2) & D_{12}k_1k_2 & D_{12}k_1k_3 \\ D_{12}k_1k_2 & A + D_{11}k_2^2 + D_{44}(k_1^2 + k_3^2) & D_{12}k_3k_2 \\ D_{12}k_1k_3 & D_{12}k_3k_2 & A + D_{11}k_3^2 + D_{44}(k_2^2 + k_1^2) \end{pmatrix}. \quad (7.4.21)$$

In a finite size ferroelectric system, such as a nanowire, thin film and capacitor structure, the polarization is normal to the surface, and the unscreened surface polarized charge will induce a depolarizing electric field E_d , which will split the degenerated longitudinal and transverse optical modes. In the absence of an external field, the depolarization field is

$$E_d = -\frac{(\mathbf{P} \cdot \mathbf{k})}{\varepsilon_b} \frac{k}{k^2}, \quad (7.4.22)$$

where ε_b is the background dielectric constant and the spontaneous polarization $\mathbf{P}$ is a vector proportional to u with a factor of the Born effective charge Ze . Therefore, the following two terms are added in the energy expression:

$$-PE_d - \frac{1}{2} \epsilon_b E_d^2. \quad (7.4.23)$$

In the following, we gave the dispersion surface along three special k -points.

(1) $k=k[1,0,0]$

The dynamical matrix has the following form after the consideration of the depolarization field:

$$\mathbf{H} = \begin{pmatrix} A + D_{11}k^2 + (Ze)^2 / \epsilon_b & 0 & 0 \\ 0 & A + D_{44}k^2 & 0 \\ 0 & 0 & A + D_{44}k^2 \end{pmatrix}, \quad (7.4.24)$$

From Eq. (7.4.24), the relationship of the longitudinal and transverse modes with the gradient coefficients is:

$$\begin{cases} \omega_L^2 = \frac{1}{M} (A + D_{11}k^2 + (Ze)^2 / \epsilon_b), \\ \omega_T^2 = \frac{1}{M} (A + D_{44}k^2). \end{cases} \quad (7.4.25)$$

(2) $k = \frac{k}{\sqrt{2}} [1,1,0]$

In this case, the depolarization field is $\frac{P_1 + P_2}{2\epsilon_b} [1,1,0]$. And the dynamical matrix is

$$\mathbf{H} = \begin{pmatrix} A + (D_{11} + D_{44})k^2 / 2 + (Ze)^2 / (2\epsilon_b) & (Ze)^2 / (2\epsilon_b) + D_{12}k^2 / 2 & 0 \\ (Ze)^2 / (2\epsilon_b) + D_{12}k^2 / 2 & A + (D_{11} + D_{44})k^2 / 2 + (Ze)^2 / (2\epsilon_b) & 0 \\ 0 & 0 & A + D_{44}k^2 \end{pmatrix}. \quad (7.4.26)$$

And the dispersion relations are:

$$\begin{cases} \omega_L^2 = \frac{1}{M} [A + (D_{11} + D_{12} + D_{44})k^2 / 2 + (Ze)^2 / \epsilon_b], \\ \omega_{T1}^2 = \frac{1}{M} [A + (D_{11} - D_{12} + D_{44})k^2 / 2], \\ \omega_{T2}^2 = \frac{1}{M} (A + D_{44}k^2). \end{cases} \quad (7.4.27)$$

Due to the perturbation of the depolarization field, the former degenerate transverse modes split into two modes.

$$(3) k = \frac{k}{\sqrt{3}} [1, 1, 1]$$

The depolarization field in this case is $\frac{P_1 + P_2 + P_3}{3\epsilon_b} [1, 1, 1]$ and the dynamical matrix is

$$\mathbf{H} = \begin{pmatrix} A + (D_{11} + 2D_{44})k^2 / 3 + (Ze)^2 / (3\epsilon_b) & (Ze)^2 / (3\epsilon_b) + D_{12}k^2 / 3 & (Ze)^2 / (3\epsilon_b) + D_{12}k^2 / 3 \\ (Ze)^2 / (3\epsilon_b) + D_{12}k^2 / 3 & A + (D_{11} + 2D_{44})k^2 / 3 + (Ze)^2 / (3\epsilon_b) & (Ze)^2 / (3\epsilon_b) + D_{12}k^2 / 3 \\ (Ze)^2 / (3\epsilon_b) + D_{12}k^2 / 3 & (Ze)^2 / (3\epsilon_b) + D_{12}k^2 / 3 & A + (D_{11} + 2D_{44})k^2 / 3 + (Ze)^2 / (3\epsilon_b) \end{pmatrix}. \quad (7.4.28)$$

And the dispersion relations are

$$\begin{cases} \omega_L^2 = \frac{1}{M} [A + (D_{11} + 2D_{12} + 2D_{44})k^2 / 3 + (Ze)^2 / \epsilon_b], \\ \omega_T^2 = \frac{1}{M} [A + (D_{11} - D_{12} + 2D_{44})k^2 / 3]. \end{cases} \quad (7.4.29)$$

From the above dispersion surface relationship, one can use the soft phonon dispersion relation measured by inelastic neutron scattering to adjust the parameter of the gradient terms, but it should be noted that the above proposed dispersion surface is derived from the cubic phase; therefore the choice of the soft phonon modes is measured from the cubic phase.

For the cubic BaTiO₃, the neutron-scattering studies of soft modes have been carried out by Harada (1971), and the corresponding dispersion relation of transverse soft-phonon modes with wave vector $\mathbf{k}$ parallel to the [100], [110] and [111] directions was given in Fig. 7.9. Similar results were obtained in the first principles calculation by Ghosez et al. (1999). The phonon dispersion curves at 150 °C can be approximated as

$$\hbar^2 \omega_T^2(100) = [(3 \pm 1) \text{ meV}]^2 + (972 \pm 50) (\text{meV} \cdot \text{\AA})^2 |\mathbf{k}|^2, \quad (7.4.30)$$

$$\hbar^2 \omega_T^2(111) = [(3 \pm 1) \text{ meV}]^2 + (4750 \pm 300) (\text{meV} \cdot \text{\AA})^2 |\mathbf{k}|^2, \quad (7.4.31)$$

$$\hbar^2 \omega_{T1}^2(110) \approx \hbar^2 \omega_T^2(111), \quad \hbar^2 \omega_{T2}^2(110) = \hbar^2 \omega_T^2(100). \quad (7.4.32)$$

Using the data given in Eq. (7.3.24), $A=2.8\times10^7\text{ C}^{-2}\cdot\text{m}^2\cdot\text{N}$ at 150°C , and comparing Eqs. (7.4.25) and (7.4.30), we can obtain the value of $D_{44}=3\times10^{-11}\text{ C}^{-2}\cdot\text{m}^4\cdot\text{N}$, and $\hbar\Omega_p \approx 190\text{ meV}$. The value of the mode-plasma frequency was very consistent with that directly calculated from first-principles calculations (Ghosez et al., 1999). From the comparison of the dispersion in Fig. 7.9 and Eq. (7.4.29), the value of $D_{11}-D_{12}+2D_{44}$ was estimated to be $51\times10^{-11}\text{ C}^{-2}\cdot\text{m}^4\cdot\text{N}$. When assuming that the first neighbor interaction was the major contribution to the gradient energy, we could obtain $D_{12}=-D_{44}$ and $D_{11}=45\times10^{-11}\text{ C}^{-2}\cdot\text{m}^4\cdot\text{N}$. The larger difference between D_{11} and D_{44} was due to the anisotropy of the soft-mode dispersion curves, and it also reflected the anisotropy of the polarization fluctuation in BaTiO₃ ferroelectrics. Due to the uncertainties associated with the domain wall widths of the 90°C and 180°C domain, there were some errors in the determination of D_{11} and D_{44} . And because of the limitation of the available data, we could not give a quantitative estimation of D_{12} by this method. In addition to that, the dependence on temperature and thickness of the gradient term coefficients is also neglected here.

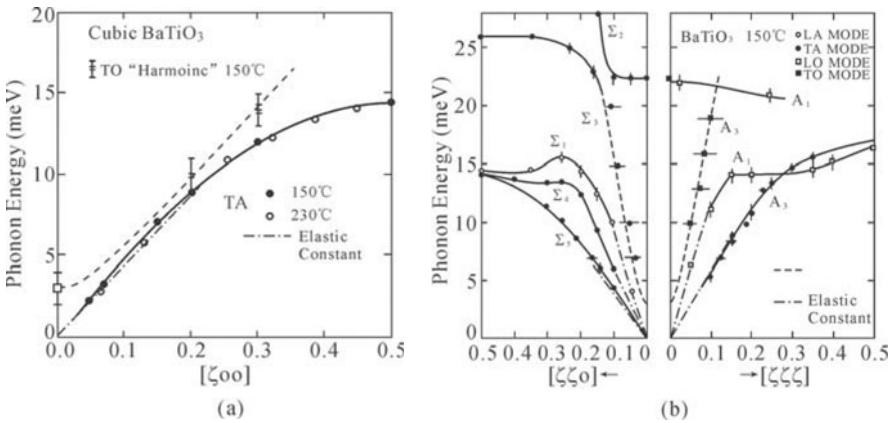


Fig. 7.9 Phonon dispersion curves for the TA and TO soft modes along (a) [100], (b) [110], [111] directions. Reprinted with permission from Harada (1971). Copyright (1971), American Institute of Physics

7.4.4 The Extrapolation Length and the Gradient Coefficient

From the concept of the extrapolation length, one can partially investigate the thickness dependence of the gradient coefficient. For simplicity, we start from a semi-infinite ferroelectrics with nonzero polarization along the c axis, which can also be generalized to the case of freestanding thin films. The free energy of the semi-infinite ferroelectrics with a free surface can be described by Eq. (7.4.2).

Since the gradient term in Eq. (7.4.2) can be converted to a surface and volume integral when the divergence theorem is applied,

$$\iiint \nabla (D(\nabla P)P) dV = \iint D[\hat{\mathbf{n}} \cdot (\nabla P)P] dS, \quad (7.4.33)$$

$$G = \iiint D[\nabla P(r)]^2 dV = \iint D[\hat{\mathbf{n}} \cdot \nabla P(r)]P(r) dS - \iiint D[P(r)\nabla^2 P(r)] dV, \quad (7.4.34)$$

where $\hat{\mathbf{n}}$ is the surface vector along the normal direction of the surface. Analogous to the system of magnetic thin films (Binder and Hohenberg, 1972; Kaganov and Omelyandchouk, 1972), the free energy can be split into several parts:

The interior free energy:

$$G_{\text{int}} = \iiint \left[\frac{A}{2} P^2 + \frac{B}{4} P^4 + \frac{C}{6} P^6 - D(P\nabla^2 P) - PE_{\text{ext}} \right] dV. \quad (7.4.35)$$

The surface contribution:

$$G_{\text{sur}} = \iint \left[D(\hat{\mathbf{n}} \cdot \nabla P)P + \alpha(T - T^*)P^2 / 2 \right] dS, \quad (7.4.36)$$

where T^* is the Curie temperature of the semi-infinite ferroelectrics or thin films, which is not necessarily equivalent to that of the bulk ferroelectrics due to the different boundary conditions. The quartic and sixth order terms are neglected in the surface free energy.

For most cases in reality, there is a mismatch between the films and the substrates. As a result, the effect of the stresses induced from the different expansion coefficients and lattice mismatches must be taken into account. For the biaxial tetragonal phase, we have $\sigma_1 = \sigma_2 = \sigma$ and $\sigma_3 = \sigma_4 = \sigma_5 = \sigma_6 = 0$ in the Voigt notation.

Therefore, the stress contribution to the free energy is given by

$$\begin{aligned} G_{\text{stress}} &= \iiint \left[-\frac{1}{2} s_{11}(\sigma_1^2 + \sigma_2^2) - s_{12}\sigma_1\sigma_2 \right] dV \\ &= \iiint -(s_{11} + s_{12})\sigma^2 dV. \end{aligned} \quad (7.4.37)$$

Besides that, in the ferroelectric phase there is interaction between the stress and the polarization, and the electrostrictive coupling energy is

$$\begin{aligned} G_{\text{coupling}} &= \iiint -Q_{12}(\sigma_1 + \sigma_2)P^2 dV \\ &= \iiint -2Q\sigma P^2 dV. \end{aligned} \quad (7.4.38)$$

At last, we obtain the variation of the free energy with respect to the polarization,

$$\begin{aligned} \frac{\delta G}{\delta P} = & \iiint \left[(A - 4Q\sigma)P + BP^3 + CP^5 - D(\nabla^2 P) - E_{\text{ext}} \right] dV \\ & + \iint \left[D(\hat{n} \cdot \nabla P) + \alpha(T - T^*)P \right] dS. \end{aligned} \quad (7.4.39)$$

To simplify the calculation, we consider only the case where polarization is along the z direction. The stationary condition for the total free energy with respect to the polarization variation leads to $\frac{\delta G}{\delta P} = 0$; thus the polarization satisfies the following nonlinear Euler-Lagrange equation together with the boundary conditions

$$\begin{cases} E_{\text{ext}} = (A - 4Q\sigma)P + BP^3 + CP^5 - D(\nabla^2 P), \\ D \frac{dP}{dz} + \alpha(T - T^*)P = 0 \Big|_{z=0^+}. \end{cases} \quad (7.4.40)$$

At the surface, the boundary condition can be rewritten as

$$\frac{dP}{dz} = -\frac{P}{\lambda} \Big|_{z=0^+}, \quad (7.4.41)$$

with

$$\lambda = \frac{D}{\alpha(T - T^*)} \quad (7.4.42)$$

as the extrapolation length. Eq. (7.4.41) gives the solution $P(z) = P_0 \exp(-z/\lambda)$, where P_0 is the polarization at the surface. As can be seen from Fig. 7.10, if $\lambda > 0$, then the polarization at the surface is smaller than that at the interior, which also means that the Curie temperature of the surface T^* is smaller than that of the bulk T_c . In contrast, when the bulk Curie temperature $T_c < T^*$, the extrapolation length $\lambda < 0$, and the polarization is higher than that of the bulk. This relationship can be generalized to ferroelectric films (Tilley and Zeks, 1984; Ong et al., 2001) with appropriate boundary conditions at the surfaces.

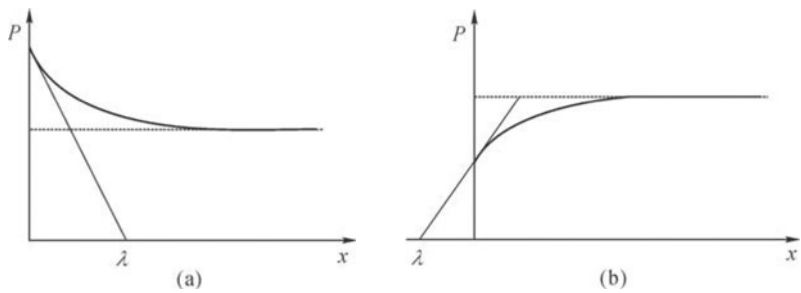


Fig. 7.10 Schematic illustration of the extrapolation length and its relationship with the spatial distribution of spontaneous polarizations. (a) $\lambda < 0$; (b) $\lambda > 0$

From Eq. (7.4.42), if we can measure the extrapolation length and the Curie temperature of the semi-infinite ferroelectrics or ferroelectric thin film, then the gradient coefficient D in the nanoscale ferroelectric systems can be determined. In practice, the extrapolation length can be determined from the surface relaxation measured from the x-ray diffraction patterns. With the help of the x-ray diffraction (XRD), the lattice parameters and surface relaxation PbTiO_3 and BaTiO_3 were measured by Ishikawa (1999), and the value of the extrapolation lengths was as follows:

$$\begin{aligned}\lambda_{\text{PTO}} &= 7 \text{ nm}, \\ \lambda_{\text{BTO}} &= 88 \text{ nm}.\end{aligned}\tag{7.4.43}$$

In the following, we summarize the parameters using the tetragonal thin films for BaTiO_3 , PbTiO_3 and $\text{Pb}(\text{Zr}_{0.5}\text{Ti}_{0.5})\text{O}_3$ materials in Table 7.4. The correlation length ξ is calculated by Eq. (7.4.11).

Table 7.4 The coefficients of Landau-Devonshire-Ginzburg for BaTiO_3 , PbTiO_3 , and PZT (Li et al., 1997; Pertsev et al., 1998; Zhang et al., 2002)

Ferroelectrics	T_c ($^{\circ}\text{C}$)	A ($\text{m}\cdot\text{F}^{-1}$)	B ($\text{m}^5\cdot\text{C}^{-2}\cdot\text{F}$)	C ($\text{m}^9\cdot\text{C}^{-4}\cdot\text{F}$)	D ($\text{m}^3\cdot\text{F}^{-1}$)	Q ($\text{m}^4\cdot\text{C}^{-2}$)	ξ (nm)
BaTiO_3	135	6.6×10^5	$1.44\times 10^7(T-175)$	3.96×10^{10}	2.7×10^{-9}	0.043	6.1
PbTiO_3	490	7.6×10^5	-2.9×10^8	1.56×10^9	2.7×10^{-9}	0.026	2.8
PZT	392	2.7×10^5	-6.0×10^8	6.4×10^8	2.7×10^{-9}	0.026	5.2

7.5 The Transverse Ising Model and Statistical Mechanics Approaches

7.5.1 Phase Transition from the Transverse Ising Model

The transverse Ising model was first introduced to describe the order-disorder phase transition in hydrogen-bonded ferroelectrics, such as KH_2PO_4 and PbHPO_4 (Blin and Zeks, 1974; Stinchcombe, 1973; Elliott and Young, 1974). Then this microscopic model was also applied to other systems successfully (Lage and Stinchcombe, 1979). In this section, we will first introduce the transverse Ising model briefly, and then give the phase transition properties calculated by the transverse Ising model under mean field approximation. Since Landau theory is based on the expansion of the free energy, it is expected that Landau theory should emerge as the continuum limit of the expressions derived from the microscopic transverse Ising model. Therefore, the relationship between the phenomenological Landau theory and the microscopic transverse Ising model can shed some light on the sound physical foundation to the coefficients of the phenomenological Landau theory.

In H bonded ferroelectrics, the order-disorder motion of the H (proton) is chosen as the order parameter. The possibility of a light atom (such as the proton) tunneling between the two equilibrium positions at the double well minimum is large and cannot be neglected, as shown in Fig. 7.11. Since the energy separation between the ground state A and the higher level vibrational state (such as B) is large compared with the splitting of the ground double minimum Δ introduced by the quantum tunneling, it is reasonable to neglect the higher excited energy states.

For the two-level (E_+ and E_-) systems, the Hamiltonian of the system should include the single particle component $H_1(i)$ and the interaction component $H_2(i,j)$:

$$H = \sum_i H_1(i) + \frac{1}{2} \sum_{ij} H_2(i, j). \quad (7.5.1)$$

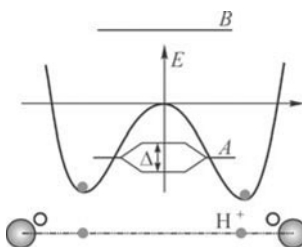


Fig. 7.11 Energy levels in H bonded double well, showing the small tunneling in the lowest level A

After introducing the spin $-\frac{1}{2}$ Pauli matrices and unit matrix,

$$\mathbf{S}^x = \frac{1}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \mathbf{S}^y = \frac{1}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \mathbf{S}^z = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \mathbf{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \quad (7.5.2)$$

We can obtain a complete set of Hermitian 2×2 matrices, and then all the two level systems can be described by a pseudo-spin operator. In the double well situation of the H bonded ferroelectric system, the polarization of each cell can be associated with fictitious spin $-\frac{1}{2}S$. And the Hamiltonian of the pseudo-spin system can be written as:

$$H = -\mathcal{Q} \sum_i S_i^x - \frac{1}{2} \sum_{ij} J_{ij} S_i^z S_j^z. \quad (7.5.3)$$

If we take the $\mathcal{Q}$ as a transverse field, Eq. (7.5.3) is equivalent to the Hamiltonian of the Ising model in a transverse field; thus this model is called the transverse field Ising model. In particular, in H bonded ferroelectrics, $\mathcal{Q}$ is the tunneling probability of the proton through the deep double well. J_{ij} represents the two-spin interaction constants, and S_i^x and S_i^z are the x and z components of the spin at lattice site i .

In the mean-field approximation, the density matrix of the system is equal to the product of single-particle density $\rho = \prod_i \rho_i$, where

$$\rho_i = (1/Z_i) \exp(\beta \mathcal{A}_i \cdot S_i), \quad (7.5.4)$$

with $\beta = 1/(k_B T)$, within Boltzmann statistical mechanics, and the distribution function

$$Z_i = \text{Tr} \exp(\beta \mathcal{A}_i \cdot S_i) = e^{\frac{1}{2}\beta \mathcal{A}_i} + e^{-\frac{1}{2}\beta \mathcal{A}_i} = 2 \cosh\left(\frac{1}{2}\beta \mathcal{A}_i\right). \quad (7.5.5)$$

The entropy S is given by

$$S = -k_B \ln \rho. \quad (7.5.6)$$

Thus, the mean field parameter $\mathcal{A}_i$ can be obtained from the minimum of the free energy $\delta G=0$, here

$$G = U - TS = \text{Tr}(\rho \langle H \rangle) + k_B T \rho \ln \rho. \quad (7.5.7)$$

The molecular mean field is

$$A_i = -\frac{\partial H}{\partial \langle S_i \rangle}, \quad (7.5.8)$$

where $\langle \cdot \rangle$ denotes the average value of the quantities. From Eqs. (7.5.3) – (7.5.8), the effective field for the i th cell is $A_i = (A_x, A_y, A_z)$, and the component of the effective field is

$$A_x = \Omega, \quad A_y = 0, \quad A_z = \sum_j J_{ij} \langle S_j^z \rangle. \quad (7.5.9)$$

With the effective Hamiltonian, the average of the pseudo-spin can be calculated as

$$\begin{aligned} \langle S_i \rangle &= \sqrt{\langle S_i^x \rangle^2 + \langle S_i^y \rangle^2 + \langle S_i^z \rangle^2} \\ &= \frac{\text{Tr} S_i \exp(\beta A_i \cdot S_i)}{\text{Tr} \exp(\beta A_i \cdot S_i)} = \frac{d(\ln Z_i)}{d(\beta A_i)}. \end{aligned} \quad (7.5.10)$$

Combined with Eq. (7.5.10) and $A_i = \left[\Omega^2 + \left(\sum_j J_{ij} \langle S_j^z \rangle \right) \right]^2$,

$$\langle S_i \rangle = \frac{1}{2} A_i^{-1} A_i \tanh \left(\frac{1}{2} \beta A_i \right). \quad (7.5.11)$$

Above the transition temperature T_c , in the paraelectric state $\langle S_i^y \rangle = \langle S_i^z \rangle = 0$ and

$$\langle S_i^x \rangle = \frac{1}{2} \tanh(\Omega / (2k_B T)). \quad (7.5.12)$$

In the ferroelectric phase, $\langle S_i^z \rangle \neq 0$, and the z component of the spin is equivalent for every cell. Thus the spontaneous polarization is

$$P_s = (2\mu / a^3) \langle S^z \rangle, \quad (7.5.13)$$

where $\langle S^z \rangle = \langle S_i^z \rangle = \langle S_j^z \rangle$, a is the unit cell dimension of the simple cubic lattice, and μ is the magnitude of the electric dipole moment of the dipole.

Once $J_0 = \sum_j J_{ij}$ is defined, we can determine the bulk Curie temperature T_c for a second order phase transition from Eq. (7.5.12):

$$\Omega / J_0 = \frac{1}{2} \tanh(\Omega / (2k_B T_c)). \quad (7.5.14)$$

It is found that the Curie temperature is determined by the tunneling frequency and the exchange coupling to its nearest neighbors. For $T < T_c$, the average spin has components in the x and z directions. While for $T > T_c$, only the x component spin exists. As shown in Eq. (7.5.13), the polarization of a ferroelectrics material is proportional to the average of S^z . Eq. (7.5.14) holds only for $2\Omega \leq J_0$; otherwise, the paraelectric phase will always exist at all temperatures.

7.5.2 Relationship of the Parameters Between Landau Theory and the Transverse Ising Model

From Eq. (7.5.13), we can see that the expression can connect the transverse Ising model with that of the Landau theory, if we can express the free energy of the system in a Taylor series expansion of the order parameter $\langle S^z \rangle$. Therefore, Tilley and Zeks (1984) first constructed the free energy of the transverse Ising model from Eq. (7.5.7) with the entropy given by Eq. (7.5.6):

$$G = -\sum_i \Omega \langle S_i^x \rangle - \frac{1}{2} \sum_{ij} J_{ij} \langle S_i^z \rangle \langle S_j^z \rangle + k_B T \sum_i \left[2S_i \operatorname{arctanh}(2S_i) + \frac{1}{2} \ln(1 - 4S_i^2) \right]. \quad (7.5.15)$$

In order to derive the Landau expansion from Eq. (7.5.15), they expanded G to the fourth order in $\langle S_i^z \rangle$ by minimizing G with respect to $\langle S_i^x \rangle$ and $\langle S_i^y \rangle$

$$\frac{\partial G}{\partial \langle S_i^x \rangle} = \frac{\partial G}{\partial \langle S_i^y \rangle} = 0. \quad (7.5.16)$$

As a result, the free energy as a function of $\langle S_i^z \rangle$ was obtained:

$$\begin{aligned}
G = & \sum_i \frac{\Omega_i \langle S_i^z \rangle^2}{\tanh(\Omega_i / (2k_B T))} - \frac{1}{2} \sum_{ij} J_{ij} \langle S_i^z \rangle \langle S_j^z \rangle \\
& + \sum_i \frac{\Omega_i \langle S_i^z \rangle^4}{\tanh^3(\Omega_i / (2k_B T))} \left(1 - \frac{\Omega_i / 2k_B T}{\sinh(\Omega_i / (2k_B T)) \cosh(\Omega_i / (2k_B T))} \right). \quad (7.5.17)
\end{aligned}$$

In the continuum limit, one assumes that $\langle S_i^z \rangle$ changes slowly in space and is a function of the coordinates and the summation is then replaced by the volume integral $\sum_i \rightarrow \frac{1}{a^3} \iiint dx dy dz$. We can express $\langle S^z \rangle_{r+a} = \langle S^z \rangle_r + (a \nabla) \langle S^z \rangle_r$, and the two-spin interaction constant on the surface and interior equals J_s and J , the two-spin interaction constant equals J_{2s} when they are on the surface and J_4 while in the bulk. The tunneling term equals Ω in the bulk and Ω_s on the surface.

With the appearance of the surface, the Landau free energy of semi-infinite ferroelectrics has an expansion of the fourth order as follows:

$$\begin{aligned}
G = & \iint dx dy \int_0^\infty dz \left[\frac{A}{2} P^2 + \frac{B}{4} P^4 + \frac{1}{2} D (\nabla P)^2 - E_{\text{ext}} P \right] \\
& + \frac{1}{2} \iint dx dy D \lambda^{-1} P^2 (z = 0). \quad (7.5.18)
\end{aligned}$$

By using the proportionality relationship between polarization and $\langle S_i^z \rangle$ in Eq. (7.5.13), one obtains the expressions for phenomenological coefficients:

$$\begin{aligned}
A = & \frac{a^3}{4\mu^2} \left(\frac{2\Omega}{\tanh \Omega / (2k_B T)} - 6J \right), \\
B = & \frac{a^9}{16\mu^2} \left(\frac{4\Omega}{\tanh^3(\Omega / (2k_B T))} \left(1 - \frac{\Omega / 2k_B T}{\sinh(\Omega / (2k_B T)) \cosh(\Omega / (2k_B T))} \right) \right), \quad (7.5.19) \\
D = & \frac{a^3}{4\mu^2} J, \\
\lambda = & \frac{2}{a} \left(\frac{5J - 4J_s}{J} + \frac{2}{J} \left(\frac{\Omega_s}{\tanh(1/2)\beta\Omega} - \frac{\Omega}{\tanh(1/2)\beta\Omega} \right) \right).
\end{aligned}$$

The coefficient A equals zero at the bulk transition temperature T_c given by Eq. (7.5.14). It is positive for $T > T_c$ and negative for $T < T_c$. The coefficient B is positive and slightly temperature dependent and can be taken as constant close to T_c . And the gradient coefficient D is proportional to the exchange interaction and the unit cell dimension, and diverges if the electric dipole moment tends to zero. It

is positive and temperature independent as has been assumed in the phenomenological theory. The surface parameter λ depends on the bulk and surface values of the transverse field and the exchange interaction constants.

In the above pseudo-spin system, only the two-body interaction is taken into account. It can explain the basic characteristics of the order-disorder ferroelectrics. But it also has some restrictions, such as it can explain only the second-order phase transition, and there is a big error between the calculated transition temperature and the experimental measurement. Recently, one of the developments of pseudo-spin theory is the inclusion of the four-spin interactions in the transverse Ising model. In that case, the first-order phase transition is explained with good agreement with the experiments.

7.5.3 Determination of Landau-Ginzburg Free Energy Expansion Coefficients from Statistical Mechanics

In this section we start from statistical mechanics to derive the expression of gradient coefficients. We use the interaction potential and some microscopic parameters to derive the coefficients of the expansion, especially the coefficients of the gradient terms. We will focus on the general characteristics. First, let us consider a ferroelectric material element with an average spontaneous polarization distribution $P^0(\mathbf{r})$ under some external electric field $\mathbf{E}$. The total Helmholtz free energy of the system can be written as

$$H = H_0 + \sum_i E_i P_i, \quad (7.5.20)$$

where H_0, E_i, P_i are intrinsic Hamiltonians of the system, external field and the order parameter at the i th cell, respectively.

According to statistical physics, the partition function $\mathbf{Z}$ and the equilibrium probability density f_0 can be written in the form

$$\mathbf{Z} = \text{Tr} \exp(-\beta H) = \int \cdots \int \exp(-\beta H) dP_1 dP_2 \cdots dP_N, \quad (7.5.21)$$

$$f_0 = \frac{1}{\mathbf{Z}} \exp(-\beta H). \quad (7.5.22)$$

It is easy to derive that the Helmholtz free energy is given by

$$F = -\beta^{-1} \log \mathbf{Z}. \quad (7.5.23)$$

Therefore, the average polarization $P^0(\mathbf{r})$ and the correlation function

$C_{ij}(\mathbf{r}, \mathbf{r}')$ is given by

$$P^0 = \langle P_i \rangle = \frac{\delta \Gamma}{\delta E_i} = -\frac{1}{\beta} \frac{\partial \log Z}{\partial E_i},$$

$$\Gamma_{ij}(\mathbf{r}, \mathbf{r}') = \frac{1}{\beta} \frac{\partial P_i^0}{\partial E_j} = \frac{1}{\beta} \frac{\delta^2 F}{\delta E_i \delta E_j} = \langle P_i P_j \rangle - \langle P_i \rangle \langle P_j \rangle = \langle P_i P_j \rangle - P_i^0 P_j^0. \quad (7.5.24)$$

Following Mermin (1965), we introduce the following functions

$$\mathcal{Q}(f) = F'(f) + \sum_{i=1}^N E_i P_i, \quad (7.5.25)$$

$$F'(f) = \text{Tr } f(H + \beta^{-1} \log f),$$

where $F'(f)$ is the intrinsic Helmholtz free energy for $E_i = 0, i = 1, \dots, N$, and $\text{Tr } f = 1$. For the equilibrium probability, we have

$$\mathcal{Q}(f_0) = -\beta^{-1} \log Z = F. \quad (7.5.26)$$

From Eqs. (7.5.21), (7.5.22) and (7.5.26), we can obtain the following relationship $\mathcal{Q}(f) = \mathcal{Q}(f_0) + \beta^{-1}(\text{Tr } f \log f - \text{Tr } f \log f_0)$, where $f \geq f_0$ is the result of the Gibbs inequality (see Eq. (7.2.2)). Therefore, we have the inequality

$$\mathcal{Q}(f) > \mathcal{Q}(f_0), \quad \text{for } f \neq f_0. \quad (7.5.27)$$

Now we consider a ferroelectric system consisting of N crystal lattice, each of which has a spontaneous polarization P_i . Neglecting the kinetic energy, we can write the Hamiltonians in the form

$$H = \sum_{i=1}^N E_i \cdot P_i + \sum_{i=1}^N U(P_i) + \sum_{i=1}^N V_{\text{ext}}, \quad (7.5.28)$$

where the first term is the interaction energy between the external field and lattice polarization, the second term is the local energy for a lattice cell, and the third term is the arbitrary external potential.

For a given U , the equilibrium polarization $P^0(\mathbf{r}) = \langle P(\mathbf{r}) \rangle = \text{Tr } f_0 P(\mathbf{r})$, and the free energy $F'(P) = \text{Tr } f_0(U + \beta^{-1} \log f_0)$ is a unique functional of the polarization. Therefore, the corresponding functional can be written as

$$\Omega(P) = F'(P) + \sum_{i=1}^N E_i P_i. \quad (7.5.29)$$

We suppose $P'(\mathbf{r})$ is the equilibrium polarization associated with another probability density f' , then

$$\begin{aligned} \Omega(f') &= \text{Tr } f'(H + \beta^{-1} \log f' + E_i P) \\ &= F'(P') + \sum_{i=1}^N E_i P'_i \\ &= \Omega(P'), \end{aligned} \quad (7.5.30)$$

where $F'(P') = \text{Tr } f'(U + \beta^{-1} \log f')$. From Eq. (7.5.27) we have

$$\Omega(P') > \Omega(P_0). \quad (7.5.31)$$

From the minimum of functional Ω (Evans, 1979), one can derive the following important results:

$$\begin{aligned} \left. \frac{\delta \Omega}{\delta P_i(\mathbf{r})} \right|_{P_0} = 0 &\Rightarrow \left. \frac{\delta F'}{\delta P_i} \right|_{P_0} = -E_i, \\ \left. \frac{\delta^2 F'}{\delta P_i \delta P_j} \right|_{P_0} &= -\frac{\delta E_i}{\delta P_j^0} = -G(\mathbf{r}_i, \mathbf{r}_j). \end{aligned} \quad (7.5.32)$$

Combining Eqs. (7.5.24) and (7.5.32), one can obtain

$$\int G(\mathbf{r}_i - \mathbf{r}_k) \Gamma(\mathbf{r}_k - \mathbf{r}_j) d\mathbf{r}_k = -\frac{1}{\beta} \delta(\mathbf{r}_i - \mathbf{r}_j). \quad (7.5.33)$$

From the fluctuation-response theorem (Schwabl, 2006), one knows that the susceptibility is related to the correlation function by

$$\chi(\mathbf{r}_l - \mathbf{r}_{l'}) = \beta \Gamma(\mathbf{r}_l - \mathbf{r}_{l'}). \quad (7.5.34)$$

Therefore, $G^T(\boldsymbol{\omega}) = -\frac{1}{\chi^T(\boldsymbol{\omega})}$.

In the following, we will pay attention to the Helmholtz free energy of the system under no external field as follows:

$$F = \iiint_V f(P) d\mathbf{r} . \quad (7.5.35)$$

Following the formal development of gradient expansion used by Hohenberg and Kohn (1964) for the interacting electron gas, and assuming that the polarization density varies slowly over the range of the lattice correlations, the average polarization can be written in the form $P^0(\mathbf{r}) = \Psi(\mathbf{r}/r_0)$, where the scale parameter $r_0 \rightarrow \infty$. For large r_0 , the energy density can be expanded as a series of polarization gradients:

$$\begin{aligned} f[P(\mathbf{r})] = & f_0[P(\mathbf{r})] + \sum_{i=1}^3 A_i(P) \nabla_i P \\ & + \sum_{i,j=1}^3 [B_{ij}(P) \nabla_i P \nabla_j P + C_{ij}(P) \nabla_i \nabla_j P] + \dots \end{aligned} \quad (7.5.36)$$

Successive terms in Eq. (7.5.36) correspond to successive powers of the inverse scale parameters r_0^{-1} . Since $f[P(\mathbf{r})]$ is a unique functional of P independent of the applied electric field, it must be invariant around $\mathbf{r}$. Therefore, the coefficients in Eq. (7.5.36) are functions of only the average polarization and independent of the directions, Eq. (7.5.36) can be expressed in the form :

$$f[P(\mathbf{r})] = f_0[P(\mathbf{r})] + f_2[P(\mathbf{r})] |\nabla P|^2 + \dots \quad (7.5.37)$$

To derive Eq. (7.5.37) we have used

$$f(P) \frac{\partial^2 P}{\partial x_i^2} = \frac{\partial}{\partial x_i} \left[f(P) \frac{\partial P}{\partial x_i} \right] - \frac{\partial f}{\partial P} \left(\frac{\partial P}{\partial x_i} \right)^2 . \quad (7.5.38)$$

And the first term will vanish on integration over the volume of the system. The $f_2[P(\mathbf{r})]$ in Eq. (7.5.37) is the coefficient of the gradient term, and should be determined in the following. In fact, the polarization is a vector with three components. Here we consider the case with only one non-zero component. For more general cases, the extension is straightforward.

We restrict our consideration to the spontaneous polarization distributions which vary slowly and exhibit only a small departure from the mean value P_0 , i.e., $|\tilde{P}| = |P(\mathbf{r}) - P_0| \ll P_0$, and the free energy functional can be expanded in powers of $\tilde{P}$ as follows,

$$\begin{aligned}
F(P) = F(P_0) + \int d\mathbf{r} \frac{\delta F}{\delta P(\mathbf{r})} \Big|_{P_0} \tilde{P}(\mathbf{r}) \\
+ \frac{1}{2} \iint d\mathbf{r} d\mathbf{r}' \frac{\delta^2 F}{\delta P(\mathbf{r}) \delta P(\mathbf{r}')} \Big|_{P_0} \tilde{P}(\mathbf{r}) \tilde{P}(\mathbf{r}') + \dots,
\end{aligned} \tag{7.5.39}$$

where $\Big|_{P_0}$ means the polarization is set to the average value. From Eq. (7.5.32) we know that for uniform distributions of spontaneous polarization, the second derivative of F depends only on $|\mathbf{r} - \mathbf{r}'|$; therefore, by using the Fourier transform form and the theorem of the exchange integral, we can obtain

$$\begin{aligned}
F(P) = F(P_0) + \int d\mathbf{r} \frac{\delta F}{\delta P(\mathbf{r})} \Big|_{P_0} \tilde{P}(\mathbf{r}) \\
+ \frac{1}{2} \left(\frac{1}{2\pi} \right)^3 \int G^T(\boldsymbol{\omega}) \tilde{P}^T(\boldsymbol{\omega}) \tilde{P}^T(-\boldsymbol{\omega}) d\boldsymbol{\omega} + \dots,
\end{aligned} \tag{7.5.40}$$

where $G^T(\boldsymbol{\omega}) = G^T(\boldsymbol{\omega})$ since G is dependent only on $|\mathbf{r} - \mathbf{r}'|$.

For slowly varying polarization distribution, only the small ω Fourier components will be important in the expansion. The function $G(\omega)$ can be expanded in the form

$$G(\omega) = a(P_0) + b(P_0)\omega^2 + c(P_0)\omega^4 + \dots \tag{7.5.41}$$

Substitution of Eq. (7.5.41) into Eq. (7.5.40) yields

$$\begin{aligned}
F(P) = F(P_0) + \int d\mathbf{r} \frac{\delta F}{\delta P(\mathbf{r})} \Big|_{P_0} \tilde{P}(\mathbf{r}) \\
+ \frac{1}{2} a(P_0) \int [\tilde{P}(\mathbf{r})]^2 d\mathbf{r} + \frac{1}{2} b(P_0) \int |\nabla \tilde{P}(\mathbf{r})|^2 d\mathbf{r} + \dots.
\end{aligned} \tag{7.5.42}$$

In order to derive the coefficient of the gradient term, we will expand $f_0(P), f_2(P)$ in Eq. (7.5.37) in the Taylor series about P_0 ,

$$\begin{aligned}
f[P(\mathbf{r})] = f_0(P_0) + f'_0(P_0) \tilde{P}(\mathbf{r}) + \frac{f''_0(P_0)}{2} \tilde{P}(\mathbf{r})^2 + \dots \\
+ [f_2(P_0) + f'_2(P_0) \tilde{P}(\mathbf{r}) + \dots] |\nabla \tilde{P}(\mathbf{r})|^2 + \dots.
\end{aligned} \tag{7.5.43}$$

Then, comparing the coefficient with Eq. (7.5.42), we find

$$\begin{aligned}
f'_0(P_0) &= \frac{\delta F}{\delta P(\mathbf{r})} \Big|_{P_0}, \\
f''_0(P_0) &= a(P_0), \\
f_2(P_0) &= b(P_0).
\end{aligned} \tag{7.5.44}$$

Since we are interested in the gradient coefficients, from Eq. (7.5.44) it is shown that f_2 is proportional to the coefficients of ω^2 in the expansion of the inverse correlation function. According to the result of Yang (1976), we have therefore

$$f_2(P_0) = b(P_0) = D = \iiint r^2 G(r) d\mathbf{r} = \iiint r^4 G(r) \cos \theta dr d\theta d\varphi. \tag{7.5.45}$$

According to the scaling hypothesis of phase transition theory (Schwabl, 2006), the Fourier-transformed susceptibility can be expressed in the form

$$\chi(\omega) = \frac{1}{\omega^{2-\eta}} \tilde{\chi}(\omega\xi), \quad \omega = |\omega|, \quad \xi = \xi_0(T - T_c)^{-\nu}. \tag{7.5.46}$$

Here, the function G can be derived by the inverse Fourier transform as follows:

$$\begin{aligned}
G(\mathbf{r}) &= \left(\frac{1}{2\pi}\right)^3 \iiint G(\omega) e^{i\omega \cdot \mathbf{r}} d\omega = -\left(\frac{1}{2\pi}\right)^3 \iiint 1/\chi(\omega) e^{i\omega \cdot \mathbf{r}} d\omega \\
&= -\left(\frac{1}{2\pi}\right)^3 \iiint \omega^{2-\eta} \tilde{\chi}^{-1}(\omega\xi) e^{i\omega \cdot \mathbf{r}} d\omega.
\end{aligned} \tag{7.5.47}$$

Let $Q = \omega r$, $r = |\mathbf{r}|$ and one obtains

$$\begin{aligned}
G(\mathbf{r}) &= -\left(\frac{1}{2\pi}\right)^3 \iiint \left(\frac{Q}{r}\right)^{2-\eta} \tilde{\chi}^{-1}\left(\frac{Q}{r}\xi\right) e^{iQ \cos \theta} \left(\frac{Q}{r}\right)^2 d\left(\frac{Q}{r}\right) ds \\
&= \frac{1}{r^{5-\eta}} \hat{G}(\xi/r).
\end{aligned} \tag{7.5.48}$$

Substitution of Eq. (7.5.48) into Eq. (7.5.45) yields

$$f_2(P_0) = D = D_0 \xi^\eta = D_0 (|T - T_c|)^{-\nu\eta} = D_0 (|T - T_c|)^{-\lambda}. \tag{7.5.49}$$

The exponents can be determined by the renormalization group approach. For the case with three components of the order parameter and three-dimensional problems, we can express the two critical exponents in the following form as (Schwabl, 2006)

$$\begin{aligned}\nu &= \frac{1}{2} + \frac{n+2}{4(n+8)}\varepsilon + O(\varepsilon^2), \\ \eta &= \frac{n+2}{2(n+8)^2}\varepsilon^2 + O(\varepsilon^3).\end{aligned}\tag{7.5.50}$$

where $\varepsilon = 4 - d$, d and n are the dimension and number of components of the order parameters, respectively. For three-dimensional ferroelectric transition problems with three components of the order parameter, one can obtain

$$\nu = 0.6114, \quad \eta = 0.021, \quad \lambda = 0.01284.\tag{7.5.51}$$

Under the molecular field approximation (Schwabl, 2006), $\eta = 0$. Therefore, no divergent properties could be found as shown in (Cottam et al., 1984; Wang et al., 1996) when the temperature approaches its critical level. From our work, we have predicted that the gradient coefficient diverges at the critical point with a universal critical exponent $\lambda = 0.01284$ (Fig. 7.12), indicating that the spatial non-uniform polarization is greatly enhanced during the transition temperatures.

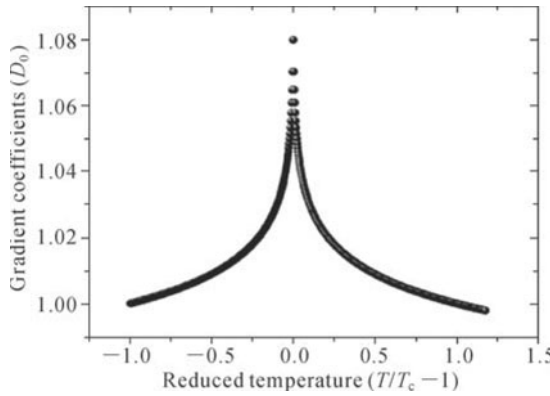


Fig. 7.12 The deduced gradient coefficients as a function of the reduced temperatures

7.6 Remarks

In summary, modeling ferroelectrics in finite dimensions using the Landau-Devonshire-Ginzburg free-energy functional approach requires the contributions from the spatial inhomogeneous distribution of the order parameter, which can be caused by effects of the surface, interfacial and other finite dimensions systems. Based on the more rigorously derived free energy of ferroelectrics by using spontaneous polarization as the order parameter, we comprehensively discussed

the derivation of the free energy expansion coefficients in BaTiO_3 and PbTiO_3 . The properties of these ferroelectrics, such as the phase transition temperature, dielectric constants and Curie-Weiss-type relation, are illustrated in ferroelectric systems.

When considering the inhomogeneous distribution of the order parameter, we introduce the gradient energy in the free energy expression, and introduce how to obtain the coefficients of the gradient term from the lattice dynamics and transverse Ising model. Based on statistical mechanics, we give a straightforward model to account for the physical meaning of the gradient terms, and give the relationship of the coefficient to the dielectric constants. Nevertheless, despite the more reasonable thermodynamic model of ferroelectrics in gradient energy advanced in this book, we consider that there are still many other outstanding challenges to the thermodynamic models based on the Landau-type theory. A comprehensive treatment of the surface, interface, background dielectric properties, polarization fluctuations, etc. of ferroelectrics still requires a better understanding of materials behavior, which may require the combined efforts of phenomenological theory, molecular dynamics, first-principle-based calculations, etc.

References

- [1] Amin, A., Haun, M.J., Badger, B., McKinstry, H., Cross, L.E.: A phenomenological Gibbs function for the single cell region of the PbZrO_3 : PbTiO_3 solid solution system. *Ferroelectrics* **65**, 107-130 (1985)
- [2] Anderson, P.W.: *Fizika Dielektrikov*. Akad Nank, Moscow (1960)
- [3] Axe, J.D.: Apparent ionic charges and vibrational eigenmodes of BaTiO_3 and other perovskites. *Phys. Rev.* **157**, 429-435 (1967)
- [4] Bachelet, G.B., Hamann, D.R., Schlüter, M.: Pseudopotentials that work: from H to Pu. *Phys. Rev. B* **26**, 4199-4228 (1982)
- [5] Bader, R.F.W.: *Atoms in Molecules: a Quantum Theory*. Clarendon Press, Oxford, UK (1990)
- [6] Bader, R.F.W., Nguyen-Dang, T.T., Tal, Y.: A topological theory of molecular structure. *Rep. Prog. Phys.* **44**, 893-948 (1981)
- [7] Baldereschi, A., Resta, R.: *Ab initio Calculation of Phonon Spectra*. Plenum Press, New York (1983)
- [8] Baroni, S., Giannozzi, P., Testa, A.: Green's-function approach to linear response in solids. *Phys. Rev. Lett.* **58**, 1861-1864 (1987)
- [9] Bell, A.J., Cross, L.E.: A phenomenological Gibbs function for BaTiO_3 giving correct e field dependence of all ferroelectric phase changes. *Ferroelectrics* **59**, 197-203 (1984)
- [10] Binder, K., Hohenberg, P.C.: Phase transitions and static spin correlations in Ising models with free surfaces. *Phys. Rev. B* **6**, 3461-3487 (1972)
- [11] Blinc, R., Zeks, B.: *Soft Modes in Ferroelectrics and Antiferroelectrics*.

- North-Holland Amsterdam (1974)
- [12] Cao, W.: Polarization gradient coefficients and the dispersion surface of the soft mode in perovskite ferroelectrics. *J. Phys. Soc. Jpn.* **63**, 1156-1161 (1994)
 - [13] Choi, K.J., Biegalski, M., Li, Y.L., Sharan, A., Schubert, J., Uecker, R., Reiche, P., Chen, Y.B., Pan, X.Q., Gopalan, V., Chen, L.Q., Schlom, D.G., Eom, C.B.: Enhancement of ferroelectricity in strained BaTiO₃ thin films. *Science* **306**, 1005-1009 (2004)
 - [14] Cottam, M.G., Tilley, D.R., Zeks, B.: Theory of surface modes in ferroelectrics. *J. Phys. C Solid. State. Phys.* **17**, 1793-1823 (1984)
 - [15] Dawber, M., Lichtensteiger, C., Cantoni, M., Veithen, M., Ghosez, P., Johnston, K., Rabe, K.M., Triscone, J.M.: Unusual behavior of the ferroelectric polarization in PbTiO₃/SrTiO₃ superlattices. *Phys. Rev. Lett.* **95**, 177601 (2005)
 - [16] Dawber, M., Rabe, K.M., Scott, J.F.: Physics of thin-film ferroelectric oxides. *Rev. Mod. Phys.* **77**, 1083-1130 (2005)
 - [17] Dawber, M., Stucki, N., Lichtensteiger, C., Gariglio, S., Ghosez, P., Triscone, J.M.: Tailoring the properties of artificially layered ferroelectric superlattices. *Adv. Mater.* **19**, 4153-4159 (2007)
 - [18] Deguchi, K., Nakamura, E.: Critical region in ferroelectric triglycine sulfate. *Phys. Rev. B* **5**, 1072-1073 (1972)
 - [19] Devonshire, A.F.: Theory of barium titanate-part I. *Philos. Mag.* **40**, 1040-1063 (1949)
 - [20] Devonshire, A.F.: Theory of barium titanate-part II. *Philos. Mag.* **42**, 1065 (1951)
 - [21] Devonshire, A.F.: Theory of ferroelectrics. *Adv. Phys.* **3**, 85-130 (1954)
 - [22] Diéguez, O., Tinte, S., Antons, A., Bungaro, C., Neaton, J.B., Rabe, K.M., Vanderbilt, D.: *Ab initio* study of the phase diagram of epitaxial BaTiO₃. *Phys. Rev. B* **69**, 212101 (2004)
 - [23] Ehrhart, P., Landolt-Börnstein: Numerical Data and Functional Relationships in Science and Technology. Springer, Heidelberg (1991)
 - [24] Elliott, R.J., Young, A.Y.: The pseudo-spin model and its application to KDP-type materials. *Ferroelectrics* **7**, 23-30 (1974)
 - [25] Evans, R.: The nature of the liquid-vapor interface and other topics in the statistical mechanics of non-uniform, classical fluids. *Adv. Phys.* **28**, 143-200 (1979)
 - [26] Fong, D.D., Kolpak, A.M., Eastman, J.A., Streiffer, S.K., Fuoss, P.H., Stephenson, G.B., Thompson, C., Kim, D.M., Choi, K.J., Eom, C.B.: Stabilization of monodomain polarization in ultrathin PbTiO₃ Films. *Phys. Rev. Lett.* **96**, 127601 (2006)
 - [27] Gao, Y.F., Suo, Z.: Domain dynamics in a ferroelastic epilayer on a paraelastic substrate. *J. Appl. Mech.* **69**, 419-424 (2002)
 - [28] Gavril'yachenko, V.G., Spinko, R.I., Martynenko, M.A., Fesenko, E.G.: Spontaneous polarization and coercive field of lead titanate. *Sov. Phys. Solid State* **12**, 1203 (1970)
 - [29] Ghosez, P., Cockayne, E., Waghmare, U.V., Rabe, K.M.: Lattice dynamics

- of BaTiO₃, PbTiO₃, and PbZrO₃: A comparative first-principles study. *Phys. Rev. B* **60**, 836-843 (1999)
- [30] Ghosez, P., Gonze, X., Charlier, J.C., Godby, R.W., Resta, R., Vigneron, J.P., Issi, J.P.: First-Principles Study of the Dielectric and Dynamical Properties of Barium Titanate. Doctor Thesis, Universite Catholique de Louvain (1997)
 - [31] Ginzburg, V.L.: The dielectric properties of crystals of seignetteelectric substances and of barium titanate. *Zh. Eksp. Teor. Fiz* **15**, 739-749 (1945)
 - [32] Ginzburg, V.L.: *ZhETF* **15**, 739 (1945). *Journ. Phys. USSR* **10**, 107 (1946)
 - [33] Ginzburg, V.L.: Some remarks on second order phase transitions and microscopic theory of ferroelectrics. *Fiz. Tverd. Tela* **2**, 111 (1960)
 - [34] Harada, J., Axe, J.D., Shirane, G.: Neutron-scattering study of soft modes in cubic BaTiO₃. *Phys. Rev. B* **4**, 155-162 (1971)
 - [35] Harwood, M.G., Popper, P., Rushman, D.F.: Curie point of barium titanate. *Nature* **160**, 58-59 (1947)
 - [36] Haun, M.J., Furman, E., Jang, S.J., Cross, L.E.: Thermodynamic theory of the lead zirconate-titanate solid solution system, part V: Theoretical calculations. *Ferroelectrics* **99**, 63-86 (1989)
 - [37] Haun, M.J., Furman, E., Jang, S.J., McKinstry, H.A., Cross, L.E.: Thermodynamic theory of PbTiO₃. *J. Appl. Phys.* **62**, 3331-3338 (1987)
 - [38] Haun, M.J., Furman, E., McKinstry, H.A., Cross, L.E.: Thermodynamic theory of the lead zirconate-titanate solid solution system, part II: Tricritical behavior. *Ferroelectrics* **99**, 27-44 (1989)
 - [39] Haun, M.J., Zhuang, Z.Q., Furman, E., Jang, S.J., Cross, L.E.: Thermodynamic theory of the lead zirconate-titanate solid solution system, part III: Curie constant and sixth-order polarization interaction dielectric stiffness coefficients. *Ferroelectrics* **99**, 45-54 (1989)
 - [40] Hlinka, J., Marton, P.: Phenomenological model of a 90° domain wall in BaTiO₃ type ferroelectrics. *Phys. Rev. B* **74**, 104104 (2006)
 - [41] Hlinka, J., Petzelt, J., Kamba, S., Noujni, D., Ostapchuk, T.: Infrared dielectric response of relaxor ferroelectrics. *Phase. Transit.* **79**, 41-78 (2006)
 - [42] Hohenberg, P., Kohn, W.: Inhomogeneous electron gas. *Phys. Rev.* **136**, B864-B871 (1964)
 - [43] Hu, H.L., Chen, L.Q.: Three-dimensional computer simulation of ferroelectric domain formation. *J. Am. Ceram. Soc.* **81**, 492-500 (1998)
 - [44] Huibregtse, E.J., Young, D.R.: Triple hysteresis loops and the free-energy function in the vicinity of the 5 °C transition in BaTiO₃. *Phys. Rev.* **103**, 1705-1711 (1956)
 - [45] Iniguez, J., Ivantchev, S., Perez-Mato, J.M., Garca, A.: Devonshire-Landau free energy of BaTiO₃ from first principles. *Phys. Rev. B* **63**, 144103 (2001)
 - [46] Ishikawa, K., Uemori, T.: Surface relaxation in ferroelectric perovskites. *Phys. Rev. B* **60**, 11841-11845 (1999)
 - [47] Junquera, J., Ghosez, P.: Critical thickness for ferroelectricity in perovskite ultrathin films. *Nature* **422**, 506-509 (2003)
 - [48] Kaganov, M.I., Omel'Yanchuk, A.N.: Phenomenological theory of phase transition in a thin ferromagnetic plate. *Sov. J. Exp. Theo. Phys.* **34**, 895

- (1972)
- [49] Kay, H.F., Wellard, H.J., Vousden, P.: Atomic positions and optical properties of barium titanate. *Nature* **163**, 636-637 (1949)
 - [50] Kittel, C.: Theory of antiferroelectric crystals. *Phys. Rev.* **82**, 729-732 (1951)
 - [51] Kittel, C.: Introduction to Solid State Physics. John Wiley & Sons, New York (2005)
 - [52] Kretschmer, R., Binder, K.: Surface effects on phase transitions in ferroelectrics and dipolar magnets. *Phys. Rev. B* **20**, 1065-1076 (1979)
 - [53] Lage, E.J.S., Stinchcombe, R.B.: The bond-diluted problem in the transverse Ising model. *J. Phys. C Solid. State. Phys.* **12**, 1319-1329 (1979)
 - [54] Landau, L.D., Lifshitz, E.M.: Electrodynamics of Continuous Media. Pergamon Press, Oxford (1960)
 - [55] Levanyuk, A.P.: *Sov. Phys. JETP* **36**, 571 (1959)
 - [56] Li, S., Eastman, J.A., Vetrone, J.M., Foster, C.M., Newnham, R.E., Cross, L.E.: Dimension and size effects in ferroelectrics. *Jpn. J. Appl. Phys.* **36**, 5169-5174 (1997)
 - [57] Li, Y.L., Cross, L.E., Chen, L.Q.: A phenomenological thermodynamic potential for BaTiO₃ single crystals. *J. Appl. Phys.* **98**, 064101-064104 (2005)
 - [58] Lines, M.E., Glass, A.M.: Principles and Applications of Ferroelectrics and Related Materials. Oxford University Press, New York (1977)
 - [59] Lupascu, D.C.: Fatigue in Ferroelectric Ceramics and Related Issues. Springer Verlag, Berlin (2004)
 - [60] Mason, W.P.: Theory of the ferroelectric effect and clamped dielectric constant of Rochelle salt. *Phys. Rev.* **72**, 854-865 (1947)
 - [61] Matthias, B.T., Miller, C.E., Remeika, J.P.: Ferroelectricity of glycine sulfate. *Phys. Rev.* **104**, 849-850 (1956)
 - [62] Mehta, R.R., Silverman, B.D., Jacobs, J.T.: Depolarization fields in thin ferroelectric films. *J. Appl. Phys.* **44**, 3379-3385 (1973)
 - [63] Mermin, N.D.: Thermal properties of the inhomogeneous electron gas. *Phys. Rev.* **137**, A1441-A1443 (1965)
 - [64] Merz, W.J.: Double hysteresis loop of BaTiO₃ at the Curie point. *Phys. Rev.* **91**, 513-517 (1953)
 - [65] Mueller, H.: Properties of Rochelle salt. *Phys. Rev.* **57**, 829-839 (1940a)
 - [66] Mueller, H.: Properties of Rochelle salt. III. *Phys. Rev.* **58**, 565-573 (1940b)
 - [67] Mueller, H.: Properties of Rochelle salt. IV. *Phys. Rev.* **58**, 805-811 (1940c)
 - [68] Nambu, S., Sagala, D.A.: Domain formation and elastic long-range interaction in ferroelectric perovskites. *Phys. Rev. B* **50**, 5838-5847 (1994)
 - [69] Nicolis, G., Prigogine, I.: Self-organization in Nonequilibrium Systems. Wiley, New York (1977)
 - [70] Noheda, B., Cox, D.E., Shirane, G., Gonzalo, J.A., Cross, L.E., Park, S.E.: A monoclinic ferroelectric phase in the Pb(Zr_{1-x}Ti_x)O₃ solid solution. *Appl. Phys. Lett.* **74**, 2059-2061 (1999)
 - [71] Ong, L.H., Osman, J., Tilley, D.R.: Landau theory of second-order phase transitions in ferroelectric films. *Phys. Rev. B* **63**, 144109 (2001)
 - [72] Pertsev, N.A., Kohlstedt, H.: Elastic stabilization of a single-domain

- ferroelectric state in nanoscale capacitors and tunnel junctions. *Phys. Rev. Lett.* **98**, 257603 (2007)
- [73] Pertsev, N.A., Zembilgotov, A.G., Tagantsev, A.K.: Effect of mechanical boundary conditions on phase diagrams of epitaxial ferroelectric thin films. *Phys. Rev. Lett.* **80**, 1988-1991 (1998)
- [74] Pintilie, L., Boldyreva, K., Alexe, M., Hesse, D.: Capacitance tuning in antiferroelectric-ferroelectric PbZrO_3 - $\text{Pb}(\text{Zr}_{0.8}\text{Ti}_{0.2})\text{O}_3$ epitaxial multilayers. *New J. of Phys.* **10**, 013003 (2008)
- [75] Rabe, K.A., Ahn, C.H., Triscone, J.M.: *Physics of Ferroelectrics: A Modern Perspective*. Springer, Berlin, Heidelberg, New York (2007)
- [76] Radescu, S., Etxebarria, I., Perez-Mato, J.M.: The Landau free energy of the three-dimensional Φ^4 model in wide temperature intervals. *J. Phys. Condens. Matter.* **7**, 585-595 (1995)
- [77] Remeika, J.P., Glass, A.M.: The growth and ferroelectric properties of high resistivity single crystals of lead titanate. *Mater. Res. Bull.* **5**, 37-45 (1970)
- [78] Rupprecht, G., Bell, R.O.: Dielectric constant in paraelectric perovskites. *Phys. Rev.* **135**, A748-A752 (1964)
- [79] Schwabl, F.: *Statistical Mechanics*. Springer-Verlag, Berlin Heidelberg (2006)
- [80] Shvartsman, V.V., Kholkin, A.L.: Domain structure of $0.8\text{Pb}(\text{Mg}_{1/3}\text{Nb}_{2/3})\text{O}_3$ - 0.2PbTiO_3 studied by piezoresponse force microscopy. *Phys. Rev. B* **69**, 14102 (2004)
- [81] Smoluchowski, R., Mayer, J.E., Weyl, W.A.: *Phase Transformations in Solids*. Wiley, New York (1951)
- [82] Stinchcombe, R.B.: Ising model in a transverse field. I. Basic theory. *J. Phys. C Solid. State. Phys.* **6**, 2459-2483 (1973)
- [83] Streiffner, S.K., Eastman, J.A., Fong, D.D., Thompson, C., Munkholm, A., Ramana Murty, M.V., Auciello, O., Bai, G.R., Stephenson, G.B.: Observation of nanoscale 180° stripe domains in ferroelectric PbTiO_3 thin films. *Phys. Rev. Lett.* **89**, 67601 (2002)
- [84] Tagantsev, A.K.: Landau expansion for ferroelectrics: Which variable to use? *Ferroelectrics* **375**, 19-27 (2008)
- [85] Tenne, D.A., Bruchhausen, A., Lanzillotti-Kimura, N.D., Fainstein, A., Katiyar, R.S., Cantarero, A., Soukiassian, A., Vaithyanathan, V., Haeni, J.H., Tian, W., Schlom, D.G., Choi, K.J., Kim, D.M., Eom, C.B., Sun, H.P., Pan, X.Q., Li, Y.L., Chen, L.Q., Jia, Q.X., Nakhmanson, S.M., Rabe, K.: Probing nanoscale ferroelectricity by ultraviolet raman spectroscopy. *Science* **313**, 1614-1616 (2006)
- [86] Tilley, D.R.: Landau theory of phase transitions in thick films. *Solid State Commun.* **49**, 823-828 (1984)
- [87] Troster, A., Dellago, C., Schranz, W.: Free energies of the Φ^4 model from Wang-Landau simulations. *Phys. Rev. B* **72**, 94103 (2005)
- [88] Tybell, T., Ahn, C.H., Triscone, J.M.: Ferroelectricity in thin perovskite films. *Appl. Phys. Lett.* **75**, 856-858 (1999)
- [89] Vanderbilt, D., Cohen, M.H.: Monoclinic and triclinic phases in higher-order

- Devonshire theory. *Phys. Rev. B* **63**, 94108 (2001)
- [90] Wang, B., Woo, C.H.: Curie temperature and critical thickness of ferroelectric thin films. *J. Appl. Phys.* **97**, 084109-084110 (2005)
 - [91] Wang, B., Woo, C.H.: Curie-Weiss law in thin-film ferroelectrics. *J. Appl. Phys.* **100**, 044114-044115 (2006)
 - [92] Wang, Y.G., Zhong, W.L., Zhang, P.L.: Lateral size effects on cells in ferroelectric films. *Phys. Rev. B* **51**, 17235-17238 (1995)
 - [93] Wang, Y.G., Zhong, W.L., Zhang, P.L.: Surface effects and size effects on ferroelectrics with a first-order phase transition. *Phys. Rev. B* **53**, 11439-11443 (1996)
 - [94] Wang, Y.L., Tagantsev, A.K., Damjanovic, D., Setter, N., Yarmarkin, V.K., Sokolov, A.I.: Anharmonicity of BaTiO₃ single crystals. *Phys. Rev. B* **73**, 132103 (2006)
 - [95] Woo, C.H., Zheng, Y.: Depolarization in modeling nano-scale ferroelectrics using the Landau free energy functional. *Appl. Phys. Mater. Sci. Process.* **91**, 59-63 (2008)
 - [96] Wul, B.M.: *J. Phys. U.S.S.R.* **10**, 95 (1946)
 - [97] Wul, B.M., Goldman, I.M.: Dielectric constants of titanates of metals of the second group (1945).
 - [98] Yang, A.J.M., Paul, D.F., Julian, H.G. III: Molecular theory of surface tension. *J. Chem. Phys.* **64**, 3732-3747 (1976)
 - [99] Zembilgotov, A.G., Pertsev, N.A., Kohlstedt, H., Waser, R.: Ultrathin epitaxial ferroelectric films grown on compressive substrates: Competition between the surface and strain effects. *J. Appl. Phys.* **91**, 2247-2254 (2002)
 - [100] Zhang, J., Wu, Z., Yin, Z., Zhang, M.S.: Stress effect on critical thickness in ferroelectric thin films. *Integr. Ferroelectric.* **43**, 19-30 (2002)
 - [101] Zheng, Y., Woo, C.H.: Tuning the properties of a vertically composition-graded ferroelectric film using a mechanical load. *J. Phys. Appl. Phys.* **41**, 175403 (2008)
 - [102] Zheng, Y., Woo, C.H.: Thermodynamic modeling of critical properties of ferroelectric superlattices in nano-scale. *Appl. Phys. Mater. Sci. Process.* **97**, 617-626 (2009)
 - [103] Zhong, W.L., Qu, B.D., Zhang, P.L., Wang, Y.G.: Thickness dependence of the dielectric susceptibility of ferroelectric thin films. *Phys. Rev. B* **50**, 12375-12380 (1994)
 - [104] Zhong, W.L., Wang, Y.G., Zhang, P.L., Qu, B.D.: Phenomenological study of the size effect on phase transitions in ferroelectric particles. *Phys. Rev. B* **50**, 698-703 (1994)

Multiferroic Materials

8.1 Background

Multiferroic materials are characterized by the coexistence of two or more properties of ferroelectricity, ferromagnetism and ferroelasticity. The coexistence of at least two ferroic forms of ordering leads to additional interactions (Fig. 8.1). A typical example is that of magnetoelectric multiferroic materials, in which an applied electric field can tune the ferromagnetic properties of the materials, and vice versa. The phenomenon is different from what is established in modern electromagnetic theory which shows an intrinsic connection between dynamic electric and magnetic fields. In multiferroic materials, physical parameters like spin, charge, dipole and/or orbital ordering occur simultaneously. Generally speaking, the corresponding states to the physical parameters include one or more of the most important states: ferroelectric, ferrotoroidic, paramagnetic, ferromagnetic, antiferromagnetic, ferrimagnetic or weakly ferromagnetic, ferroelastic, etc. Such couplings present a fascinating playground to explore the interactions and the multi-control of the charge, spin, orbital and lattice degrees of freedom, which eventually leads to new state of matter and is envisioned in a wide range of applications including electrically controlled microwave phase shifters or ferromagnetic resonance devices, magnetically controlled electro-optic or piezoelectric devices, broadband magnetic field sensors, spintronics, millimeterwave devices, miniature antennas and magnetoelectric memory devices, etc. Additionally, the coexistence of the optic-electric and optic-magnetic properties also opened vistas into novel physical phenomena and a wide range of applications.

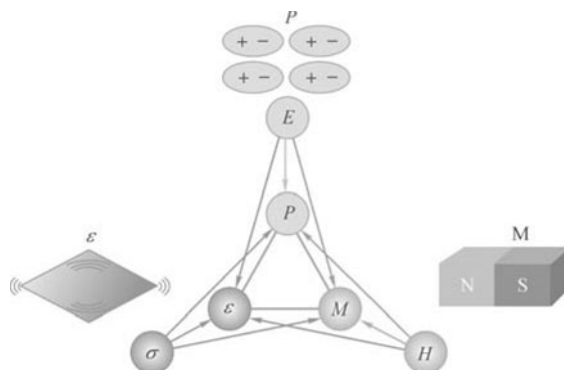


Fig. 8.1 Schematic diagram of the couplings (Fiebig, 2005)

Although the references in the field involve terms such as ‘multiferroic’ and ‘magnetoelectric’, the overlap between them is incomplete, as shown in Fig. 8.2 (Eerenstein et al., 2006). By the original definition, a single-phase multiferroic material is the one that possesses two or all three of the so called ‘ferroic’ properties; i.e., the magnetic and electric order parameters are coupled in one phase and can manipulate one through changing the other. While magnetoelectric coupling may exist in more broad material systems, especially in composites in which magnetoelectric coupling usually arises indirectly from interfaces via strain, linear coupling relations were usually used to describe the induction of magnetization by an electric field or polarization by a magnetic field (Martin et al., 2010). Recently, the research on multiferroics (or magnetic ferroelectrics) was also intimately interwoven with research on the magnetoelectric effect including linear, quadratic, and higher order magneto-electric effects. In a more general concept, forms of cross-correlation between magnetic and electric order parameters can be named ‘magneto-electric effect’, and the corresponding materials can be called ‘multiferroic materials’.

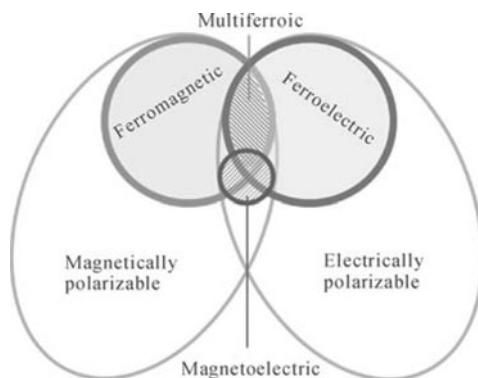


Fig. 8.2 Relationship between multiferroic and magnetoelectric materials (Eerenstein et al., 2006). Copyright (2006) by the Nature Publishing Group

It is a well-known fact that electrons possess charges and spins. Since the charge can be operated by the electric field and the spin by the magnetic field, the coupling effect brings additional degrees and becomes the key to much complex physical behaviors. However, there are some restrictions mainly for single materials, such as symmetry, electronic properties and chemical properties. There are only 13 point groups of the crystalline materials that can give rise to multiferroic behavior. Additionally, ferroelectrics by definition are insulators; for example, 3d transition metal based oxides typically possess ions that have a formal d_0 electronic state, while itinerant ferromagnets need conduction electrons; even in double exchange ferromagnets such as manganites, magnetism is mediated by incompletely filled 3d shells. Thus the contradiction makes the scarcity of single phase multiferroics (Martin et al., 2008). Fortunately, there are still some systems in other kinds of mechanisms. Research mainly focuses on the boracites ($\text{Ni}_3\text{B}_7\text{O}_{13}\text{I}$, $\text{Cr}_3\text{B}_7\text{O}_{13}\text{Cl}$) (Ascher et al., 1966), fluorides (BaMF_4 , $M=\text{Mn, Fe, Co, Ni}$) (Fox et al., 1980), magnetite Fe_3O_4 (Kato et al., 1983), hexagonal manganites ($(\text{Y/Yb})\text{MnO}_3$, and perovskites PbVO_3 , LuFe_2O_4 , $\text{R}_{1-x}\text{Ca}_x\text{MnO}_3$, and Bi-based compounds, namely BiMnO_3 or BiFeO_3 (Wang et al., 2009).

Because of the scarcity of single phase multiferroic materials and the weak coupling effects which usually exist at low temperatures, an alternative route is to make multiferroic composites composed of piezoelectric and magnetostrictive materials. It is the only way by far to achieve technological commercial availability. Such magnetoelectric composites are usually composed of the piezoelectric phase (such as $(\text{PbZr})\text{TiO}_3$ (PZT), BaTiO_3 (BT), PMN-PT) and the magnetostrictive phase (such as LSMO, CoFe_2O_4 , Terfenol-D, NiFe_2O_4). The elastic strain induced by either electric or magnetic field plays a very important role in such systems. Nowadays, more research focuses on the multiferroic nanocomposites which provide the possibility for many interesting nanodevices (Wu et al., 2006).

In fact, the birth of the magnetoelectric (ME) effect and multiferroics can be traced back to 1888, when Röntgen (1888) discovered that a moving dielectric became magnetized when placed in an electric field, and tens years later the reverse effect, polarization of a moving dielectric in a magnetic field, was observed (Wilson, 1905). The ME effect was first theoretically envisaged by Pierre Curie who pointed out the possibility of intrinsic ME behavior of (non-moving) crystals on the basis of symmetry considerations in 1894 (Curie, 1894). In the following years, many researchers were focused on such an effect. One milestone is that Astrov (1960) first observed the small electric-field-induced magnetization and a magnetic field-induced polarization in Cr_2O_3 , both linearly changed with the applied field. Later on, van Suchtelen (1972) reported the successful fabrication of $\text{BaTiO}_3\text{-CoFe}_2\text{O}_4$ composite, whose ME coupling coefficient was two orders higher than that of Cr_2O_3 , but the material still obtained less attention because of the limitation of such a small ME effect for applications. Early works on the ME effect are summarized in Table 8.1.

Table 8.1 Works on ME effect

References	Descriptions
Röntgen, 1888	A moving dielectric became magnetized when placed in an electric field
Curie, 1894	First proposal of the effect on symmetry grounds
Perrier & Staring, 1922	Proposed the possibility of the effect and reported their first experimental results on iron
Perrier & Staring, 1923	Claimed to have definitely observed the effect in iron
Piccard, 1924	Suggested the impossibility of the effect on symmetry grounds
Szivessy, 1925	Expected to find the effect in liquid crystals but no result
Perrier & Borel, 1925	Found no effect in nickel and suggested that their experimental results in 1923 on iron were wrong
Debye, 1926	Suggested the effect was impossible and gave the word 'Magnetoelectric'.
Huber, 1926	Studied NO, FeCl ₃ etc. and found no such effect
van Vleck, 1932	Devoted a section of his book to the reasons why no ME effect could exist
Condon, 1937	Gave the first quantum mechanical description of a system which might show the effect
Telegen, 1948	Proposed the concept of magnetoelectric composite
Landau & Lifshitz, 1957	Showed that the ME effect should exist in magnetic crystals
Smolensky & Ioffe, 1958	Introduced a magnetic ion to B-site of the perovskite ABO ₃ and fabricated Pb(Fe ₁₅ Nb ₁₅)O ₃
Dzyaloshinskii, 1959	Showed that AFM chromium oxide had a magnetic symmetry which showed the effect
Astrov, 1960; Folen, 1961; Rado, 1964	Successfully observed the ME effect in a single crystal of chromium oxide
Brown, 1968	Obtained up-limit value using the method of perturbation theory of quantum mechanics
van Suchtelen, 1972	Proposed the ME effect in composites
van den Boomgard, 1974	Fabricated BaTiO ₃ /CoFe ₂ O ₄ which showed the ME effect
Hans Schmid, 1994	Gave the name 'multiferroic'

From the end of the 1990s, great progress has been achieved both in experimental and theoretical studies. One aspect of fundamental interest to the study of multiferroics is the production of high quality samples of the materials for potential novel devices. With the improvement in the growth and characterization technology of thin film, multiferroic structures (both single-phase and composite) with good performance revealed rich and fascinating phenomena, and became a hot topic in multifunctional materials. In December 2007, *Science* Magazine listed multiferroic materials as one out of ten 'Areas to watch in 2008', the only entry from the materials science/condensed matter area that was included in this list.

The developments of new thin film growth techniques that allow for the production of non-equilibrium phases of materials and strain engineering of existing materials represent a significant step forward in the study of functional complex oxide materials. The growth techniques include sputtering, spin coating, pulsed laser deposition (PLD), sol-gel processes, metal-organic chemical vapor deposition (MOCVD), molecular beam epitaxy (MBE), and more. Up to now, more than 100 compounds that exhibit the magnetoelectric effect have been discovered or synthesized (Wang et al., 2009).

Due to the immense potential for technological device applications and the rich and interesting fundamental physics in single phase materials, great effort is also devoted to theoretical studies. However, multiferroics forms a very diverse class of materials. Almost every material has a unique physical mechanism and has to be studied in its own right (Picozzi and Ederer, 2009). Fortunately, the development of the numerical simulation technique provides the possibility to study the physical properties at the atomic level, such as the first-principle method with various advanced techniques including beyond-LDA/GGA approaches for the treatment of strongly correlated transition metal oxides, mostly LSDA+U (Anisimov et al., 1991; 1997), methods for the treatment of non-collinear magnetism (Hobbs et al., 2000) and spin-orbit coupling (MacDonald et al., 1980), the Berry-phase approach to calculate electric polarization (King-Smith and Vanderbilt, 1993; Resta, 1994), combined with a further analysis using maximally localized Wannier functions (Marzari and Vanderbilt, 1997), and many more. Such modern theoretical predictions really promote the renaissance of multiferroic single phase materials. Other phenomenological theories like the Landau theory to study the multiferroic composites (Lu et al., 2009), the inverse DM model to study the spiral spin-order-induced multiferroicity (Sergienko and Dagotto, 2006), and the KNB model to study the spin current system (Katsura et al., 2005) were also developed to study some kinds of single phase materials.

Generally speaking, the mechanics models of magnetoelectric composites are proper, since the ME effect can be realized in multiferroic composites through a strain/stress-mediated interaction, which enables the effective energy transfer between electric and magnetic fields. The ME composites are by far the only choice for efficient ME couplings for commercial applications because of the large ME coupling effect at room temperature. Intensive experimental and theoretical studies focused on the multiferroic composites both in ceramic or nanostructure forms. Relative progress can be seen (Nan et al., 2008; Ryu et al., 2002; Zhai et al., 2008). The largest ME effect reported in the experiment is 90 V/cmOe at electromechanical resonance (Ryu et al., 2002), but the ME effect is still much smaller than the theoretical predictions at the low frequency stage. This may suggest that there is still a large room for improvement in the material design and fabrication. Reliable theoretical models are also much needed. In the past few years, a variety of models have been proposed including the effective magnetoelectroelastic model (Harshe et al., 1993a; 1993b; Nan, 1994; Li and Dunn, 1998) and the thermodynamic theory. By using the magnetoelectroelastic model, the expressions for the ME coupling coefficients and the effective moduli

of fibrous and laminated composites, as well as the exact connections between the effective thermal moduli and the effective magneto-electroelastic moduli of two-phase composites were studied in detail. However, only linear equations with respect to the polarization, magnetization and external fields were used. Phenomenological formulae (Nan et al., 2005; Liu et al., 2006) were also constructed by using the Landau theory. The critical phase transition properties of multiferroic composites were studied in detail (Wang and Woo, 2008; Lu et al., 2009). Since the ME effect is highly dependent on the frequency, the resonant ME coupling effect was also extensively studied recently.

As for the trends toward device miniaturization and high-density data storage, an integration of multifunctions into one material system or in the form of a nanostructured composite has become highly desirable. Many related issues such as the size effect, the surface effect and quantum effect should be carefully considered. Strong ME coupling is still critical for multiferroic devices. Recently, many new routes toward mutual control of the spin and charge were also proposed with the cooperation of colossal magnetoresistance (CMR), ferromagnetic resonance (FMR) and multiferroic tunneling junctions (MTJ), and so on. Novel physics of strongly correlated electronic systems in multiferroics leads to advances in the high temperature superconductor and CMR manganites, etc.

8.2 Coupling Mechanism of Multiferroic Materials

8.2.1 *Single Phase Multiferroic Materials*

Multiferroic single phase materials exhibit a rich spectrum of properties, including ferro-/ferri-magnetic, ferroelectric, strongly correlated electron behavior, superconductivity and magnetoresistance. Although many single phase multiferroic materials show a relatively low coupling effect which exists at very low temperature, they have gained great attention and become an extremely interesting class of systems because of the amazing physical mechanism and the potential applications. The attempt to combine the dipole and spin orders into one system with consideration for the similarity of the ferroelectrics and ferromagnetics can be traced back to the 1960s (Smolenskii and Ioffe, 1958). The first discovered multiferroic material was Nickel iodine boracite ($\text{Ni}_3\text{B}_7\text{O}_{13}\text{I}$) (Ascher et al., 1966) which has complex structures with many atoms per formula unit and more than one formula unit per unit cell. In recent years, great strides have been made and lots of materials including the families of boracites, fluorides, perovskites, magnetite and some others were identified or invented.

Since most multiferroic materials discovered so far are transitional metal oxides with perovskite structures, which are strongly correlated electronic systems, the correlations between spins, charges/dipoles, orbital and lattice/phonons are significant, which leads to possible strong intrinsic ME coupling. Therefore, the

multi-latitude landscape of interactions between these orders is complex and even vague, making the physics of multiferroicity extremely complicated. In general, the ME coupling mechanism can be divided into two categories: one is a heterogenous-source (different constituent ions in one unit serve as different sources for electric and magnetic properties), and the other is a homologous-source (the electric and magnetic properties come from one constituent ion). The former, which is intensively studied, shows weak coupling, while the latter is the potential candidate material for strong ME interactions (Brink and Khomskii, 2008).

8.2.1.1 Incompatibility Between Ferroelectricity and Ferromagnetism

Since many magnetic materials and classical ferroelectric materials belong to the perovskites with transition metal (TM), the first attempt to create multiferroic materials was mostly concentrated on this class of compounds. From the point of view of symmetry consideration, ferroelectricity needs broken spatial inverse symmetry while the time reverse symmetry can be invariant. In contrast, the broken time-reversal symmetry is the prerequisite for magnetism (spin order). Therefore, a multiferroic system that is both ferromagnetic and ferroelectric requires the simultaneous breaking of the spatial-inversion and time-reversal symmetries. Among all of the 233 Shubnikov magnetic point groups, only several point groups allow the simultaneous appearance of spontaneous polarization and magnetization (Fuentes et al., 2007). Although there are hundreds of magnetic perovskites and hundreds of more ferroelectric perovskites, little overlap exists in the magnetism and FE perovskites. It seems that magnetism and FE perovskites exclude one another. So, how come this mutual exclusion? Empirical observations and first-principle calculations tell us: All conventional FE perovskites contain TM ions with a formal configuration d_0 , such as Ti^{4+} , Ta^{5+} , W^{6+} at B-sites (i.e., the TM ions with an empty d -shell), and magnetism. In contrast, we require TM ions at the B-site with partially filled shells (always d - or f -shells), such as Cr^{3+} , Mn^{3+} , Fe^{3+} , because of the spins of electrons. So it is obvious that the difference in filling the TM ion d -shells at the B-site makes these two ordered states mutually exclusive. It should be noted that this question was first raised in 1999 during the workshop on quantum magnetism at the ITP in Santa Barbara (Khomskii, 2001), and was largely elaborated by Hill (2000) based on first-principle calculations. The possible explanation in more detail can be seen in (Khomskii, 2006; Wang et al., 2009).

As mentioned above, the origin of strong magnetism in insulators is induced by the partially filled inner shells (d - or f -levels), while the off-centre shifts of the TM ions provide the main driving force for FE. But why do we need empty d -shells for that? The *ab initio* calculations give a qualitative answer: empty d -states of TM ions like Ti^{4+} may be used to establish strong covalency with the surrounding oxygens and it may be favorable to shift the TM ion from the centre of O_6 octahedra towards one (or three) oxygen(s), to form a strong covalent bond

with this particular oxygen(s) at the expense of weakening the bonds with other oxygens (Wang et al., 2009), see Fig. 8.3a. The hybridization matrix element t_{pd} is changed to $t_{pd}(1 \pm gu)$, where u is the distortion and g is the coupling constant. In the linear approximation with respect to charge transfer gap Δ , the corresponding terms in the energy $\sim (-t_{pd}^2 / \Delta)$ cancel each other (Khomskii, 2006). But in the second order in u , an additional energy difference can be gained:

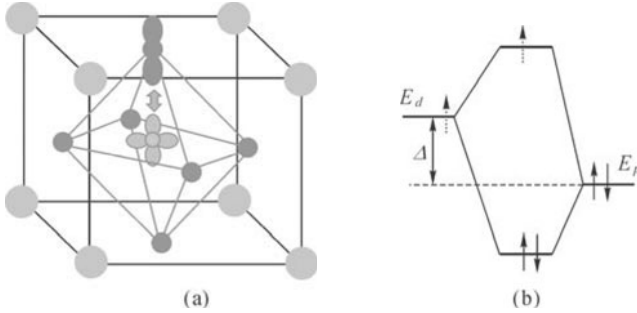


Fig. 8.3 (a) Orbital configuration of O-TM-O chain unit in perovskites ABO_3 cell; (b) The corresponding energy levels. The B-site TM ions with d_0 configuration tend to move toward one of the neighboring oxygen anions to form a covalent bond

$$\begin{aligned} \delta E &\cong -t_{pd}^2 (1 + gu)^2 / \Delta - t_{pd}^2 (1 - gu)^2 / \Delta + 2t_{pd}^2 / \Delta \\ &= -2t_{pd}^2 (gu)^2 / \Delta. \end{aligned} \quad (8.2.1)$$

If the corresponding total energy gain $\sim u^2$ exceeds the energy loss due to the ordinary elastic energy of the lattice distortion, such a distortion would be energetically favorable and the system would become ferroelectric.

From Fig. 8.2, one observes that only the bonding bands would be occupied (solid arrows) if the TM ion has an empty d -shell, a process that allows only for electronic energy. If there is an additional d -electron on the corresponding d -orbital (dashed arrow), this electron will occupy an anti-bonding hybridized state, thus suppressing the total energy gain. This may be one factor suppressing the tendency of magnetic ions to FE for magnetic ions.

But surely, the incompatibility between ferroelectricity and magnetism has even more complicated origins than the above model and this is not enough to explain the almost total mutual exclusion of magnetism and FE in perovskites. For example, CaMnO_3 or RCrO_3 (R is a rare-earth ion) is not FE. More details of the mechanism still deserve further attention.

8.2.1.2 Coupling Models of Single Phase Multiferroic Materials

As stated above, ferroelectric perovskite oxides need B-site TM ions with an

empty d -shell to form ligand hybridization with the surrounding anions. This type of electronic structure likely excludes magnetism. In fact, not all experimental and theoretical results support the argument. There are some routes for the coexistence of the two orders. Generally, there are two groups which are named type-I (heterogenous-source) and type-II (homologous-source) multiferroics as shown in Table 8.2 (Brink and Khomskii, 2008).

Table 8.2 Multiferroics typical systems

	Mechanism	Materials
Type I	FE induced by lone pair	BiMnO ₃ , BiFeO ₃
	FE induced by geometric distortion	Hexagonal RMnO ₃ (YMnO ₃ , HoMnO ₃)
	Charge ordering	Fe ₃ O ₄ , (PrCa)MnO ₃ , LuFe ₂ O ₄
Type II	Combined site and bond in-equivalence in a magnetic order	RMn ₂ O ₅ , Ca ₃ CoMnO ₆ , CoCr ₂ O ₄
	Magnetic E-type ordering	RNiO ₃
	Inverse DW effect	TbMnO ₃ , Ni ₃ V ₂ O ₈ , MnWO ₄

In type I, ferroelectricity and magnetism have different origins; i.e., they are induced by different active subsystems where the ferroelectric and magnetic order parameters break the spatial inversion symmetry and time reversal symmetry, respectively. However, the two order parameters can be coupled and coexist by some mechanisms including the presence of bismuth or lead in magnetic materials with lone pairs of Bi³⁺ and Pb²⁺ and the presence of geometric ferroelectricity caused by the rotation of rigid M-O polyhedra where the magnetic order is also induced by the magnetic atoms. In these type-I materials, the Curie temperature of ferroelectricity usually is much higher than that of magnetic one (Brink and Khomskii, 2008).

In type II, the ferroelectric and magnetic order parameters occur in one subsystem; i.e., ferroelectricity sets in at the same temperature as a certain type of magnetic ordering and is driven by it. There may be two microscopic origins. One of them probably is the most common system like TbMnO₃, Ni₃V₂O₈ and MnWO₄, multiferroic pyroxenes and some other systems. Ferroelectric order is induced by the inverse Dzyaloshinskii-Moriya effect, which operates in systems with non-collinear, usually spiral spin orders. The other one is based on the exchange striction induced by the inequivalent magnetic ions with different charges. For example, the ferroelectricity of BiMn₂O₅ is induced by the simultaneous presence of sites with different charges and with inequivalent bonds occurring in a magnetically ordered state. Some other examples include rare earth nickelates (RNiO₃) and Ca₃CoMnO₆ (Choi et al., 2008).

A simple way to reach ME coupling of type I began in the 1950s taken by the Russian groups with the replacement of some of the B (d_0) cations in ferroelectric perovskite oxides by magnetic ions. Many such combinations of the type AB_{1-x}B'_xO₃ have been found; for example, (1-x)PbFe_{0.66}W_{0.33}O₃-xPbMg_{0.5}W_{0.5}O₃, PbFe³⁺_{1/2}Nb⁵⁺_{1/2}O₃, and PbFe³⁺_{2/3}W⁶⁺_{1/3}O₃ with Mg²⁺, Nb⁵⁺ and W⁶⁺ caused ferroelectricity and Fe³⁺(d^5)

ions are responsible for the magnetic ordering (Smolensky et al., 1961). Other examples include B site disordered $\text{Pb}_2(\text{Co,W})\text{O}_6$ and $\text{Pb}_2\text{FeTaO}_6$ and $\text{Pb}_2(\text{FeTa})\text{O}_6$ (Brixel et al., 1988). However, because of the independent subsystems, the coupling is weak, but still provides a very useful approach in searching for multiferroic materials.

By far, all arguments presented above are applicable to systems in which TM ions are responsible for FE. However, in some systems, the FE has nothing, or very little, to do with the presence of TM ions. In more recent research, great efforts were devoted to search for ferroelectricity that is intrinsically generated by special spin orders where the spin exchange interaction can be mapped into the virtual hopping of electrons between the adjacent ions. In such systems, the uncompensated spin exchange interaction between adjacent magnetic ions is induced by the long-range spin order and macroscopic magnetization. This enables not only an effective combination of the two orders but also the spontaneous mutual control of them. Some other novel multiferroic materials, in which ferroelectricity was induced by a geometric distortion and a helical/conical spin order, as well as a charge-ordered (CO) structure, have also been synthesized.

● *Bi perovskites: ferroelectricity induced by lone-pairs*

In the perovskites family, there are two apparent exceptions from the exclusion rule mentioned in subsection 8.2.1.1. They are BiMnO_3 and BiFeO_3 , where the B-site ions, namely magnetic TM ions $\text{Fe}^{3+}(d^5)$ and $\text{Mn}^{3+}(d^4)$, contribute to the magnetism and the A-site Bi^{3+} ions with two electrons in a $6s$ orbit (lone pair) via the lone pair mechanism lead to the ferroelectricity. As is well known, Bi^{3+} , and also Pb^{2+} , are known to have the so-called lone pairs (two valence electrons which can participate in chemical bonds using (sp)-hybridized states). The existence of lone pairs or dangling bonds may also favor breaking the inversion symmetry and create local dipoles, thus inducing and stabilizing the FE or anti-FE order.

Bulk BiMnO_3 possesses a distorted perovskite structure, having monoclinic symmetry with space group $C2$ and lattice parameters $a=9.5323(6)$ Å, $b=5.6064(3)$ Å, $c=9.8535(7)$ Å and $\beta=110.667(5)$ Å (Atou et al., 1999; Moreira dos Santos et al., 2002). Schematic illustrations of the crystal structure of BiMnO_3 are shown in Fig. 8.4. The distortion is caused by a polarized Bi $6s^2$ lone pair as calculated from first-principles electronic structure calculations (Seshadri and Hill, 2001). Every Mn^{3+} cation reveals Jahn–Teller distortion and ordering of a vacant $d_{x^2-y^2}$ orbital is suggested to play an important role for the ferromagnetism (Atou et al., 1999). The observed magnetic moment at a temperature of 5 K and magnetic field of 5 T reaches $3.6 \mu_B$, which is slightly smaller than that of the fully aligned spin value of $4 \mu_B$ for Mn^{3+} (Atou et al., 1999). The lattice instability to the off-centered displacement is due to the strong covalent bonding between Bi $6p$ and O $2p$ states which lead to the ferroelectricity (Shishidou et al., 2004).

On the other hand, BiMnO_3 has a triclinically distorted perovskite structure

with $a=c=3.935$ Å, $b=3.989$ Å, $\alpha = \gamma = 91.4^\circ$, $\beta = 91.0^\circ$ (Chiba et al., 1997). The lower-perovskite-symmetry structure has been explained by the highly polarized $6s^2$ character of Bi^{3+} (Sugawara et al., 1968). Ferroelectric phase transition occurred near 490°C with a first-order structural change from triclinic to tetragonal. The remanent polarization is about 62 nC/cm^2 with a coercive field of about $8.1\text{ kV}\cdot\text{cm}^{-1}$ at 87 K (Moreira dos Santos et al., 2002).

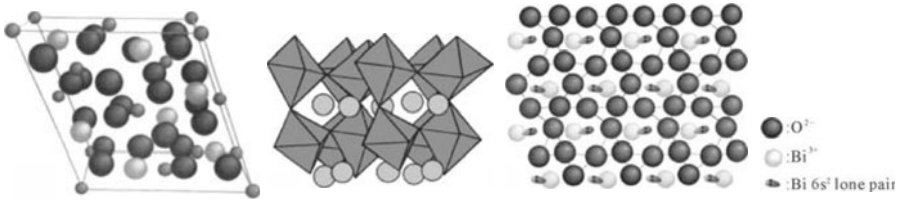


Fig. 8.4 Schematic illustrations of crystal structure of BiMnO_3 . (a) Unit cell view; (b) A part of the crystal structure on the base of the perovskite unit with MnO_6 Octahedra and bismuth cations; (c) The close packed plane which consists of bismuth cations and oxide anions. Polarized Bi $6s^2$ lone pairs are shown schematically (Atou et al., 1999). Copyright (1999) Elsevier B.V.

Properties of BiMnO_3 thin film are largely dependent on the substrate. The Curie temperature is a few degrees lower than that of the bulk material, which may be attributed to strain, non-stoichiometry, or size effects (Moreira dos Santos et al., 2004). Ferromagnetic phase transition of BiMnO_3 thin film on (100) Si polycrystalline (111) orientation occurred at 105 K and on LaAlO_3 with lattice parameter $a=3.792$ Å at 50 K , and on SrTiO_3 substrate with lattice parameter $a=3.905$ Å at 97 K . The remanent magnetization on the LaAlO_3 is larger than that on the $\text{Pt/TiO}_2/\text{SiO}_2/\text{Si}$ substrate indicating the magnetic ordering of the epitaxial thin film is better than that of the polycrystalline thin film (Son et al., 2004). The strain and temperature effect were also investigated by the resonant X-ray scattering as shown in Fig. 8.5 (Yang et al., 2006).

Another Bi perovskite BiFeO_3 (BFO) also gained much attention recently which may be stimulated by the paper published in *Science* in 2003 by the group of Ramesh (Wang et al., 2003). Perovskite BFO is one rare example of multiferroic material with ME coupling effect above room temperature. Early studies indicated that bulk BFO with $R3c$ rhombohedral structure ($a=5.58$ Å and $\alpha = 89.5^\circ$) was a G-type antiferromagnet with a Néel temperature of about 643 K (Kiselev et al., 1963; Fischer et al., 1980) and possessed a cycloidal spin structure with a period of about 620 Å (Sosnowska et al., 1982). Ferroelectric phase transition occurred at about 1103 K (Teague et al., 1970). The spontaneous polarization by current voltage measurements is $3.5\text{ }\mu\text{C/cm}^2$ along the $[100]$ direction (corresponding to $6.1\text{ }\mu\text{C/cm}^2$ along the polar $[111]$ axis). The values are small compared with the first-principles results, which may be reduced due to the high conductivity of the available samples (Ederer and Spaldin, 2005). The structure of bulk BFO can be characterized by two distorted perovskite blocks

connected along their body diagonal or the pseudocubic (111) to build a rhombohedral unit cell (Fig. 8.6) (Martin et al., 2010).

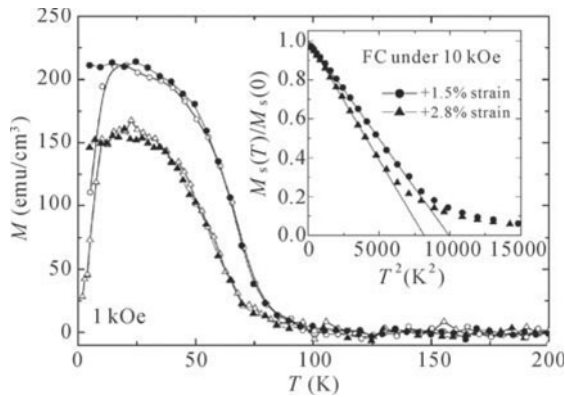


Fig. 8.5 Zero-field cooled (open symbol) and field-cooled (solid symbol) magnetic properties of the strained BiMnO_3 films (+1) on SrTiO_3 (001) are shown as a function of temperature, respectively. The in-plane field strength was 1 kOe. Inset displays that the saturation magnetization decreases with the T^2 behavior (Yang et al., 2006). Copyright (2006) by the American Institute of Physics

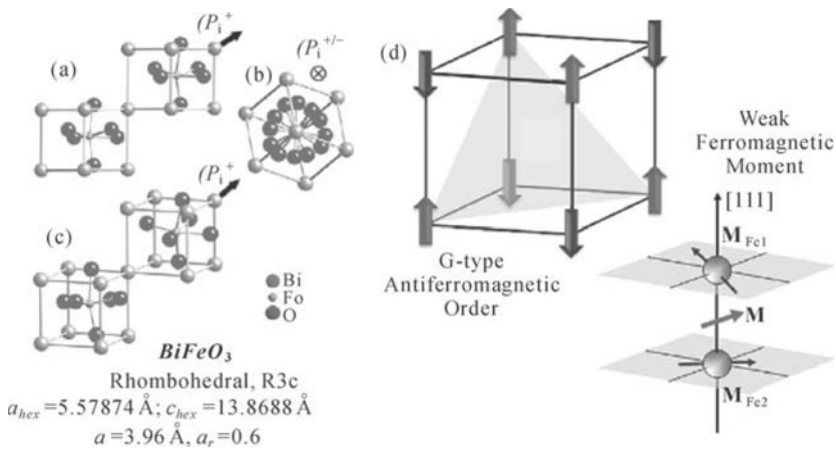


Fig. 8.6 Structure of BiFeO_3 . (a) Looking down the pseudocubic 110; (b) Looking down the pseudocubic 111 polarization direction; (c) A general three-dimensional view of the structure; (d) The magnetic structure including the G-type antiferromagnetic ordering and the formation of the weak ferromagnetic moment (Martin et al., 2010). Copyright (2010) Elsevier B.V.

Recently, studies have shown that the BFO grown on SrTiO_3 substrate revealed an unexpectedly large remnant polarization of about $90 \mu\text{C}/\text{cm}^2$, 15 times larger than previously seen in bulk, together with very large ferromagnetism of ca. 1.0 Bohr magneton (μ_B) per unit cell ($M_r \sim 150 \text{ emu/cc}$). A large ME coupling

effect was found at room-temperature. The epitaxial strain plays a very important role in the high value of P_r and M_r . However, a recent study showed that the epitaxial strain does not enhance M_r in BFO thin films (Eerenstein et al., 2005). It is believed that the heteroepitaxy induces significant and important structural changes in BFO thin films, which may lead to higher values for P_r than found in single crystals (Palai et al., 2008). Moreover, research has revealed that the in-plane strains in the thin films could drive a rotation of the spontaneous polarization on the (110) plane, while the polarization magnitude itself remains almost constant, which is responsible for the strong strain tunability of the out-of-plane remnant polarization in (001)-oriented BiFeO₃ films (Jang et al., 2008).

The crystal structure of (001) epitaxial BFO thin films on STO substrates has led to many controversies in the literature (tetragonal (Wang et al., 2003; Singh et al., 2005), rhombohedral (Qi et al., 2005; Das et al., 2006), and monoclinic (Xu et al., 2005)). Palai et al. (2008) reported that the BFO (001) thin films are clearly monoclinic, in agreement with synchrotron data but contrary to the bulklike rhombohedral and tetragonal structures reported earlier; the β phase is unambiguously orthorhombic, in disagreement with all other recent publications. BFO is usually prepared from equal parts of Bi₂O₃ and Fe₂O₃, and under high temperatures, which can be seen in the phase diagram of the Bi₂O₃-Fe₂O₃ system given by Palai et al. (2008) (Fig. 8.7). More information about the phase diagram and the material structure can be seen in (Palai et al., 2008).

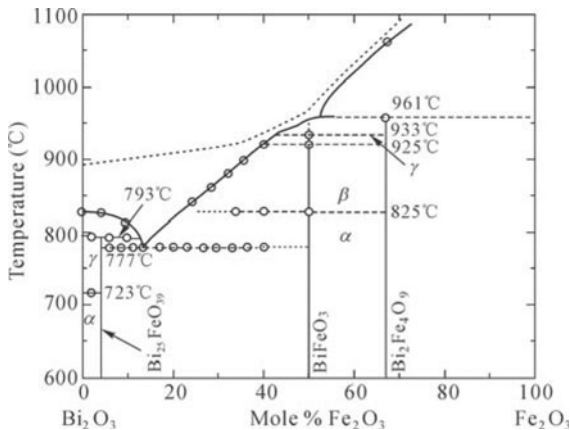


Fig. 8.7 Phase diagram of the Bi₂O₃-Fe₂O₃ system. Open circles show the data points obtained by DTA. The dotted line above the liquidus represents the approximate temperature limit not to be surpassed to avoid decomposition, otherwise correct equilibrium DTA peaks are no longer observed upon a second heating. The α , β and γ phases are rhombohedral, orthorhombic and cubic, respectively (Palai et al., 2008). Copyright (2008) by the American Physics Society

• Hexagonal manganites: geometric ferroelectricity

The compounds REMnO₃, where RE is a rare-earth element, have attracted much

attention since the middle of the 20-th century. Rare earth perovskite compounds with large RE ionic element (La and Ce-Dy) crystallize in a distorted orthorhombic perovskite structure (space group: $Pnma$) where the Mn atom is octahedrally coordinated by O atoms and the octahedra forms a corner-shared three-dimensional network (Yakel, 1955). A colossal magnetoresistance effect, i.e., a metal-insulator transition that changes the conductivity by many orders of magnitude at the Curie temperature, had been found (Ramirez, 1997). With the reduction of the ionic radius of the A cation ($A=Y, Ho, Er, Tm, Yb, Lu$) as one goes from Y to Lu, the orthorhombic and the Jahn-Teller distorted perovskite structure become less stable and the structure of $REMnO_3$ often becomes hexagonal (space group: $P6_3cm$) as seen in Fig. 8.8 (Prellier et al., 2005). The small ionic radius lanthanides, Ho-Lu, crystallize in a hexagonal structure where the Mn ions are coordinated by a trigonal bipyramid of O atoms, forming a pseudo-layered structure (Yakel et al., 1963). Magnetic ordering occurs in both structures at very low temperature, less than or around 100 K; However, ferroelectricity is possible only for the hexagonal phase which has a non-centrosymmetric structure at a very high temperature ($T_{FE} \sim 900$ K).

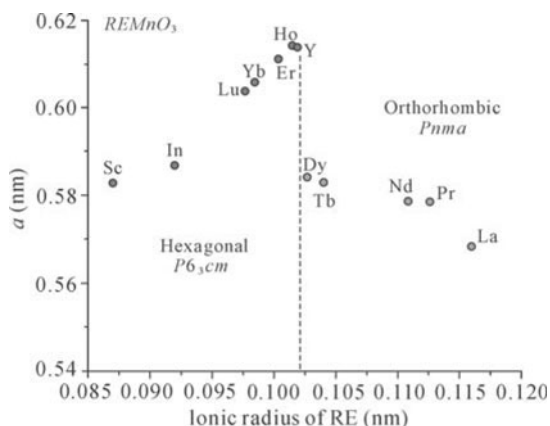


Fig. 8.8 Evolution of the lattice structure in $REMnO_3$ as a function of the size of the rare earth (Prellier et al., 2005). Copyright (2005) by Institute of Physics

In the adjacent region, as shown in Fig. 8.8, the structure and properties are very different from the early works published in the 1960s. For example, early research showed that the hexagonal $YMnO_3$ possessed the $P6_3cm$ space group with lattice parameters $a=6.125$ Å and $c=11.41$ Å at 290 K (Yakel et al., 1963). Its structure consists of layers of two-dimensional connected distorted MnO_5 bipyramids separated by distorted layers of rare earths, leading to frustrated $Mn^{3+}-O^{2-}-Mn^{3+}$ in-plane exchange paths and $Mn^{3+}-O^{2-}-O^{2-}-Mn^{3+}$ interplane exchange paths (Lonkai et al., 2003). Recently, studies showed that $YMnO_3$ was hexagonal under ambient conditions (Brinks et al., 1997). However, the orthorhombic structure can be stabilized by either low-temperature or high-

pressure synthesis (Waintal and Chevanas, 1967), or epitaxial thin film growth (Salvador et al., 1998). Bulk YMnO_3 revealed an A-type antiferromagnetic order with non-collinear Mn spins oriented in a triangular arrangement (Bertaut et al., 1963). The magnetoelectric effect was small, even in the presence of a 5 T field at its low antiferromagnetic stage of about 80 K (Huang et al., 1997).

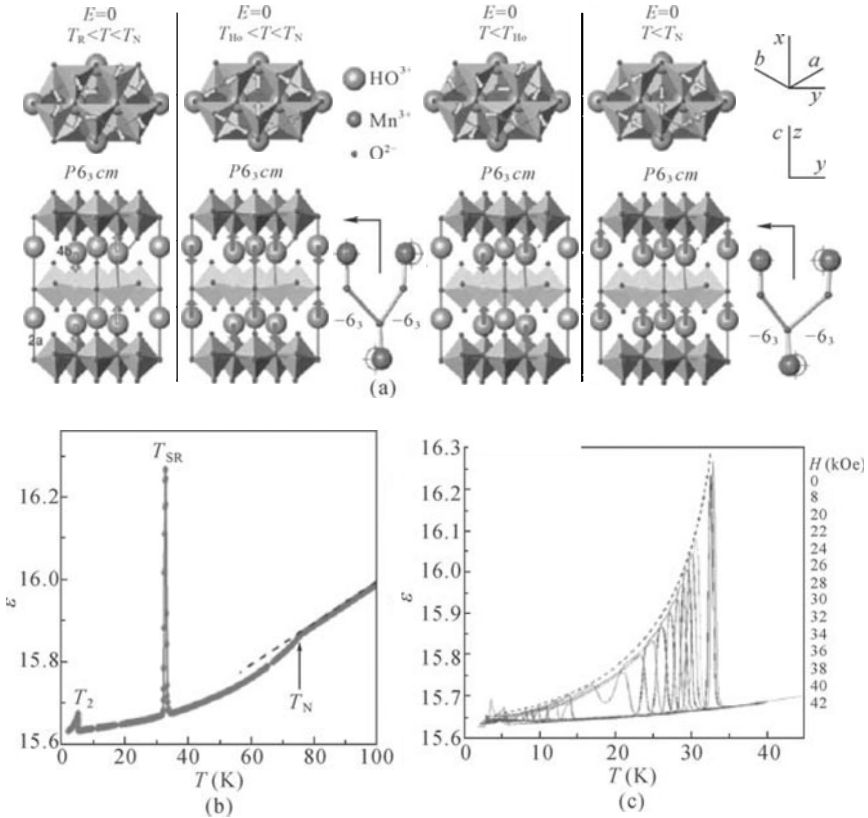


Fig. 8.9 Spin configurations and lattice symmetry of HoMnO_3 in different temperature ranges with and without electric field (Lottermoser and Lonkai, 2004). Copyright (2004) by the Nature Publishing Group

Strain engineering is now very useful in improving the ME coupling properties of YMnO_3 film. The YMnO_3 films were first grown via radio-frequency magnetron sputtering and epitaxial (0001) films on MgO (111) and ZnO (0001)/sapphire (0001) and polycrystalline films on Pt (111)/ MgO (111) were obtained (Norifumi et al., 1996). The structure of the YMnO_3 turned out to be perovskite instead of the hexagonal phase. This is a perfect example of strain engineering. Later on, many different substrates were used and results showed that the induced epitaxial strain could drive the hexagonal phase of YMnO_3 to a metastable, nonferroelectric orthorhombic perovskite phase by growth on the

appropriate oxide substrates including SrTiO_3 (001) and NdGaO_3 (101) (Salvador et al., 1998). This work was of great interest because it was the first evidence of a competition between hexagonal and orthorhombic YMnO_3 phases and how epitaxial thin film strain could be used to manipulate the structure of this material.

In fact, another fascinating effect with hexagonal RMnO_3 is that the magnetic phase can be controlled by an electric field, as demonstrated in HoMnO_3 . Using an optical SHG technique, it was observed that at T_N , an external electric field could drive HoMnO_3 into a magnetic state different from that under a zero electric field, thereby modulating the magnetic order of the Mn^{3+} sublattice, as shown in Fig. 8.9.

Moreover, compared with YMnO_3 , HoMnO_3 has an extra magnetic sublattice consisting of Ho^{3+} ions, which shows an interesting response to an electric field. In the presence of an electric field, the para- or antiferromagnetic state under a zero field is converted into a ferromagnetic order with strong macroscopic magnetization. The proposed mechanism for this phase control is the microscopic magnetoelectric coupling originating from the interplay of the Ho^{3+} - Mn^{3+} interactions and ferroelectric distortion (Lottermoser et al., 2004).

• *Charge ordering multiferroics*

Conventionally, ferroelectricity originates from the relative displacement between anions and cations (such as the covalent bonding between $3d_0$ transition metal and oxygen), the lattice distortion associated with the second-order Jahn-Teller effect (polarization of $6s^2$ lone pair of Bi or Pb), and the geometric deformation (like RMnO_3). The other two kinds of ferroelectricity can be induced by the charge ordering, which is also called electric ferroelectrics, and magnetic ordering which is named magnetic ferroelectrics (Cheong and Mostovoy, 2007). In many strong electronic correlation systems, typically narrow-band metal oxides, charge carriers may become localized at low temperature and form a periodic structure. For example, iron oxides, such as magnetite Fe_3O_4 , are also famous charge ordered materials with ferrimagnetic ordering occurring below about 860 K and the metal-insulator transition (the Verwey transition) at approximately 120 K below, with a rather complex iron charge order pattern and ferroelectric order in the insulating state below about 120 K (Verwey and Haayman, 1941). It is expected that there is a special group of multiferroics in which a non-symmetric charge order may induce electric polarization. Recently, multiferroicity was found in perovskite manganites $(\text{PrCa})\text{MnO}_3$, frustrated charge ordered system LuFe_2O_4 , magnetite Fe_3O_4 , the complex manganites RMn_2O_5 , $\text{Ca}_3\text{CoMnO}_6$ and RNiO_3 .

The mechanism of the ferroelectricity due to the charge ordering can be explained in Fig. 8.10 (Brink and Khomskii, 2008). Fig. 8.10a shows the state of a homogeneous crystal with an equal charge on each site. Fig. 8.10b shows the same chain after a charge ordering at which the sites become inequivalent. This process does not break spatial inversion symmetry. Fig. 8.10c shows the charge ordering when the system dimerizes. In this case, the bonds are inequivalent as strong and weak bonds alternate, but the sites remain equivalent. The combined condition can be illustrated in Fig. 8.10d where inversion symmetry is broken, which results in a

net dipole moment and the ferroelectricity of the whole system (Efremov et al., 2004).

The simultaneous presence of inequivalent sites and bonds has many different origins, each of which may be studied separately on a microscopic level. In some cases, bonds are inequivalent just because of the crystallographic structure; in some other cases, ions with different valences cause the structural distortion, which induces a ferroelectric order. Of course, the two cases can occur simultaneously.

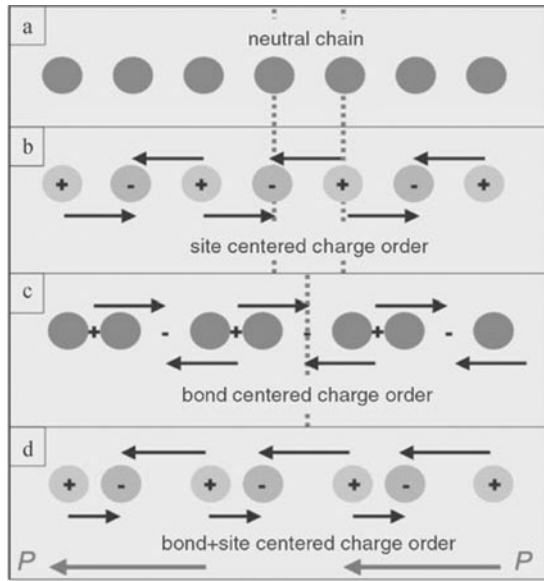


Fig. 8.10 (a) Example of a neutral one-dimensional chain exhibiting (b) site-centered charge ordering, (c) bond-centered charge ordering, and (d) a linear combination of these two that is ferroelectric. The arrows indicate the polarization, which is in total zero in (b) and (c), but develops a macroscopic moment, indicated by the red arrow in (d). The dashed lines in (a), (b) and (c) indicate mirror planes of the system (Brink and Khomskii, 2008). Copyright (2008) by Institute of Physics

The simultaneous presence of inequivalent sites and bonds is quite common in manganites $R_xA_{1-x}MnO_3$, in which ions formally have a mixed valence, such as $La_xCa_{1-x}MnO_3$, and $Pr_xCa_{1-x}MnO_3$. For example, half-doped manganites $Pr_{0.5}Ca_{0.5}MnO_3$ or $La_{0.5}Ca_{0.5}MnO_3$ have one extra electron/hole per two Mn. The Mn^{3+} and Mn^{4+} ions can be arranged in a particular case, as shown in Fig. 8.11 (Tokura and Nagaosa, 2000; Nakamura et al., 2010).

The validity of the conventional CO picture in which the Mn^{3+} and Mn^{4+} ions orbitally order in a checkerboard arrangement has been debated for a long time. Recently, a kind of superposition of these two different charge-ordered patterns was predicted. The local dipole moments that add up to a macroscopic ferroelectric polarization can be illustrated by this superposition (Ederer and

Spaldin, 2004; Efremov et al., 2004). Moreover, electrons around the atoms may have several choices among their energetically equivalent (or degenerate) electronic orbits. This orbital degree of freedom allows for a manifold of possible electronic states to be chosen which may not be independent.

The feature of site CO versus bond CO was successfully treated by using a double exchange model with the consideration of the double degeneracy of active e_g electrons and the underlying magnetic structure. Results showed that there were different charge-ordered states with very close energy levels. These results were also confirmed by density functional calculations. More details can be found in the review (Brink and Khomskii, 2008) and the references therein.

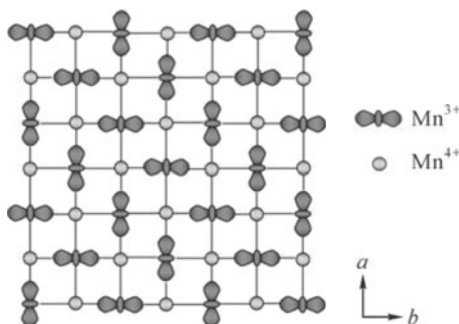


Fig. 8.11 CO-OO structure schematic view of $\text{Pr}_{0.5}\text{Ca}_{0.5}\text{MnO}_3$

Charge frustration system LuFe_2O_4 also exhibits a remarkable magnetoelectric coupling effect. Although its chemical formula is similar to that of a spinel, the structure is quite different from a spinel. Instead, a double layer structure with a triangular iron lattice within each layer is shown in Fig. 8.12 (Xiang and Whangbo, 2007).

LuFe_2O_4 belongs to a class of materials with a general formula RFe_2O_4 (R: Dy–Lu) and crystallizes in the rhombohedral structure (space group $R\bar{3}m$, $a = 3.44 \text{ \AA}$ and $c = 25.28 \text{ \AA}$). In LuFe_2O_4 , an equal number of Fe^{2+} and Fe^{3+} ions coexist on the same site of the triangular lattice, which results in an average valence of $\text{Fe}^{+2.5}$ and a possible charge order. Indeed, at about 330 K, the structure forms the charge-ordered state where the Coulomb force prefers pairing with oppositely signed charges Fe^{2+} and Fe^{3+} to be in the lowest energy. An electric polarization as high as about 0.24 C/m^2 was measured, which is comparable to the polarization of 0.26 C/m^2 in BaTiO_3 (Ikeda et al., 2005). In response to temperature variation, significant decaying of the polarization occurs at about 250 K, i.e., the magnetic transition point, and at about 330 K the CO superstructure of Fe^{2+} and Fe^{3+} disappears (Subramanian et al., 2006).

There are still some other types of charge ordering multiferroics which can be categorized into the type-II multiferroics as mentioned at the beginning of this section. The coupling mechanism of each material is unique and should be analyzed in a separated way. More information about these kinds of multiferroics can also be

found in the review (Brink and Khomskii, 2008) and the references therein.

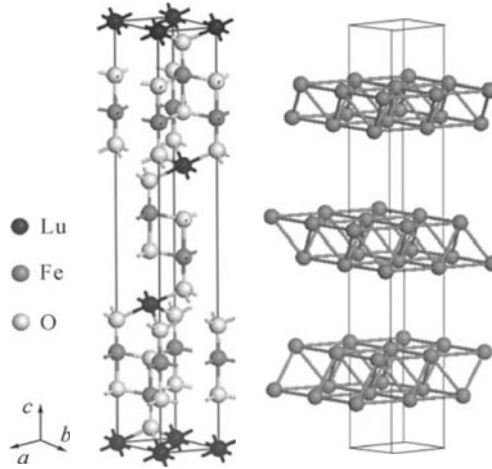


Fig. 8.12 Crystal structure of LuFe₂O₄ showing the layered arrangement of Lu (large dark-grey spheres), Fe (small black spheres), and oxygen (large white spheres) along the *c*-axis (left). The Fe double layers are shown with a triangular interconnectivity (right)

• *Multiferroics by the same atoms: spiral spin order*

So far, the above mentioned multiferroic mechanisms are somewhat disappointing because the weak coupling originates from different ions or subsystems. To fabricate the strong coupling materials, one possible way is by ferroelectricity induced directly by the spin order, meaning that an intrinsic magnetoelectric coupling occurs between the ferroelectric and magnetic order parameters. Keeping this in mind, the primary problem is how to overcome the inter-exclusion between ferroelectricity and ferromagnetism so that any special spin order can induce ferroelectricity.

With the consideration of the symmetry, a multiferroic system that is both ferromagnetic and ferroelectric requires the simultaneous breaking of the spatial-inversion and time-reversal symmetries. In the simplest case of cubic symmetry, the magnetoelectric coupling term can be written as (Mostovoy, 2006; Betouras et al., 2007)

$$\Phi_{em}(\mathbf{r}) = \mathbf{P} \cdot \{ \gamma [\mathbf{M}(\nabla \cdot \mathbf{M}) - (\mathbf{M} \cdot \nabla) \mathbf{M}] + \dots \}, \quad (8.2.2)$$

where, $\mathbf{r}$, $\mathbf{P}$ and $\mathbf{M}$ are vectorized spatial coordinates, polarization and magnetization, and γ is the coupling coefficient. The omitted terms can be written as the total derivative $\nabla f(\mathbf{M})$ and do not contribute to the uniform polarization. If the system is stable towards ferroelectricity in the absence of magnetism, only the quadratic term in the ‘electric part’ of the thermodynamic potential needs to be considered. The electric energy term Φ_e associated with $\mathbf{P}$ can be written as $\mathbf{P}^2 / 2\chi_e$, where

χ_e is the dielectric susceptibility. From the variation of $\Phi_{em} + \Phi_e$ with respect to the polarization for a given magnetization, one can find the value of the static polarization field:

$$\mathbf{P} = \gamma\chi_e[\mathbf{M}(\nabla \cdot \mathbf{M}) - (\mathbf{M} \cdot \nabla)\mathbf{M}]. \quad (8.2.3)$$

This simple symmetry argument predicts the possible multiferroicity in spinfrustrated systems which always prefer to have spatially inhomogeneous magnetization owing to the competing interactions. The frustrated state is characterized by a spiral spin order, also named spiral spin-density wave (SDW):

$$S_n = S_1 e_1 \cos \mathbf{Q} \cdot \mathbf{r} + S_2 e_2 \sin \mathbf{Q} \cdot \mathbf{r}, \quad (8.2.4)$$

where the unit vectors e_i ($i=1,2,3$) form an orthogonal basis, e_3 is the axis around which spins rotate and vector $\mathbf{Q}$ is given by $\cos(\mathbf{Q}/2) = -J''/(4J)$. If only S_1 or S_2 is non-zero, this equation describes a sinusoidal SDW, which cannot induce any FE order because it is invariant upon the spatial inversion operation. If the S_1 and S_2 are both non-zero, Eq. (8.2.4) described a spiral spin order with the spin rotation axis e_3 . Using Eqs. (8.2.3) and (8.2.4), one can find that the average polarization is transverse to both e_3 and $\mathbf{Q}$:

$$\bar{\mathbf{P}} = \frac{1}{V} \int d^3 \mathbf{x} \mathbf{P} = \gamma' \chi_e S_1 S_2 [\mathbf{e}_3 \times \mathbf{Q}]. \quad (8.2.5)$$

The above simple model can be extended to two- or three-dimensional spin systems. In general, two or more competing magnetic interactions can induce the spin frustration and the spiral (helical) spin order which, in turn, breaks the spatial inversion and time-reversal symmetries simultaneously, thus establishing the ferroelectric order. However, whether the spiral spin order is a prerequisite for the origin of ferroelectricity is still unclear. More details can be found in (Wang et al., 2009).

In fact, multiferroics with enhanced cross-coupling effects exhibit magnetic orders with broken centrosymmetry, and are associated with magnetism-induced ferroelectricity. The exchange striction is a tensorial quantity with symmetric parts and antisymmetric parts. The symmetric parts are relevant to multiferroicity in systems such as RMn_2O_5 , and the antisymmetric parts which constitute the so-called inverse Dzyaloshinskii-Moriya (DM) interaction, are active when ferroelectricity is induced by spiral-magnetic orders in systems (Park et al., 2007) such as quasi-1D systems: $LiCu_2O_2$, $(Ba,Sr)_2Zn_2Fe_{12}O_{22}$; quasi-2D systems $Ni_3V_2O_8$, $CuFeO_2$, $RbFe(MoO_4)_2$ (RFMO) and $CuCrO_2$, $LiCrO_2$ and $NaCrO_2$; and quasi-3D systems: $TbMnO_3$, $DyMnO_3$, $CoCr_2O_4$, and $MnWO_4$ (Sergienko and Dagotto, 2006; Qu et al., 2000). In addition to the inverse DM model, theories such as the spin current model (KNB model) (Katsura et al., 2005), and the electric current cancellation model (Hu, 2008) were also developed and led to the

description of the mechanism of La_2CuO_2 , Cr_2O_3 , Fe_2O_3 , and so on. In the following, the parallel experimental evidence hinted that the current DW model would be presented.

The spin configuration of LiCu_2O_2 can be approximately treated as a quasi-1D spin chain system with magnetic Cu^{2+} ions bonded with oxygen ions and isolated by the non-magnetic copper ion layer. The single phase of LiCu_2O_2 is orthorhombic with a space group $Pnma$, and lattice parameters $a=5.734(4)$ Å, $b=2.856(2)$ Å, and $c=12.415(6)$ Å. Within the theoretical framework addressed above, the ferroelectric polarization along the c -axis (P_c) would be expected and was experimentally evident in LiCu_2O_2 . The intrinsic magnetoelectric coupling between the spin order and ferroelectric order, which is evident in the response of polarization to an external magnetic field, can be clearly seen in the measurements with the application of the external magnetic field. The behavior shown in Figs. 8.13a–8.13f indicates that a ferroelectric state emerges when a magnetic order sets in. By applying a magnetic field H_b , P_c decreases, and becomes negligible at $H_b=9$ T. Such experimental results can also be found in Figs. 8.13e and 8.13f. The polarization rotated from c -axis to a -axis with the application of the external magnetic field along the b -axis.

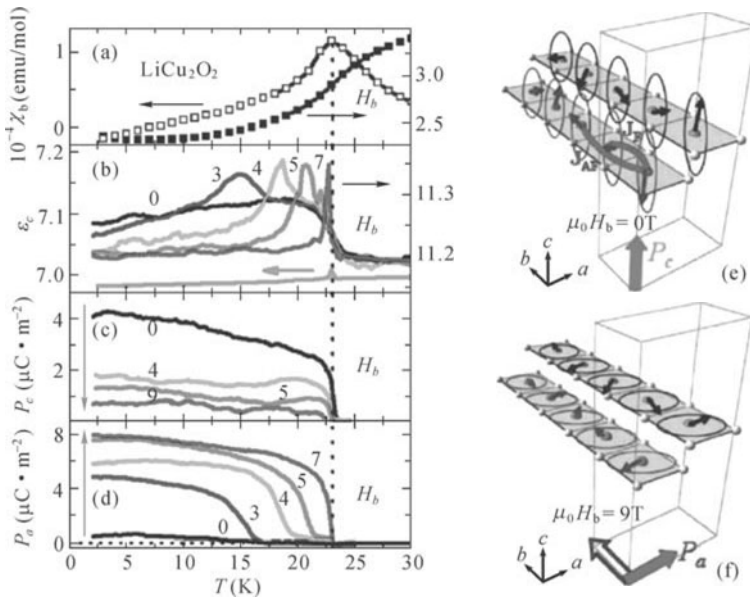


Fig. 8.13 Physical properties of LiCu_2O_2 as a function of temperature. (a) Magnetic susceptibility along the b -axis and its temperature derivative; (b) Dielectric constant along the c -axis; (c) Polarization along the c -axis and that along the a -axis; (d) Under various magnetic fields as numbered (in T). Spiral arrangements of the Cu spin ladders and the corresponding polarization under (e) zero magnetic field and (f) 9.0 T applied along the b -axis (Park et al., 2007)

Kagome staircase $\text{Ni}_3\text{V}_2\text{O}_8$ is a typical example of the two-dimensional (2D) frustrated spin system (Lawes et al., 2005) and can also be explained by the

inverse DM model (Fig. 8.14). However, the origin of ferroelectricity in $\text{RbFe}(\text{MoO}_4)_2$ (RFMO) which exhibits the typical easy-plane triangular lattice with space group $P\bar{3}m1$ at room temperature can be explained by neither the KNB model nor the inverse DM model. In addition to RFMO and CuCrO_2 , LiCrO_2 , CuFeO_2 and NaCrO_2 also exhibit the 2D triangular-lattice structure, but they do not exhibit any electrical polarization over the whole temperature range since they are probably antiferroelectrics due to a different rock salt structure. More details can be found in (Wang et al., 2009) and the references within it.

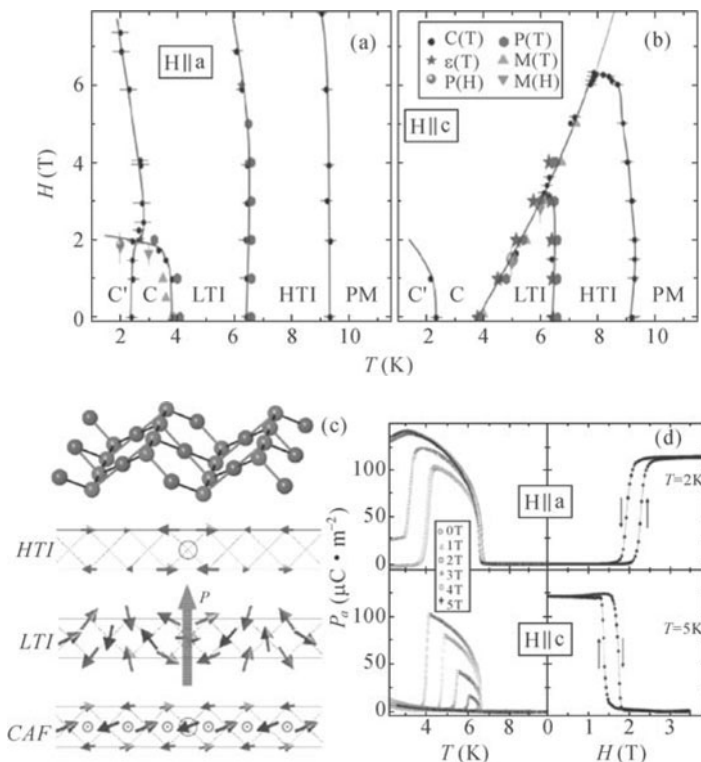


Fig. 8.14 Phase diagrams of magnetic field against temperature for $\text{Ni}_3\text{V}_2\text{O}_8$ under magnetic field along the a -axis (a) and c -axis (b); (c) Lattice structure and three spin arrangements in Ni sublattice; (d) Polarization along the b -axis as a function of temperature and magnetic field applied along the a -axis and c -axis, respectively (Jang et al., 2008). Copyright (2008) by the American Physics Society

Another typical 3D frustrated spin system is the perovskite manganite TbMnO_3 . At room temperature, TbMnO_3 is orthorhombic with space group $Pbnm$. The onset of ferroelectricity in TbMnO_3 is clearly related to the appearance of spiral magnetic ordering at $\sim 28\text{ K}$ (Kenzelmann et al., 2005). In the applied magnetic fields, TbMnO_3 shows a spin-flop transition, at which the polarization vector rotates by 90° (see Fig. 8.15). Theoretical investigation predicted that the

Jahn-Teller distortion, together with the relatively weak next-nearest-neighbour superexchange coupling in perovskite multiferroic manganites is shown to be essential for the spiral spin order (Kimura et al., 2003).

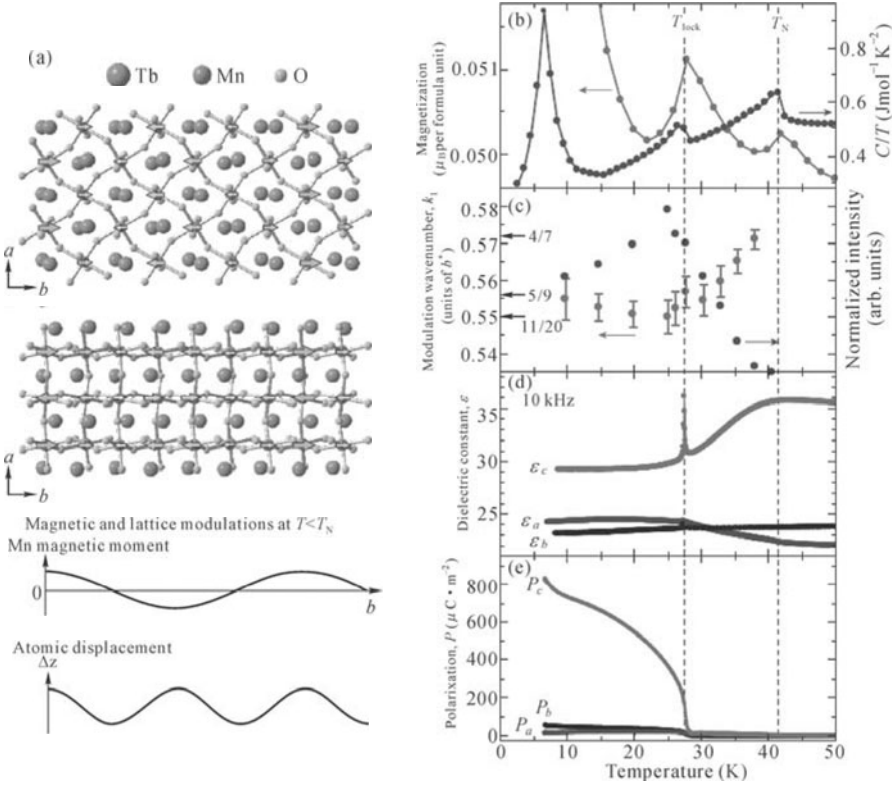


Fig. 8.15 Appearance of ferroelectricity below a lock-in transition temperature in TbMnO₃. (a) Rough sketches of crystal structure, spatial variation of Mn magnetic moment, and atomic displacement below T_N ; (b) Temperature profiles of magnetization and specific heat divided by temperature C/T ; (c) Wavenumber of lattice modulation; (d) Dielectric constant ϵ at 10 kHz; (e) Electric polarization P along the principal axes (Kimura et al., 2003). Copyright (2003) by the Nature Publishing Group

As stated in the above experimental results, the ferromagnetism seems to be very weak since no spontaneous magnetization is available due to the helical or spiral spin order. One way to avoid this problem is to form a conical spin state, in which the spontaneous component $S_{\parallel}$ (homogeneous ferromagnetic part) and spiral component of the magnetization coexist. An example of the conical spin structure is chromite spinel CoCr₂O₄ as shown in Fig. 8.16 (Yamasaki et al., 2006). If the spiral component lies in the (e_1, e_2) plane and $S_{\parallel}$ points to the e_3 -axis, the spin moment is $S_n = S_1 e_1 \cos(\mathbf{Q} \cdot \mathbf{r}) + S_2 e_2 \sin(\mathbf{Q} \cdot \mathbf{r}) + S_{\parallel} e_3$ where $\mathbf{Q}$ is the wave

vector and r is the space coordinate. More details can be found in (Kimura et al., 2006; Yamasaki et al., 2006; Aguilar et al., 2008).

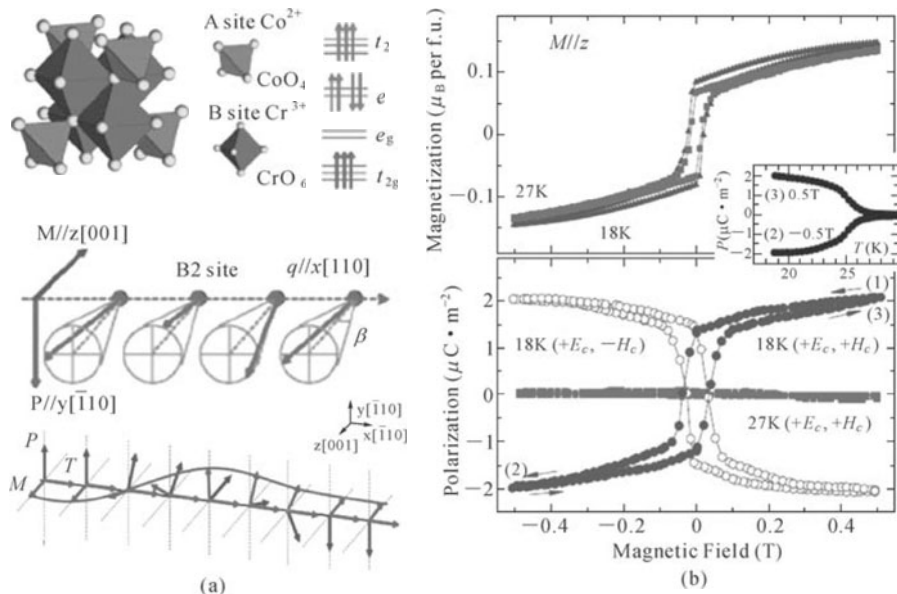


Fig. 8.16 (a) Up: Lattice structure and electronic configurations of CoCr_2O_4 . Bottom: The spin canting and the synchronized reversal of the polarization at 18 K with change of the external magnetic field; (b) Magnetic-field dependence of magnetization and electric polarization at temperatures above (27 K) and below (18 K) the ferroelectric transition temperature ($T_s=25$ K) (Yamasaki et al., 2006). Copyright (2006) by the American Physics Society

8.2.2 Magnetolectric Composites

8.2.2.1 Piezoelectric-Magnetostrictive Ceramics

Since the coupling effect of the single-phase materials is still too small and usually exists in low temperatures, an alternative route is to make multiferroic composites consisting of piezoelectric materials and magnetostrictive materials. The elastic interaction plays a very important role in such systems. The coupling typically refers to the linear magnetoelectric effect, which can be described as the product effect of the piezoelectric and magnetostrictive effect, i.e., an applied electric field creates a piezoelectric strain in the ferroelectrics, which produces a corresponding strain in the ferri-/ferro-magnetics and a subsequent magnetic property change by piezomagnetic effect, and vice versa. Following van Suchtelen (1972), who first proposed the concept of a product property in two-phase composite materials, composites combining magnetostrictive and piezoelectric phases were developed

and a large ME effect ($\frac{dE}{dH} = 0.13 \text{ V/cm Oe}$), which is a hundred times larger than that in single-phase materials, in ceramic composites of $\text{BaTiO}_3\text{--CoFe}_2\text{O}_4$, was observed in the Philips Laboratory (van den Boomgaard et al., 1974; Boomgaard et al., 1976; van den Boomgaard and Born, 1978). However, the used unidirectional solidifications need to be well in control of the process, and were not widely used at that time. Then, in the early 1990s, Newnham's group (Harshe et al., 1993) and Russian scientists (Lopatin et al., 1994) prepared particulate ceramic composites of ferrites and BaTiO_3 or $\text{Pb}(\text{ZrTi})\text{O}_3$ (PZT) by a conventional sintering processing. The sintered ceramic composites were much easier to fabricate in comparison to eutectic composites, which provided the opportunity to combine phases with widely different crystal structures even though these sintered ceramic composites had lower ME coefficients than the prior eutectic composites of Philips.

An upsurge in multiferroic ME composite research appeared in the early 2000s (Ryu et al., 2002). Various ceramic composites with different connectivity schemes including 0-3 type particulate composites and 2-2 type laminate composites have been reported. The ME effect in composite materials is known as a product tensor property (Nan, 1994) which results from the interaction between different orderings of the two phases in the composite:

$$\begin{aligned} \text{ME}_H \text{ effect} &= \frac{\text{magnetic}}{\text{mechanical}} \cdot \frac{\text{mechanical}}{\text{electric}}, \\ \text{ME}_E \text{ effect} &= \frac{\text{electric}}{\text{mechanical}} \cdot \frac{\text{mechanical}}{\text{magnetic}}. \end{aligned} \quad (8.2.6)$$

This is a coupled electrical and magnetic phenomenon via elastic interaction. So composites with large piezoelectric and magnetostrictive properties are needed. Lots of materials were used as subsystems which are summarized in Table 8.3.

Table 8.3 Materials used as sub-phases in ME composites

Piezoelectric/ferroelectric phase	Magnetic/magnetostrictive phase
$\text{Pb}(\text{Zr,Ti})\text{O}_3$	TbDyFe_2 (Terfenol-D)
$\text{Pb}(\text{Mg}_{0.33}\text{Nb}_{0.67})\text{O}_3\text{--PbTiO}_3$ (PMN-PT)	NiFe_2O_4 , CoFe_2O_4 , LiFe_2O_4
BaTiO_3	$\text{Ni}_{0.8}\text{Zn}_{0.2}\text{Fe}_2\text{O}_4$
BZT	$\text{La}_{0.7}\text{Sr}_{0.3}\text{MnO}_3$, $\text{La}_{0.7}\text{Ca}_{0.3}\text{MnO}_3$
$(\text{Sr,Ba})\text{Nb}_2\text{O}_5$	Yttrium iron garnet (YIG)

Using the concept of phase connectivity introduced by Newnham et al. (1978), the structures of a two-phase composite can be described using the notations 0-3, 2-2, 1-3, etc. (as shown in Fig. 8.17), in which each number denotes the connectivity of the respective phase. In view of the easy fabrication and flexibility to apply external fields, many kinds of magnetoelectric composites are in the form of 0-3, 2-2 and 1-3 type structures (Nan et al., 2008). The advantages of 2-2

connectivity include easy manipulation and large ME coupling effect, etc. The maximum ME coefficient in the 2-2 type laminate ceramic composites is several hundreds mV/cmOe at room temperature and reaches 90 V/cm Oe in terfenol-PZT laminates at electromechanical resonance frequency (Laletsin et al., 2004).

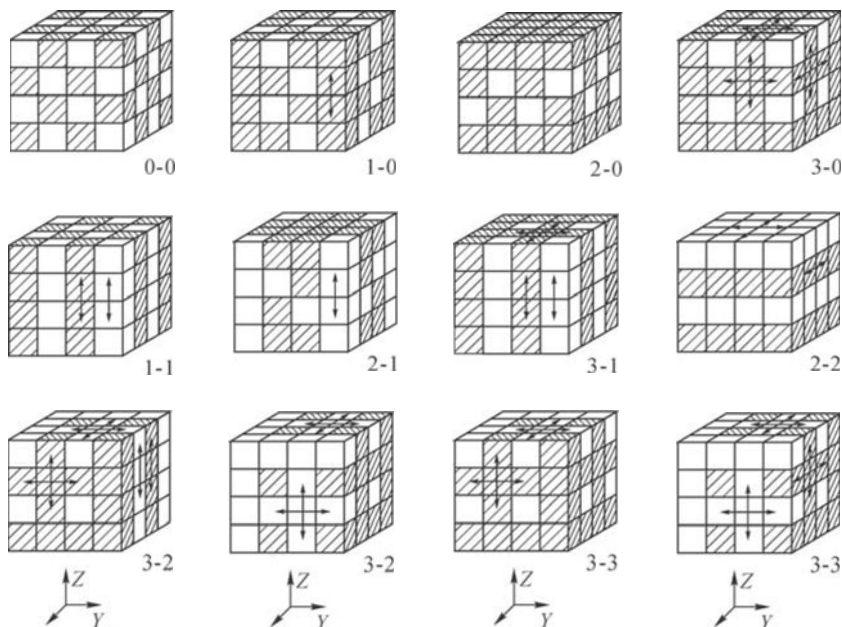


Fig. 8.17 Connectivity of composites (Newnham et al., 1978). Copyright (1978) Elsevier B.V.

Although bulk composites still face many challenges, they are the most promising type for applications because of the significant ME coupling effect. Other kinds of magnetoelectric composites including co-sintered granular composites and column-structure composites were also developed (Nan et al., 2008). Recently, in order to achieve sufficient bulk density while avoiding possible reactions between the constituent phases, hot pressing (Srinivasan et al., 2004) and spark plasma sintering (Jiang et al., 2007) techniques have been employed to replace the conventional sintering. In the composite ceramics synthesized by direct mechanical mixing of the two-phase powders, the randomly mixed ferrite particles have a low percolation threshold. In order to achieve good dispersion of a high concentration of ferrite particles in the composite ceramics, wet chemical processing (e.g., sol-gel method) has recently been employed to synthesize the homogeneously mixed composite powders of the piezoelectric and ferrite oxides (Ren et al., 2005; Wan et al., 2006; Weng et al., 2007). The *in-situ* synthesized samples showed a piezoelectric d_{33} constant approximately six times larger and an ME voltage coefficient approximately three times larger than these of the corresponding mechanically milled samples. Some of the magnetoelectric

voltage coefficients are summarized in Table 8.4. It should be noted that the SI unit for the voltage magnetoelectric coefficient is V/A. However, most works in recent years provide α_E in V/cm Oe. For conversion purposes, 1 V/cm Oe=1.257 V/m.

Table 8.4 Magnetoelectric coupling coefficients from experiments (1 V/cm Oe=1.257 V/m)

Materials and structures	ME coupling coefficients
MnZnFe ₂ O ₄ /Tb _{1-x} Dy _x Fe _{2-y} /Pb, ZrTiO ₃	3 V/cm Oe (Dong et al., 2003)
Layered PZT/Terfenol-D	2 V/cm Oe (Ryu et al., 2002)
Layered PZT/Permendur	0.8 V/cm Oe (Folen et al., 1961)
Layered PZT/Fe-Ga	0.4 V/cm Oe (Nan et al., 2001)
Layered PZT/NiFe ₂ O ₄	0.45 V/cm Oe (Dong et al., 2005)
0-3 type Ni _{0.8} Zn _{0.2} Fe ₂ O ₄ /Sr _{0.5} Ba _{0.5} Nb ₂ O ₆	0.0266 V/cm Oe (Li et al., 2006)
Bilayer Ni _{0.8} Zn _{0.2} Fe ₂ O ₄ -PZT@275.5 kHz	2.916 V/cm Oe (Zhang et al., 2007)
Trilayer Ni _{0.8} Zn _{0.2} Fe ₂ O ₄ -PZT@339.4 kHz	8.408 V/cm Oe (Ryu et al., 2002)
Bilayer Ni ₂ MnGa/PbZr _{0.52} Ti _{0.48} O ₃	0.0266 V/cm Oe (Zhao et al., 2006)
Layered Bi _{3.15} Nd _{0.85} Ti ₃ O ₁₂ (BNT)-CoFe ₂ O ₄	0.0345 V/cm Oe (Zhong et al., 2007)
Layered Sr _{0.5} Ba _{0.5} Nb ₂ O ₆ -CoFe ₂ O ₄	0.025 V/cm Oe (Panchapakesan et al., 2004)
Layered terfenol-PZT@330 kHz	90 V/cm Oe (Laletsin et al., 2004)

8.2.2.2 Multiferroic Nanostructures

The renaissance of multiferroics has been driven, in part, by the development of thin film growth techniques that allow for the production of non-equilibrium phases of materials and strain engineering of existing materials. The breakthrough in terms of the giant magnetoelectric effect in composites was achieved in nano-heterostructures reported by Zheng et al. (2004). They used the pulse-laser deposition technique to fabricate BaTiO₃-CoFe₂O₄ nanostructures on a rigid SrTiO₃ substrate and found that the magnetoelectric coupling in such nanostructure can be understood on the basis of the strong elastic interactions between the two phases. In the following years, many samples with different structures have been fabricated and large ME effects were achieved. In the past few years, many kinds of multiferroic materials have appeared, including nanolayers, superlattices, nanoheterostructures and ultra-thin films. New physical phenomena at artificially constructed heterointerfaces were observed and present an exciting area of condensed-matter science.

One issue that should be mentioned is that the trend towards miniaturization and integration of devices makes the length-scale reduction more and more interesting. On the nano scale, efficient design is vital. It will be crucial to find an optimum stress/strain distribution in the film in order to create the required interaction for generating multiferroism. Thus, the multiferroic properties in the nano-scale heterostructures and superlattices are dependent on the details of the FE and FM layers. Recent works focus on the layered films in which one phase is used as a substrate and the other phase is the film (Eerenstein et al., 2006; Wu et al., 2006). Typically, two applications were developed: one uses the electric field to control the magnetic properties and the other uses the magnetic field to control

the electric polarization properties. In particular, it is important to use the electric field as a control field because a voltage control can generally be designed with greater simplicity, efficiency and robustness than a magnetic control, and is less prone to generate cross interaction and large noise. These features are attractive for a range of applications, such as high-sensitive field sensors, voltage-modulated filters and transducers, and data storage systems.

• *Nanolayers and Superlattices*

With the innovation of nanotechnology and nano-mechanical testing techniques, multiferroic nanostructures of high quality were successfully fabricated, which led to a blossoming of miniaturized multifunctional devices (Ramesh and Spaldin, 2007). Motivated by on-chip integration in microelectronic devices, nanostructured composites of ferroelectric and magnetic oxides have recently been fabricated. The elastic coupling interaction between nanosized ferroelectric and magnetic oxides is also responsible for the magnetoelectric effect in the nanostructures. The availability of high-quality nanostructured composites makes it easier to tailor their properties through epitaxial strain, atomic-level engineering of chemistry and interfacial coupling.

Thin film heterostructures made by combining ferroelectrics such as $\text{Ba}_{0.6}\text{Sr}_{0.4}\text{TiO}_3$, BaTiO_3 , and PMN-PT with the ferromagnet $\text{Pr}_{0.85}\text{Ca}_{0.15}\text{MnO}_3$ and Tb-Fe/Fe-Co multilayers were studied. The advantages of this 2-2 connectivity are that it is easy to manipulate and has great application potential. Unfortunately, these attempts were unable to achieve large magnetoelectric coefficients because of the in-plane clamping effect induced by the rigid substrate (Zheng et al., 2004).

To avoid such clamping effect in the layered structure, an efficient way is to use FE or FM materials as a substrate and the other phase as the film deposited on the substrate. Stein et al. (2005) deposited 4 μm $\text{Tb}_{40}\text{Fe}_{60}/\text{Fe}_{50}\text{Co}_{50}$ alloy on 100 μm PMN-PT single-crystal substrate and got a maximum ME voltage coefficient of about 24 mV/cm Oe. Wu et al. (2006) grew a 110 nm PZT thin film on $\text{La}_{1.2}\text{Sr}_{1.8}\text{Mn}_2\text{O}_7$ single-crystal substrate. The ME effect was significant at the LSMO's phase transition temperature 120 K and reached 600 mV/cm Oe. Eerenstein et al. (2006) studied the LSMO/ BaTiO_3 heterostructure and observed an electric field induced sharp nonreversible decrease in magnetization, which was attributed to the change in magnetic anisotropy induced by the electric-field-controlled strain. Thiele et al. (2007) reported the electric-field-induced change in saturation magnetization and Curie temperature of manganite in the perovskite manganites/FE heterostructure. Yang et al. (2009) fabricated multiferroic heterostructures by growing the CoFe_2O_4 film on $\text{Pb}(\text{Mg}_{1/3}\text{Nb}_{2/3})_{0.7}\text{Ti}_{0.3}\text{O}_3$ substrate and found that the magnetization switched on and off alternately as a function of time with 50 V (0.83 kV/cm). Such nanostructures provide the possibility of direct ME coupling and may have wide potential applications in novel miniaturized multifunctional devices. In such structures, the ferroelectric or ferromagnetic materials served as both substrate and strain mediator. It should be mentioned that using an electric field to control the magnetic properties has many

potential applications and that the electric field induced change in magnetic anisotropy via strain plays an important role in the magnetoelectric coupling in the heterostructures.

Another type of horizontal structure is in the form of a superlattice (Hatt and Spaldin, 2007; Chaudhuri et al., 2007; Singh and Prellier, 2007). In such structures, compositional ordering that breaks inversion symmetry has produced systems with enhanced polarizations and large nonlinear optical responses because of the large epitaxial strain induced by the alternation of sublattices. Other complex interplays such as the interfacial charge transfer, chemical bonding and electrostatic screening also give rich phenomena and interfacial ME coupling effects in superlattice structures (Duan et al. 2006; Hatt and Spaldin, 2007).

• Nanoheterostructures

Monolayer self-assembled structures are believed to have strong interaction between the two constituent phases, thus enhancing the ME effect. In particular, vertically aligned structures might overcome substrate clamping effects. Numerous investigations of similar systems were reported, such as those with different perovskite (BTO, BFO, PbTiO_3 (PTO), and PZT) and spinel (CFO and NFO) phases, different spinel/perovskite or S/P atomic ratios (1/2, 1/1, and 2/1) (Qiao and Bi, 2008). To improve their piezoelectric properties, perovskite-spinel self-assembled epitaxial thin films have also been deposited onto various substrates with different lattice parameters (Table 8.5) and different substrate orientations.

Table 8.5 Lattice parameters of materials (unit: Å)

Material	LaAlO_3	SrTiO_3	MnO	MgAl_2O_4	BaTiO_3	PZT	CoFe_2O_4
Lattice parameter	3.792	3.905	4.21	8.08	3.9998	~4.1	8.69

It is well known that the misfit stress induced by the substrate can strongly affect the properties of ferroelectric (FE) and ferromagnetic (FM) materials on the nanoscale. By adjusting the stress induced by the substrate, properties of the film can be manipulated. Good connectivity and large contact areas in 1-3 type nanoheterostructures may result in a significant coupling effect because of the elastic interaction in both the interface of the two portions along the vertical direction and the interface of film/substrate (Zheng et al., 2004; Yan et al., 2007; MacManus-Driscoll et al., 2008). The stress distribution in such structures is very complex because of the existence of coupling effects, different thermal expansion coefficients, the dislocations, etc. Also, many factors, such as the lattice parameters of each part, structure, elastic properties, electrostrictive and magnetostrictive properties can also greatly affect the stress state. In a perfectly vertical epitaxial multiferroic composite thin film, the misfit stress induced by the FE/FM film/substrate interfaces is a key factor in the ME coupling effect (Lu et al., 2007; Deng et al., 2008). Stress fields in the composite also have a significant

effect on the phase morphologies of the deposited film (Slutsker et al., 2006). So an optimum design of the materials with proper stress distribution is vital to the achievement of a large ME coupling effect.

8.3 Theories of Magnetoelectric Coupling Effect at Low Frequency

8.3.1 *Energy Formulation for Multiferroic Composites*

Although experimental studies of these ME composites in the 1990s did not represent a great step forward, the experiments did inspire significant theoretical work on ME ceramic composites. A few models, such as Green's function approach and various micromechanics approaches, were developed to understand the coupling between the two ferroic phases and to predict the resultant ME tensor responses in bulk ceramic composites. The early analysis of the ME effect in bulk ceramic is just conceptually straightforward, providing little physical insight. More rigorous treatment of the ME behavior of the composites has been performed lately. The ME voltage coefficient can be obtained by solving the magnetostatic and electrostatic equations. The magnetoelectric composites were theoretically studied and the expressions for the effective magnetoelectric coefficients were obtained (Harshe et al., 1993a; 1993b). Later on, micromechanics models were used to estimate the effective properties of piezoelectric/piezomagnetic composite materials (Du and Wang, 1998; Nan, 1994; Huang et al., 2009). By using some micromechanics models, as shown in Chapter 3, some equations were obtained for the effective magnetoelastoelectric moduli of fibrous composites (Benveniste, 1995; Li and Dunn, 1998). The closed-form expressions for the effective moduli of fibrous and laminated composites had been derived (Li, 2000).

The above models mainly focus on the effective properties of the composite, especially the ME coupling coefficient. However, the linear constitutive equations with respect to the polarization, magnetization and external fields cannot predict the nonlinear properties (Slutsker et al., 2007). In general, to study the nonlinear behavior of the materials, one needs to establish the free energy expressions. Using the micromechanics model together with the Landau free energy expansion, Lu et al. (2007) built a comprehensive model for 1-3 type composites with consideration of the lattice mismatch and the surface and size effect. Critical behaviors were investigated for the multilayered system of alternating FE and FM thin films on a substrate of finite thickness (Wang and Woo, 2008). Phase field models were also developed to predict the ME coupling in a nanocomposite thin film (Zhang et al., 2007; Wu et al., 2010). Analytical evolution fields of the spontaneous polarization, magnetization and stress were obtained with the use of a generalized Eshelby model and the Fourier transformation (Ni et al., 2010). However, finite size may be an obstacle in the use of the Fourier transformation.

In fact, the sets of the physical field equations of such heterogeneous

magnetoelectroelastic solids under applied loadings are easy to establish but not easy to solve analytically because of the position-dependent constitutive coefficients (Ni et al., 2010). Thus, the estimates of the effective ME coefficients of the ME composites are needed, and usually can be obtained by using various mean field approximations.

With the stresses of the inclusions and matrix solved by using the classical elastic Eshelby solutions, the total free energy can be established for the composites, as shown in Chapter 3. Based on the phenomenological theory, the coupling effect and physical properties of nano-structured multiferroic composites can be studied by using the time-dependent Ginzburg-Landau theory.

Suppose that the FE thin film with imbedded FM inclusions was grown on a rigid substrate under high temperature. The film is in-plane infinite and homogenous on a large scale. That is, the homogeneous body with uniformly distributed inclusions is under a macroscopic stain-controlled boundary condition. At the fabrication temperature, above both the phase transition temperatures of ferroelectric and ferromagnetic phases, no phase transition occurs and both phases are in their para-states. This homogeneous state under high temperature is named the prototype phase. FE phase transition and FM phase transition occurred when the temperature was cooled down below the phase transition points. Correspondingly, FE transformation strain $\varepsilon_{ij}^P = Q_{ijkl} P_k P_l$ and FM phase transition strain $\varepsilon_{ij}^M = \lambda_{ijkl} M_k M_l / M_s^2$ were induced, where Q_{ijkl} is the electrostrictive coefficient, P_k is the spontaneous polarization vector, λ_{ijkl} is the magnetostrictive coefficient, M_k is the spontaneous magnetization vector, and M_s is the saturation magnetization. The notation of tensor quantity is bold-faced and used only where necessary. It should be noted that the thermal strain induced during the fabrication is ignored (Zheng et al., 2004).

Under applied initial strain induced by the substrate, the total free energy for the thin film composed of the intrinsic Landau free energies of the separated FE and FM portions, the coupling elastic energy and the surface energy (Lu et al., 2009) is:

$$F = (1-f) \int_v F_L^E dv + f \int_v F_L^M dv + \int_v F_{Elas} dv + F_{Surf}, \quad (8.3.1)$$

where f is the volume fraction of the FM phase, F_L^E and F_L^M are the Landau type bulk free energy densities of the FE and FM phases, respectively, with consideration of the gradient items induced by the inhomogeneous polarization and magnetization due to the surface and interface. F_{Surf} is the total surface energy and F_{Elas} is the elastic coupling energy density.

Under mechanical stress-free and free-charge electric boundary conditions, the intrinsic Landau-Ginzburg free energy for the cubic FE portion can be expressed as (Chen et al., 2003):

$$\begin{aligned}
F_L^E = & \alpha_1(P_1^2 + P_2^2 + P_3^2) + \alpha_{11}(P_1^4 + P_2^4 + P_3^4) + \alpha_{12}(P_1^2 P_2^2 + P_1^2 P_3^2 + P_2^2 P_3^2) + \\
& \alpha_{111}(P_1^6 + P_2^6 + P_3^6) + \alpha_{112}[P_1^4(P_2^2 + P_3^2) + P_2^4(P_1^2 + P_3^2) + P_3^4(P_1^2 + P_2^2)] \\
& + \alpha_{123}P_1^2 P_2^2 P_3^2 + D_{11}^P(P_{1,1}^2 + P_{2,2}^2 + P_{3,3}^2) + D_{44}^P(P_{1,2}^2 + P_{1,3}^2 + P_{2,1}^2 + P_{2,3}^2 + P_{3,1}^2 + P_{3,2}^2),
\end{aligned} \tag{8.3.2}$$

where α_i, α_{ij} and α_{ijk} are the expansion coefficients of the Landau free energy. D_{11}^P and D_{44}^P are the corresponding gradient coefficients.

Similar to the above description of dielectrics, the magnetic energy density can be expanded up to the 4-th order for the second-order cubic FM material with consideration of the magnetization gradients:

$$\begin{aligned}
F_L^M = & \beta_1(M_1^2 + M_2^2 + M_3^2) + \beta_{11}(M_1^4 + M_2^4 + M_3^4) \\
& + \frac{K_1}{M_s^4}(M_1^2 M_2^2 + M_1^2 M_3^2 + M_2^2 M_3^2) + \frac{K_2}{M_s^6} M_1^2 M_2^2 M_3^2 \\
& + D_{11}^M(M_{1,1}^2 + M_{2,2}^2 + M_{3,3}^2) + D_{44}^M(M_{1,2}^2 + M_{1,3}^2 + M_{2,1}^2 + M_{2,3}^2 + M_{3,1}^2 + M_{3,2}^2),
\end{aligned} \tag{8.3.3}$$

where β_1 and β_{11} are the expansion parameters for bulk FM material, K_1 and K_2 are anisotropic coefficients related with the anisotropic energy, and D_{11}^M and D_{44}^M are the corresponding gradient coefficients.

The surface energy F_{Surf} describing the relaxation or restriction energy of the FE and FM parts at the surface and the interface can be written as

$$F_{\text{Surf}} = \int_{s_p} (D_{11}^P \frac{P_i^2}{2\delta_i^P} + \sum_{i,j} D_{44}^P \frac{P_i^2}{2\delta_j^P}) ds + \int_{s_M} (D_{11}^M \frac{M_i^2}{2\delta_i^M} + \sum_{i,j} D_{44}^M \frac{M_i^2}{2\delta_j^M}) ds, \tag{8.3.4}$$

where, s_p and s_M are the surface areas of the FE and FM phase, respectively. δ_i^P and δ_i^M ($i=1,2$) are the extrapolation lengths for the FE and FM phases, respectively. We employ the usual suffix mutation. A repeated suffix is summed over the values 1, 2, 3 and suffixes preceded by a comma denote differentiation. The extrapolation lengths describe the difference of the surface and the bulk (Cottam et al., 1984; Qu et al., 1997).

At high temperature, no phase transition occurred and the elastic fields both in the matrix and the inclusions can be solved by using Eshelby's results for elastic media. As assumed, the composite is homogeneous on a macroscopic scale, and the phases themselves are isotropic as well as uniform. Since their elastic moduli differ, the stress and strain fields throughout the mixture are microscopically inhomogeneous under any external loading. When the temperature was cooled

down below the transition points, phase transitions occurred and Eshelby's solution should be expanded with consideration of the transformation strains. However, the elastic Eshelby solution can still be used for simplicity. When one considers the film on a substrate, the film, which is clamped by the substrate because of the lattice mismatch between the substrate and film, will be subjected to uniform mismatch strains $\varepsilon^0 = \varepsilon_{11}^0 = \varepsilon_{22}^0$ during the fabrication process (Lu et al., 2007). The average elastic strains in the matrix and inclusions can be artificially derived by using the Mori-Tanaka method (Benveniste, 1987).

The total coupled elastic energy density F_{Elas} can be finally calculated by combining the elastic energy in each portion:

$$F_{\text{Elas}} = \frac{f}{2} \langle \sigma_{ij}^I \rangle \langle \varepsilon_{ij}^I \rangle + \frac{1-f}{2} \langle \sigma_{ij} \rangle \langle \varepsilon_{ij} \rangle, \quad (8.3.5)$$

where $\langle \sigma_{ij}^I \rangle$, $\langle \varepsilon_{ij}^I \rangle$ and $\langle \sigma_{ij} \rangle$, $\langle \varepsilon_{ij} \rangle$ are average stress, strain in the inclusion and matrix, respectively.

8.3.2 Phase Transition Behaviors in Layered Structures

With a surge in the study of the coupling effect in layered structures, phase transition behavior is also becoming of keen interest. Here, we first consider a multilayer thin-film system with alternating ferroelectric (FE) and ferromagnetic (FM) layers on a substrate of finite thickness H . The FE and FM thicknesses are h_e and h_m , respectively (Fig. 8.18).

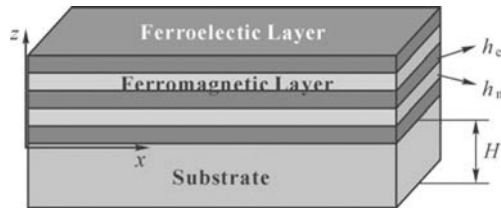


Fig. 8.18 Schematic of layered structure

The behavior of the ferroelectric layers is described in terms of self-polarization (to be distinguished from spontaneous polarization) as an order parameter. Without loss of generality, one can consider only the component P of the spontaneous polarization perpendicular to the film surface and M along the x -direction in the film plane. The total free energy is the summation of the Landau free energies for the FE and FM phases, the elastic coupling energy and the surface energy, and can be written as

$$\Phi = F_l^E + F_l^M + F_{\text{surf}} + F_{\text{elas}}. \quad (8.3.6)$$

F_{elas} is the coupled elastic energy induced by the elastic interaction of the misfit between the respective film and the substrate, and eigenstrains associated with the ferroelectric and ferromagnetic transformations.

$$\begin{aligned} F_{\text{elas}} &= F_{\text{elas}}^E + F_{\text{elas}}^M \\ &= n_e C^E \int_0^{h_e} (\boldsymbol{\varepsilon}^{0E} - \boldsymbol{\varepsilon}^{\text{Et}} + \boldsymbol{\varepsilon}^S)^2 dz + n_m C^M \int_{h_e}^{h_e+h_m} (\boldsymbol{\varepsilon}^{0M} - \boldsymbol{\varepsilon}^{\text{Mt}} + \boldsymbol{\varepsilon}^S)^2 dz + H C^0 (\boldsymbol{\varepsilon}^S)^2, \end{aligned} \quad (8.3.7)$$

where $\boldsymbol{\varepsilon}^{0E}$, $\boldsymbol{\varepsilon}^{0M}$ and $\boldsymbol{\varepsilon}^S$ are the misfit strains in the ferroelectric, ferromagnetic and substrate layers, respectively. C^E , C^M and C^0 are the corresponding elastic moduli. $\boldsymbol{\varepsilon}^{\text{Et}}$ and $\boldsymbol{\varepsilon}^{\text{Mt}}$ are the eigenstrains of ferroelectric and ferromagnetic transformations, respectively.

The transformation eigenstrains can be written in terms of the order parameters P and M via the electrostrictive and magnetostrictive (volumetric plus anisotropic) coefficients $\mathbf{Q}$ and $\mathbf{Q}_m$, using the relations $\boldsymbol{\varepsilon}^{\text{Et}}(z) = \mathbf{Q}P^2(z)$, and $\boldsymbol{\varepsilon}^{\text{Mt}}(z) = \mathbf{Q}_m M^2(z)$, respectively.

Denoting the non-zero components of the misfit strains along the x - and y -direction by ε_{11}^{0E} , ε_{22}^{0E} , ε_{11}^{0M} , ε_{22}^{0M} respectively and assuming that the multilayer system is grown hetero-epitaxially from a substrate, $\varepsilon_{\alpha\alpha}^{0E}$ and $\varepsilon_{\alpha\alpha}^{0M}$ can be expressed as usual in terms of the lattice constants of the films a_{α}^{Ef} and a_{α}^{Mf} and substrate a_{α}^s :

$$\varepsilon_{\alpha\alpha}^{0E} = \frac{a_{\alpha}^s - a_{\alpha}^{\text{Ef}}}{a_{\alpha}^{\text{Ef}}}, \quad \varepsilon_{\alpha\alpha}^{0M} = \frac{a_{\alpha}^s - a_{\alpha}^{\text{Mf}}}{a_{\alpha}^{\text{Mf}}}. \quad (8.3.8)$$

The misfit strains $\boldsymbol{\varepsilon}^{0E}$, $\boldsymbol{\varepsilon}^{0M}$ and $\boldsymbol{\varepsilon}^S$ in Eq. (8.3.8) can be calculated by adopting Timoshenko's method for thermal stresses. Thus, the integrity and coherency conditions of the multilayer system require stresses $\sigma_{ij}^E = C_{ijkl}^E (\varepsilon_{kl}^{0E} - \varepsilon_{kl}^{\text{Et}})$ and $\sigma_{ij}^M = C_{ijkl}^M (\varepsilon_{kl}^{0M} - \varepsilon_{kl}^{\text{Mt}})$ on the ferroelectric and ferromagnetic layers, respectively. If we assume that the in-plane is infinite and there is no torsion in the in-plane, the stresses will be simplified as $\sigma^E = \sigma_{11}^E = \sigma_{22}^E$ and $\sigma^M = \sigma_{11}^M = \sigma_{22}^M$. These stresses create a resultant force at the ends of the plates, given by

$$\mathbf{F} = n_e \int_{-h_e/2}^{h_e/2} \sigma^E dz + n_m \int_{-h_m/2}^{h_m/2} \sigma^M dz. \quad (8.3.9)$$

For the entire system to be in a stress-free state, a force equal and opposite to $\mathbf{F}$ must be applied. If the system is not symmetrical, the balancing force will introduce a bending moment. To simplify our analysis, the system is assumed to be symmetrical and the bending moment can be neglected. $\boldsymbol{\varepsilon}^S$ is then the uniform strain produced in the substrate by the balancing force. At final equilibrium the strain $\boldsymbol{\varepsilon}^S$ will be applied to the entire system (i.e., including the substrate). Then, the corresponding balancing equation can be written as

$$n_c h_c \boldsymbol{\sigma}^{Eb} + n_m h_m \boldsymbol{\sigma}^{Mb} + H \boldsymbol{\sigma}^{Sb} = -\mathbf{F}, \quad (8.3.10)$$

with

$$\boldsymbol{\sigma}^{Eb} = \mathbf{C}^E \boldsymbol{\varepsilon}^S, \boldsymbol{\sigma}^{Mb} = \mathbf{C}^M \boldsymbol{\varepsilon}^S, \boldsymbol{\sigma}^{Sb} = \mathbf{C}^0 \boldsymbol{\varepsilon}^S, \quad (8.3.11)$$

where the indices α, i, j, k, l take values of 1 or 2. In terms of $\boldsymbol{\varepsilon}^S$, Eqs. (8.3.10) and (8.3.11) can be written as

$$(n_c h_c \mathbf{C}^E + n_m h_m \mathbf{C}^M + H \mathbf{C}^0) \boldsymbol{\varepsilon}^S = -\mathbf{F}, \quad (8.3.12)$$

from which $\boldsymbol{\varepsilon}^S$ is linearly related to $\mathbf{F}$ via a constant system compliance $\mathbf{A}$:

$$\boldsymbol{\varepsilon}^S = -\mathbf{A} \mathbf{F} \text{ with } \mathbf{A} = (n_c h_c \mathbf{C}^E + n_m h_m \mathbf{C}^M + H \mathbf{C}^0)^{-1}. \quad (8.3.13)$$

If we define a force-constant matrix $\mathbf{K}$ by $K_{\alpha\beta} = n_c h_c C_{\alpha\beta}^E + n_m h_m C_{\alpha\beta}^M + H C_{\alpha\beta}^0$, both $\mathbf{K}$ and $\mathbf{A}$ are real and symmetric. As an example, in the particular case of isotropic elasticity

$$\begin{pmatrix} \varepsilon_1^S \\ \varepsilon_2^S \end{pmatrix} = \frac{-1}{\|\mathbf{K}\|} \begin{pmatrix} K_{11} & -K_{12} \\ -K_{12} & K_{11} \end{pmatrix} \begin{pmatrix} F_1 \\ F_2 \end{pmatrix}. \quad (8.3.14)$$

Thus, from Eqs. (8.3.9), (8.3.10) and (8.3.13), we can write

$$\begin{aligned} \boldsymbol{\varepsilon}^S &= -n_c \mathbf{A} \int_{-h_c/2}^{h_c/2} \boldsymbol{\sigma}^E dz - n_m \mathbf{A} \int_{-h_m/2}^{h_m/2} \boldsymbol{\sigma}^M dz \\ &= -n_c \mathbf{A} \mathbf{C}^E \int_{-h_c/2}^{h_c/2} \boldsymbol{\varepsilon}^{0E} - \boldsymbol{\varepsilon}^{Et} dz - n_m \mathbf{A} \mathbf{C}^M \int_{-h_m/2}^{h_m/2} (\boldsymbol{\varepsilon}^{0M} - \boldsymbol{\varepsilon}^{Mt}) dz. \end{aligned} \quad (8.3.15)$$

Substitution of Eq. (8.3.15) into Eq. (8.3.7) yields an interaction term between the ferroelectric and ferromagnetic layers through the substrate, which produces a magneto-electric coupling.

With the expirations of Landau free energy for FE and FM phases, the total energy can be derived. The time-dependent dynamic equations of the FE and FM phases can be written as the variation of the total free energy with respect to the polarization and magnetization, respectively,

$$\begin{aligned}
\frac{\partial P}{\partial t} &= -\mu_c \frac{\delta \Phi}{\delta P} \\
&= -\mu_c \left[A^E P + \frac{n_c}{\varepsilon} P + B^E P^3 + n_c C^E P^5 - n_c D^E \frac{\partial^2 P}{\partial z^2} \right], \\
\frac{\partial M}{\partial t} &= -\mu_m \frac{\delta \Phi}{\delta M} = -\mu_m [A^M M + B^M M^3 - n_m D_m \frac{\partial^2 M}{\partial z^2}],
\end{aligned} \tag{8.3.16}$$

where, μ_c and μ_m are the kinetic coefficients related to the FE and FM domain wall mobilities, respectively, and

$$\begin{aligned}
A^E &= n_c \left[A_c (T - T_{c0}^e) + \alpha_0^E + \alpha_p^E \langle P^2 \rangle + \alpha_m^E \langle M^2 \rangle \right], \\
B^E &= n_c (B_c + \beta^E), \\
A^M &= n_m \left[A_m (T - T_{c0}^m) + \alpha_0^M + \alpha_p^M \langle P^2 \rangle + \alpha_m^M \langle M^2 \rangle \right], \\
B^M &= n_m (B_m + \beta^M).
\end{aligned} \tag{8.3.17}$$

Performing functional variation on the surface terms yields the boundary conditions,

$$\begin{aligned}
\frac{\partial P}{\partial z} &= \mp \frac{P}{\delta_c}, \text{ for } z = \pm \frac{h_c}{2}, \\
\frac{\partial M}{\partial z} &= \mp \frac{M}{\delta_m}, \text{ for } z = \pm \frac{h_m}{2}.
\end{aligned} \tag{8.3.18}$$

Here, a local coordinate system has been used to simplify the formulation, in which the midpoint of each layer is chosen as the zero point of the z -axis. The stability of the stationary state (P_0, M_0) is probed by applying an infinitesimal perturbation (Δ_c, Δ_m) to the system. Retaining only terms linear in Δ_c, Δ_m , the evolution equations of the perturbations can be obtained from Eq. (8.3.16) as

$$\begin{aligned}
\frac{\partial \Delta_c}{\partial t} &= -\mu_c \left[\left(\frac{\delta A^E}{\delta P} \right) \Big|_{\substack{P=P_0 \\ M=M_0}} P_0 + A^E + \frac{n_c}{\varepsilon} P_0 + 3B^E P_0^2 + \frac{\delta B^E}{\delta P} \Big|_{\substack{P=P_0 \\ M=M_0}} P_0^3 + 6C^E P_0^5 - n_c D^E \frac{\partial^2}{\partial z^2} \right] \Delta_c \\
&\quad + \left(\frac{\delta A^E}{\delta M} \Big|_{\substack{P=P_0 \\ M=M_0}} P_0 + \frac{\delta B^E}{\delta M} \Big|_{\substack{P=P_0 \\ M=M_0}} P_0^3 \right) \Delta_m,
\end{aligned} \tag{8.3.19}$$

$$\begin{aligned} \frac{\partial \Delta_m}{\partial t} = & -\mu_m \left[\left(\frac{\delta A^M}{\delta M} \right) \Big|_{\substack{P=P_e \\ M=M_e}} M_0 + A^M + 3B^M M_0^2 + \frac{\delta B^M}{\delta M} \Big|_{\substack{P=P_e \\ M=M_e}} M_0^3 - n_m D_m \frac{\partial^2}{\partial z^2} \right) \Delta_m \\ & + \left(\frac{\delta A^M}{\delta P} \Big|_{\substack{P=P_e \\ M=M_e}} M_0 + \frac{\delta B^M}{\delta P} \Big|_{\substack{P=P_e \\ M=M_e}} M_0^3 \right) \Delta_e \Big]. \end{aligned} \quad (8.3.20)$$

Evaluating the variation derivatives from Eqs. (8.3.19) and (8.3.20), and taking into account that M_0 and P_0 are spatially segregated so that $M_0 P_0 = 0$, the evolution equation of (Δ_e, Δ_m) can be written as

$$\frac{\partial}{\partial t} \begin{pmatrix} \Delta_e \\ \Delta_m \end{pmatrix} = - \begin{pmatrix} U_e - \mu_e n_e D_e \frac{\partial^2}{\partial z^2} & 0 \\ 0 & U_m - \mu_m n_m D_m \frac{\partial^2}{\partial z^2} \end{pmatrix} \begin{pmatrix} \Delta_e \\ \Delta_m \end{pmatrix}, \quad (8.3.21)$$

where

$$\begin{aligned} U_e &= \mu_e \left[A^E + \frac{n_e}{\varepsilon} + (3B^E + 2n_e \alpha_p^E) P_0^2 + 6C^E P_0^5 \right], \\ U_m &= \mu_m \left[A^M + (3B^M + 2n_m \alpha_m^M) M_0^2 \right]. \end{aligned} \quad (8.3.22)$$

Eq. (8.3.21) is subject to the boundary conditions on the ferroelectric and ferromagnetic film surfaces $\pm h_e$ and $\pm h_m$, that is

$$\frac{\partial}{\partial z} \begin{pmatrix} \Delta_e \\ \Delta_m \end{pmatrix} = - \begin{pmatrix} \mp \delta_e^{-1} & 0 \\ 0 & \mp \delta_m^{-1} \end{pmatrix} \begin{pmatrix} \Delta_e \\ \Delta_m \end{pmatrix}. \quad (8.3.23)$$

To solve the temporal and spatial dependence of (Δ_e, Δ_m) , we separate the space and time coordinates by writing

$$\begin{pmatrix} \Delta_e \\ \Delta_m \end{pmatrix} = \exp(i\omega t) \begin{pmatrix} \varphi_e(z) \\ \varphi_m(z) \end{pmatrix}. \quad (8.3.24)$$

For a symmetric configuration of the multilayered system, the conditions $\frac{d\Delta_e}{dz} = 0$, $\frac{d\Delta_m}{dz} = 0$ at $z = 0$ must also hold. Substituting into Eq. (8.3.21), φ_e and φ_m can be solved as $\cos(k_e z)$ and $\cos(k_m z)$. ω is then governed by the following determinantal equation,

$$\begin{vmatrix} U_e + \mu_e n_e D_e k_e^2 + \omega & 0 \\ 0 & U_m + \mu_m n_m D_m k_m^2 + \omega \end{vmatrix} = 0, \quad (8.3.25)$$

which has the form $(\omega - \omega_>)(\omega - \omega_<) = 0$, in which $\omega_> = \max[\omega_e(P_0, M_0), \omega_m(P_0, M_0)]$ and $\omega_< = \min[\omega_e(P_0, M_0), \omega_m(P_0, M_0)]$ with

$$\begin{aligned} \omega_e(P_0, M_0) &\equiv -U_e - \mu_e n_e D_e k_e^2 = -\mu_e [A^E + \frac{n_e}{\varepsilon} + (3B^E + 2n_e \alpha_p^E) P_0^2 + 6C^E P_0^5 + n_e D_e k_e^2], \\ \omega_m(P_0, M_0) &\equiv -U_m - \mu_m n_m D_m k_m^2 = -\mu_m [A^M + (3B^M + 2n_m \alpha_m^M) M_0^2 + n_m D_m k_m^2]. \end{aligned} \quad (8.3.26)$$

Note here that the coupling of P_0 and M_0 comes from A^E and A^M (see Eq. (8.3.17)). The constants k_e , k_m are determined from the following pair of transcendental equations derived from the boundary conditions in Eq. (8.3.35):

$$\tan\left(\frac{k_e h_e}{2}\right) = \frac{1}{k_e \delta_e}, \quad \tan\left(\frac{k_m h_m}{2}\right) = \frac{1}{k_m \delta_m}. \quad (8.3.27)$$

To facilitate further discussion without loss of generality, we shall specify and assume that $\omega_>(0,0) = \omega_e(0,0)$; that is, the ferroelectric transformation occurs first. Then, after the first bifurcation, the system takes on a new configuration, with the polarized ferroelectric layers. Further cooling turns $\omega_> = \omega_m[P(T_{c2}), 0]$ positive at some temperature $T_{c2} < T_{c1}$, the corresponding ferromagnetic transformation takes place. The critical condition $\omega_e(0,0) = 0$ of the first bifurcation defines a relation among the set of critical parameters, such as the supercooling transition temperature T_{c1} , electric field, applied stress, system geometry as defined by the thicknesses of the films and the substrate, etc. of the ferroelectric layers. The critical condition $\omega_m[P(T_{c2}), 0] = 0$ then yields the critical properties of the second bifurcation, such as phase transition temperature, Curie-Weiss relation of the ferromagnetic layers.

To evaluate the condition $\omega_< = 0$ for the second bifurcation from Eq. (8.3.26), the stable state of the system after the first bifurcation needs to be defined as a function of environmental parameters such as temperature, applied stress, electric or magnetic fields, etc. It can be determined by solving the steady-state version of Eqs. (8.3.16) and (8.3.17). The conditions for the first and second bifurcations $\omega_> = 0$ and $\omega_< = 0$ thus determine the characteristics of the supercooling phase transitions of the system, such as the critical temperatures, critical thicknesses, Curie-Weiss relations, etc. Similarly, the characteristics of the superheating phase transitions of the system are determined by the first and second bifurcations $\omega_< = 0$ and $\omega_> = 0$ respectively. In this regard, it is of practical interest to note that in this region near the second bifurcation, the system exhibits the strongest

magnetoelectric effect. In this region, manipulation of the polarization (magnetization) in the ferroelectric (ferromagnetic) layers via the environmental parameters controls the stability of the magnetic (electric) state of the other component, thus allowing magnetic switching caused by an electric input, and vice versa. Thus, if $T_{c2} < T_{c1}$, it is easy to see that an oscillating electric field can create a magnetic hysteresis.

In general, due to the large difference between the quantities $\omega_m(0,0)$, $\omega_m(0,M)$, $\omega_m(P,0)$ and $\omega_m(P,M)$ and $\omega_e(0,0)$, $\omega_e(0,M)$, $\omega_e(P,0)$ and $\omega_e(P,M)$, the system behavior during heat-up and cool-down, or under an applied oscillation field (elastic, electric and magnetic) would be very complex but, at the same time, very interesting.

Before finishing this section, we note that the bifurcation conditions also depend on the thicknesses of the layers via the values of k_e , k_m from Eq. (8.3.26). The magnetoelectric effect due to the coupled elastic interaction will change their phase transition temperatures. K_e , k_m , from Eq. (8.3.27), can take infinite values. The solution corresponding to the smallest values of k_e , k_m gives the meaningful solution for our problems.

As a simple example to demonstrate the foregoing results, we consider a multilayer structure composed of alternating layers (10 each) of ferroelectric BaTiO₃ and ferromagnetic EuO on a SrTiO₃ (001) substrate. The material parameters used in the calculation are listed in Table 8.6.

Table 8.6 Parameters of the films and substrate

BaTiO ₃ /SrTiO ₃	EuO
$C_{11}^E = 1.76 \times 10^{11} \text{ N/m}^2$, $C_{12}^E = 0.846 \times 10^{11} \text{ N/m}^2$	$C_{11}^M = 1.79 \times 10^{11} \text{ N/m}^2$
$C_{11}^0 = 3.16 \times 10^{11} \text{ N/m}^2$, $C_{11}^0 = 1.02 \times 10^{11} \text{ N/m}^2$	$C_{12}^M = 0.77 \times 10^{11} \text{ N/m}^2$
$a_p = 0.39994 \text{ nm}$ ($a_s = 0.3905 \text{ nm}$),	$a_m = 0.838 \text{ nm}$, $\tilde{a}_m = 0.419 \text{ nm}$,
$A_e = 6.6 \times 10^5 \text{ m/FK}$, $T_{c0}^p = 383 \text{ K}$,	$M_s = 1049 \times 10^3 \text{ A/m}$, $T_{c0}^M = 69.15 \text{ K}$
$P_s = 0.26 \text{ C/m}^2$, $D_e = 0.3 \times 10^{-9} \text{ (m}^3/\text{F)}$	$A^M = 0.034$, $D_m = 5.15 \times 10^{-14} \text{ Jm/A}^2$
$\delta_e = 4 \text{ nm}$, $Q = -0.043 \text{ m}^4/\text{C}^2$	$\delta_m = 8 \text{ nm}$, $\lambda = -20 \times 10^{-6}$

Starting with an initial state of the system in which all layers are in the para-state, i.e., $P_0 = M_0 = 0$, Fig. 8.19 shows the respective ferroelectric and the ferromagnetic transition temperatures for different thicknesses of the substrate. It can be seen that the former is higher than the latter in all cases under consideration. Furthermore, it also shows that the transition temperatures decrease as expected with decreasing film thicknesses.

To compare the critical temperatures of ferromagnetic transition (treated as a second bifurcation) with and without the misfit stresses, misfit stress related phase transition temperatures are shown in Fig. 8.20. About a 300% increase in the transition temperature clearly shows the effect of the coupling of the misfit-stresses presently considered.

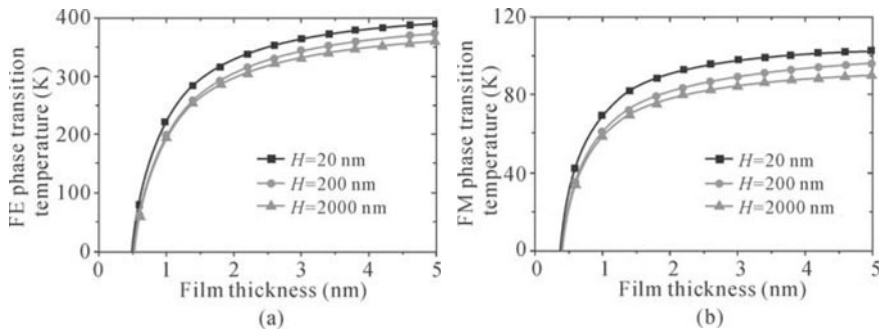


Fig. 8.19 FE phase transition temperature (a) and FM phase transition temperature (b) under different substrate thicknesses

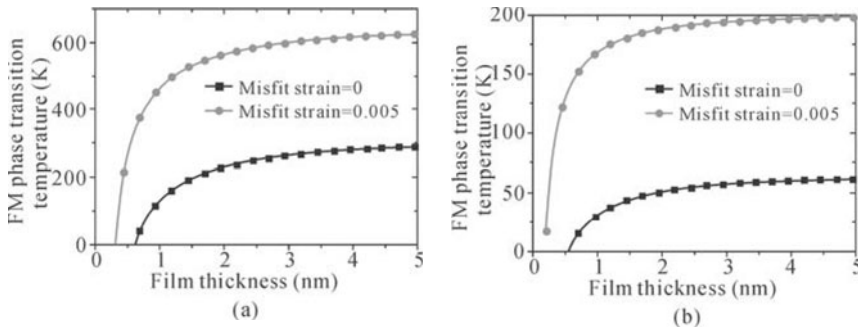


Fig. 8.20 Ferroelectric phase transition temperature (a) and ferromagnetic phase transition temperature (b) under different misfit strains

Fig. 8.21 shows the magnetoelectric effect on the para/ferromagnetic transition temperature (the second bifurcation) when the spontaneous polarization (0.26 C/m^2) is turned on and off. The large shift in the transition temperature will produce a large change in the magnetic susceptibility via the Curie-Weiss relation.

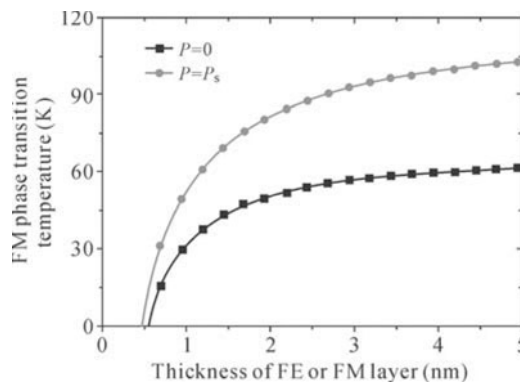


Fig. 8.21 Ferromagnetic phase transition temperature under different spontaneous polarizations

8.3.3 Magnetoelectric Coupling Coefficients in Layered Structures

The magnetoelectric coupling coefficient is an important parameter measuring the strength of the coupling effect in piezoelectric/piezomagnetic composites. Among a series of ME coupling coefficients with different field orientations, there are two important cases: (a) the longitudinal $\alpha_{E,33}$ for the fields of both perpendicular to the sample plane and (b) the transverse $\alpha_{E,31}$ for electric fields along a direction perpendicular to the sample plane and magnetic fields in the plane of the sample. The constitutive equation for describing the coupled mechanical-electric-magnetic response in the ME composites under the linear approximation can be written by direct notation for tensors as

$$\begin{aligned}\boldsymbol{\sigma} &= \mathbf{c}\mathbf{S} - \mathbf{e}^T \mathbf{E} - \mathbf{q}^T \mathbf{H}, \\ \mathbf{D} &= \mathbf{e}\mathbf{S} + \boldsymbol{\varepsilon}\mathbf{E} + \boldsymbol{\alpha}\mathbf{H}, \\ \mathbf{B} &= \mathbf{q}\mathbf{S} + \boldsymbol{\alpha}^T \mathbf{E} + \boldsymbol{\mu}\mathbf{H},\end{aligned}\tag{8.3.28}$$

where $\boldsymbol{\sigma}$, $\mathbf{S}$, $\mathbf{D}$, $\mathbf{E}$, $\mathbf{B}$, and $\mathbf{H}$ are the stress, strain, electric displacement, electric field, magnetic induction and magnetic field, respectively; $\mathbf{c}$, $\boldsymbol{\varepsilon}$, and $\boldsymbol{\mu}$ are the elastic stiffness, dielectric constant and magnetic permeability, respectively; $\mathbf{e}$ and $\mathbf{q}$ are the piezoelectric and piezomagnetic coefficients, respectively; $\boldsymbol{\alpha}$ is the ME coefficient. The superscript T means the transpose of the tensor.

With the consideration of a simple bilayered composite of piezoelectric and magnetic ceramic phases with a free boundary and the polarization direction coinciding with the axis x_3 , the boundary value problem involved is tractable for such a simple bilayered structure (Avellaneda and Harshé, 1994; Getman, 1994). Thus, the constitutive equation can be directly solved by an averaging method to estimate effective material parameters. For homogeneous bilayers with short-circuit condition and ideal interface coupling, the classical longitudinal magnetoelectric voltage coefficient is (Harshe et al., 1993)

$$\alpha'_{E,31} = \frac{-2f(1-f)d_{31}^p q_{31}^m}{[(s_{11}^m + s_{12}^m)f + (s_{11}^p + s_{12}^p)(1-f)]\varepsilon_{33}^p - 2(d_{31}^p)^2(1-f)},\tag{8.3.29}$$

and the transverse magnetoelectric voltage coefficient is

$$\alpha'_{E,33} = \frac{-2f(1-f)d_{31}^p(q_{11}^m + q_{21}^m)}{[(s_{11}^m + s_{12}^m)f + (s_{11}^p + s_{12}^p)(1-f)]\varepsilon_{33}^p - 2(d_{31}^p)^2(1-f)},\tag{8.3.30}$$

where f is the volume fraction of the ferroelectric phase and can be expressed as $f = h/(h + h_s)$. s_{ij}^m is the compliance of the substrate. q_{ij}^m is the piezomagnetic coefficient. d_{ij}^p and ε_{ij}^p are the piezoelectric coefficients and permittivity of the

ultra-thin film, respectively.

With parameters taken from (Liu et al., 2006; Lu et al., 2007), the ME coupling coefficients of $\text{BaTiO}_3\text{-CoFe}_2\text{O}_4$ are shown in Fig. 8.22.

The longitudinal ME coupling coefficient is usually lower than the transverse one, mainly caused by the anisotropy of the piezomagnetic coefficients. Since the piezoelectric and dielectric properties also contribute to the ME coupling effect and can be easily manipulated by strain engineering, a significant change in the longitudinal ME coupling can be expected when the bilayer is subjected to a uniform out-of-plane stress which can also be named x_3 -clamping. But for the transverse mode, this clamping leads to a substantial reduction in α_{E31} in comparison to the unclamped case.

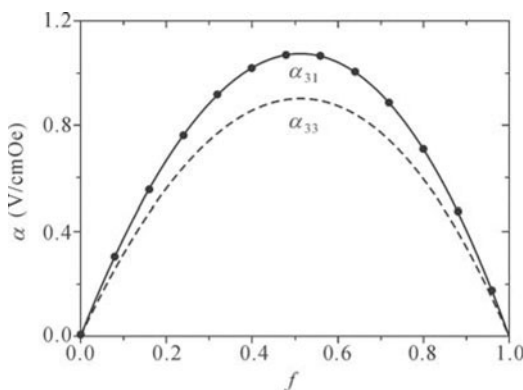


Fig. 8.22 Magnetolectric voltage coefficients vs. the ferroelectric volume fraction using parameters calculated based on the Landau theory

8.4 Magnetolectric Coupling at Resonance Frequency

In composites made of magnetostrictive and piezoelectric phases, the ME coupling is induced by the interaction of the mechanical strains of each phase. As is well known, the strains induced by the electric field in piezoelectric materials and by the magnetic field in piezomagnetic materials are highly frequency-dependent. It is obvious that the frequency can also affect the ME coupling effect mediated by the strain and an enhancement of ME interactions can be expected when the frequency of the AC field is tuned to resonant modes. Recent studies have shown that the ME coupling can be greatly enhanced by the mechanical and magnetic resonances of the respective subsystems. In the vicinity of the bending mode in each subsystem, electromechanical resonance (EMR) of the piezoelectric component or the ferromagnetic resonance (FMR) of the magnetostrictive component, the ME effects are usually several tens times stronger than those at a low frequency (10 Hz–10 kHz). Besides, there is also another resonance at the

overlap of EMR and FMR which is named as the magnetoacoustic resonance (MAR). Possible quantum resonance may also occur with the further miniaturization of the samples. By far, the suggested ME composites used for sensors of magnetic fields, dual electric-field and magnetic-field-tunable devices can be operated at 1–110 GHz.

There is usually a series of frequencies which can be defined by the shape of the vibration mode, which is related to the size and the material properties. Even the first order frequency at EMR or FMR in layered composites is quite high, usually of the order of hundreds of kilohertz or million hertz for nominal sample dimensions, respectively (Bichurin et al., 2003; Zhai et al., 2008). The operating frequency, of course, can be reduced by the increase in laminate size. Another way is to apply the frequency of ac fields to the bilayer at bending oscillations resonance, which is only several tens of kilohertz (Petrov et al., 2009).

8.4.1 Magnetolectric Coupling at Bending Modes

Bending modes include bending oscillations and radial or thickness acoustic modes. At bending resonance, the strain in the structure reaches maximum and strengthens the related electric and magnetic coupling properties. With the bending deformations, the strain, magnetic and electric fields within the sample are all thickness dependent. The frequency dependent ME coupling coefficient can be obtained by solving the electrostatic, magnetostatic and elastodynamic equations.

A small-amplitude oscillation of a bilayer which is formed by magnetostrictive and piezoelectric layers is illustrated in Fig. 8.23. The total thickness of the sample t was assumed to be small compared to other dimensions, i.e., its width w and its length L . In that case, only one component of strain and stress needs to be considered. The bilayer was assumed to be poled with an electric field along z . A bias magnetic field along z or along x was applied to the magnetostrictive component corresponding to the longitudinal ME effect and transverse ME effect, respectively.



Fig. 8.23 Bilayer of magnetostrictive and piezoelectric components. The middle plane is defined as $z=0$ and the interface of the bilayer is located at $z=z_0$

Without any axial force, the equation of bending motion for the neutral plane of bilayer which is located at $z=0$ has the form (Timoshenko and Young, 1955)

$$\nabla^2 \nabla^2 u(x) + \frac{\rho t}{D} \frac{\partial^2 u(x)}{\partial \tau^2} = 0, \quad (8.4.1)$$

where $\nabla^2 \nabla^2$ is the biharmonic operator, $u(x)$ is the deflection along the z direction. t and ρ are the thickness and average density of the bilayer, respectively, and τ is the time. D is the cylindrical stiffness. The thickness and the density of the bilayer can be written as $t = {}^p t + {}^m t$ and $\rho = ({}^p \rho {}^p t + {}^m \rho {}^m t) / t$, respectively. ${}^p t$, ${}^m t$ and ${}^p \rho$, ${}^m \rho$ are the thicknesses and densities of the piezoelectric and magnetic layers, respectively.

The position of the neutral plane is generally defined as the point where the stress along the x -direction is zero. The neutral plane for the bilayer z_0 can be determined by

$$z_0 = \frac{1}{2} \frac{{}^p Y {}^p t^2 - {}^m Y {}^m t^2}{{}^p Y {}^p t + {}^m Y {}^m t}, \quad (8.4.2)$$

where ${}^p Y$ and ${}^m Y$ are the elastic modules of piezoelectric and piezomagnetic layers at zero electric and magnetic fields, respectively. The cylindrical stiffness D of the bilayer can be calculated by using the standard technique and taking into account Eq. (8.4.2):

$$D = {}^p D + {}^m D + 3 \left(\frac{(1-\nu)^2}{{}^m D} + \frac{\nu^2}{{}^p D} \right)^{-1}, \quad (8.4.3)$$

where ν is the volume fraction of piezoelectric material, ${}^p D$ and ${}^m D$ are cylindrical stiffnesses which can be defined by the well known expressions: ${}^p D = {}^p Y {}^p t^3 / 12$, ${}^m D = {}^m Y {}^m t^3 / 12$. The general solution of Eq. (8.4.1) can be written for harmonic vibrations as:

$$u(x) = C_1 \sinh(kx) + C_2 \cosh(kx) + C_3 \sin(kx) + C_4 \cos(kx), \quad (8.4.4)$$

where the wave number k is defined by $k^4 = \omega^2 \rho t / D$.

Considering the layer is homogeneous without external fields, the constants C_1, C_2, C_3 , and C_4 are determined by the boundary conditions at the two end points $x=0$ and $x=L$. Three typical boundary conditions are listed as follows (Petrov et al., 2009):

(1) Bilayer with both ends free:

$$M_x = 0, V_x = 0 \text{ at } x = 0, x = L, \quad (8.4.5)$$

where M_x is the turning moment relative to the x -axis which can be expressed as $M_x = u''(x)D_wL$ and V_x is the transverse force which can be defined as $V_x = \frac{\partial M_x}{\partial x}$.

(2) Bilayer with both ends simply supported:

$$u = 0, M_x = 0 \text{ at } x = 0, x = L. \quad (8.4.6)$$

(3) Bilayer with one end free and one end simply supported:

$$u = 0, \frac{\partial w}{\partial x} = 0 \text{ at } x = 0; \quad (8.4.7)$$

$$M_x = 0, V_x = 0 \text{ at } x = L. \quad (8.4.8)$$

For the first case (Case (1)) with both ends free, the boundary conditions can be written as

$$\begin{aligned} C_1 (\sinh(kL) - \sin(kL)) + C_2 (\cosh(kL) - \cos(kL)) &= 0, \\ C_1 (\cosh(kL) - \cos(kL)) + C_2 (\sinh(kL) + \sin(kL)) &= 0, \end{aligned} \quad (8.4.9)$$

with $C_1 = C_3$, and $C_2 = C_4$. The frequency equation can be written as $\cosh(kL)\cos(kL) = 1$, which is a transcendental equation and can be solved approximately. The first wave vector k is zero, the second one is about $4.73/L$, and the third one is $7.8532/L$.

In the second case (Case (2)) with both ends simply supported, the boundary condition can be written as:

$$\begin{aligned} C_1 \sin(kL) + C_3 \sinh(kL) &= 0, \\ -C_1 \sin(kL) + C_3 \sinh(kL) &= 0. \end{aligned} \quad (8.4.10)$$

Since C_1 and C_3 are nonzero, the determinant of the coefficient with respect to C_1 and C_3 must be zero, and the frequency equation can be derived as $\sin(kL) = 0$. The wavenumber k thus can be written as:

$$k_n L = n\pi, \quad n = 1, 2, \dots, \infty. \quad (8.4.11)$$

In the third case (Case (3)) with one end free and one end simply supported, the boundary condition can be written as:

$$\begin{aligned} C_1 (\sinh(kL) + \sin(kL)) + C_2 (\cosh(kL) + \cos(kL)) &= 0, \\ C_1 (\cosh(kL) + \cos(kL)) - C_2 (\sinh(kL) - \sin(kL)) &= 0, \end{aligned} \quad (8.4.12)$$

with $C_1 = -C_3$ and $C_2 = -C_4$. One can derive that $\cosh(kL)\cos(kL) = -1$. The first wave vector k is $1.875/L$, the second one is about $4.694/L$, and the third one is $7.855/L$. The last two are similar to those in Case (1).

The intrinsic frequency ω , with the relationship $k^4 = \omega^2 \rho t / D$, can be derived as

$$\omega = k^2 \sqrt{\frac{D}{\rho t}}. \quad (8.4.13)$$

The intrinsic frequency is dependent on the vibration mode, intrinsic materials parameters (elasticity and density), and the geometric parameters such as the total thickness, length and the volume fractions. With parameters listed in Table 8.7, the frequency in different boundary conditions can be calculated.

Table 8.7 Parameters (elastic modulus Y , density ρ , piezomagnetic coefficient g , piezoelectric coefficient d , permeability μ , and permittivity ϵ) for PZT and Permendur taken from (Petrov et al., 2009). Copyright (2009) by the American Institute of Physics

Material	Y ($\times 10^{11}$ N/m ²)	ρ (kg/m ³)	q_{11} ($\times 10^{-12}$ m/A)	q_{31} ($\times 10^{-12}$ m/A)	d_{31} ($\times 10^{-12}$ m/V)	μ_{11} / μ_0	$\epsilon_{33} / \epsilon_0$
PZT	0.65	7500	—	—	-175	—	1750
Permendur	1.28	8300	-1200	300	—	80	—

For beams with both ends free, the frequency of the first vibration mode is zero. Using the parameters given in Table 8.7, the frequencies of the second vibration mode ($n=2$) and the third vibration mode ($n=3$) vs. the volume fraction of PZT are shown in Fig. 8.24. These two frequencies are above hundreds of kilohertz, which are comparable with that in EMR and FMR.

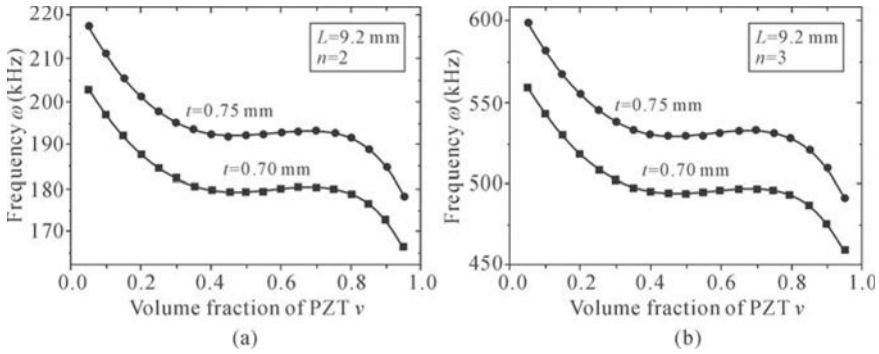


Fig. 8.24 Frequency of PZT/Permendur bilayer with both ends free. (a) The second vibration mode ($n=2$); (b) The third vibration mode ($n=3$)

For beams with both ends simply supported, the frequencies of the first vibration mode ($n=1$) and the second vibration mode ($n=2$) vs. the volume fraction of PZT are shown in Fig. 8.25.

For a beam with one end free and the other end simply supported, the frequencies of the first vibration mode and the second vibration mode are shown in Fig. 8.26. In this case, the first mode frequency is lower than that in Case (2).

With different boundary conditions, the intrinsic frequency of the first mode varies from 0 to about 100 kHz, which is lower than the electromechanical resonance frequency (about 100 kHz–10 MHz). The intrinsic frequency of the beam with both ends simply supported is about three times higher than that of the beam with one end free and one end simply supported. The second mode is also highly dependent on the boundary conditions and the total thickness of the layers. The higher modes are higher in the range above 300 kHz, which is around the electromechanical resonance frequency. With the overlap of the EMR and bending mode, ME coupling effect can be greatly strengthened.

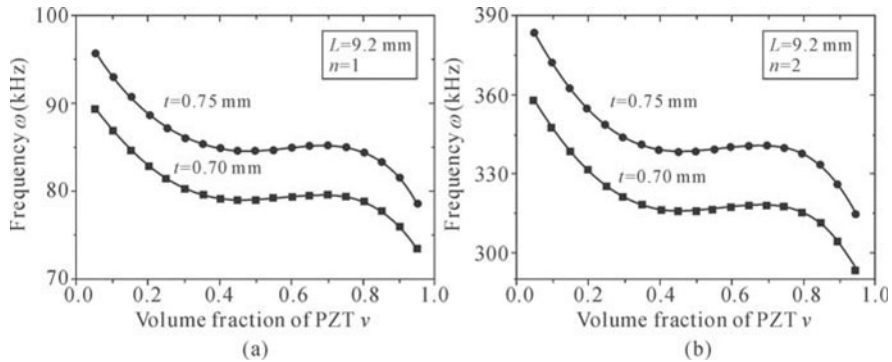


Fig. 8.25 Frequency of PZT/Permendur bilayer with both ends simply supported. (a) The first vibration mode ($n=1$); (b) The second vibration mode ($n=2$)

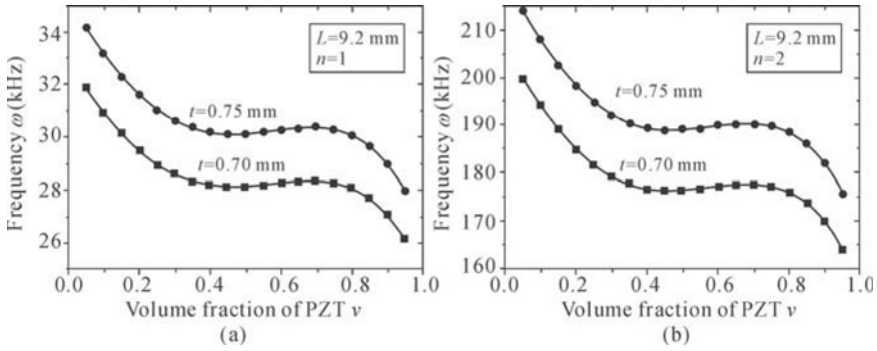


Fig. 8.26 Frequency of PZT/Permendur bilayer with one end free and the other end simply supported. (a) The first vibration mode ($n=1$); (b) The second vibration mode ($n=2$)

The resonance with consideration of the external electric and magnetic fields will be different. Now we consider the bending mode corresponding to the half wavelength over the length of the sample, that is $n=1$. With the general solution, Eq. (8.4.4), for harmonic vibrations of Eq. (8.4.1), the strain as functions of frequency ω and the location of x and z of the two layers can be written as:

$$\begin{aligned} {}^pS_1 &= -z \frac{\partial^2 u}{\partial x^2} = -zk^2 C(x) (z_0 - t_p < z < z_0), \\ {}^mS_1 &= -z \frac{\partial^2 u}{\partial x^2} = -zk^2 C(x) (z_0 < z < z_0 + t_m). \end{aligned} \quad (8.4.14)$$

As an example, we studied the longitudinal ME effect for the bilayer with both ends simply supported. With consideration of ${}^pT_1 = {}^pY({}^pS_1 - d_{31}E_3)$ and ${}^mT_1 = {}^mY({}^mS_1 - q_{31}H_3)$, and the boundary condition of Eq. (8.4.6), the following equations can be derived:

$$\begin{aligned} x=0: \quad C_2 + C_4 &= 0, \\ k^2(C_2 - C_4)(y_p + y_m) &= y_d + y_q; \\ x=L: \quad C_1 \sinh(Lk) + C_3 \sin(Lk) &= -C_2 \cosh(Lk) - C_4 \cos(Lk), \\ (y_p + y_m)C(Lk) &= y_d + y_q. \end{aligned} \quad (8.4.15)$$

In the above equations, $y_p = {}^pY[z_0^3 - (z_0 - t_p)^3]/3$, $y_m = {}^mY[(z_0 + t_m)^3 - z_0^3]/3$, $y_d = {}^pYd_{31} \times E_3[z_0^2 - (z_0 - t_p)^2]/2$ and $y_q = {}^mYq_{31}H_3[(z_0 + t_m)^2 - z_0^2]/2$.

From Eq. (8.4.15), we can derive C_1, C_2, C_3 , and C_4 as follows:

$$C_2 = \frac{y_d + y_q}{2(y_p + y_m)}, \quad C_4 = -C_2, \quad C_1 = \frac{1 - \cosh(kL)}{\sinh(kL)} C_2, \quad C_3 = \frac{\cos(kL) - 1}{\sin(kL)} C_2. \quad (8.4.16)$$

It should be noted that $y_{dq} = y_d + y_q$ is a function of the external electric field E and magnetic field H . Taking into account ${}^pD_3 = {}^p d_{31} {}^p T_1 + {}^p \epsilon_{33} {}^p E_3$ and using the open circuit boundary condition $\int_0^L D_3 dx = 0$, we can find the following relation:

$$[-d_{31}^2 Y_p + \epsilon_{33} + d_{31}^2 Y_p \frac{z_0^2 - (z_0 - t_p)^2}{2} \bar{C}] E_3 = H_3 [-d_{31} Y_p q_{31} Y_m \frac{(z_0 + t_m)^2 - z_0^2}{2} \bar{C}]. \quad (8.4.17)$$

The ME coupling coefficient can be written as

$$\alpha_{E,33} = \frac{\langle E_3 \rangle}{H_3} = \frac{d_{31} Y_p q_{31} Y_m [(z_0 + t_m)^2 - z_0^2] \bar{C}}{2d_{31}^2 Y_p - 2\epsilon_{33} - d_{31}^2 Y_p^2 [z_0^2 - (z_0 - t_p)^2] \bar{C}} \frac{t_p}{t}, \quad (8.4.18)$$

with the average value $\bar{C} = \frac{-k^2}{y_{dq} L} \frac{t_p}{2} \int_0^L C(kx) dx$.

Peaks in α_E occur at the frequency of bending oscillations which have similar values, with and without consideration of the external field. The peak values are two or three orders of magnitude higher than those at low frequencies (Fig. 8.27).

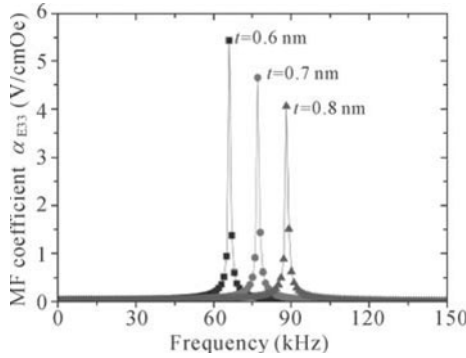


Fig. 8.27 ME coupling coefficient of PZT/Permendur (50/50) vs. frequency with different total thickness. $L=9.2$ mm

8.4.2 Magnetoelectric Coupling at Electromechanical Resonance

The ME coupling coefficients show significant enhancement at the bending resonance. However, experiments and applications require smaller deformations other than bend. ME resonance was first investigated taken out under a wide range

of frequencies, i.e., $f=10$ Hz–500 kHz (Bichurin et al., 2003). The sample was initially p-poled in an electric field and then subjected to a bias magnetic field H and an ac magnetic field δH . The only difference in the study of the EMR-dependent ME coupling coefficient is the frequency, which should be high enough to cover the EMR. Such a resonance phenomenon usually occurs at 100 kHz to 10 MHz. Although the frequencies are higher than in bending mode, they are still much smaller than magnon, optical-phonon or FMR frequencies (Bichurin et al., 2003).

To experimentally study the ME response at resonance frequency, the first step is to measure the ME response at a given ac magnetic field frequency as a function of the static magnetic field to determine the static field corresponding to the peak response. The static field is then set to this value and the frequency of the ac magnetic field frequency is swept to observe the resonance behavior (Bichurin et al., 2003; Zgür et al., 2009). As shown in Fig. 8.28, the longitudinal ME voltage coefficient measured for a 10 mm diameter multilayer sample consisting of one layer of 13 μm NiFe_2O_4 and 10 layers of 26 μm PZT shows a resonance peak more than 40 times higher than the low frequency values, at a frequency of about 350 kHz (Bichurin et al., 2003). The frequency dependence of the dielectric constant of a single 10 mm-diameter PZT disk shows a resonance (discontinuity) close to the peak position of the ME coefficient, indicating that the enhancement in ME coefficient is due to electromechanical coupling.

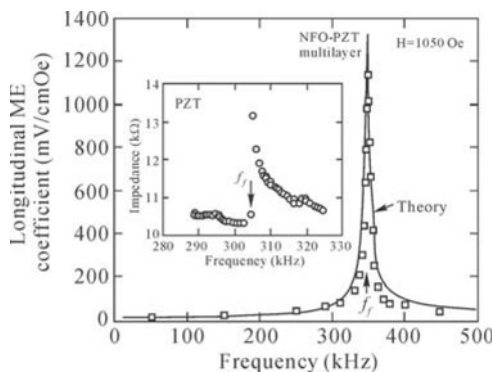


Fig. 8.28 Longitudinal ME coefficient as a function of the AC magnetic field frequency for a multilayer containing one layer of 13 μm NiFe_2O_4 and 10 layers of 26 μm PZT. The solid line represents the theoretical calculations. The inset shows the impedance for a 10 mm disk of PZT as a function of the frequency (Bichurin et al., 2003). Copyright (2003) by the American Physics Society

Theoretical studies were also developed for such a phenomenon (Bichurin et al., 2003). If the layer thickness is small compared to wavelengths for the acoustic modes, one can assume that the composite is homogeneous. The coupling properties can be described by effective parameters, such as compliance, piezoelectric and magnetostrictive coefficients that can be determined from

parameters for individual phases. The ME coupling coefficient at EMR can be probed based on the exact solution of the equations of motions of a medium which can be written as

$$\bar{\rho} \frac{\partial^2 w}{\partial t^2} = \nu \frac{\partial^p T_{ij}}{\partial x_j} + (1-\nu) \frac{\partial^m T_{ij}}{\partial x_j}, \quad (8.4.19)$$

where $\bar{\rho}$ is the average density of the layered structure. It should be noted that the size of the composite is smaller than the electromagnetic wavelength and that the gradients of the electric and magnetic fields inside the sample can be ignored. It is interesting that the mechanical oscillations of the magnetoelectric composite can be generated by alternating both magnetic and electric fields.

In the last section, we know that the size and shape can greatly affect the ME coupling coefficients besides the orientation of the electric and magnetic fields. In this section, we will take a rectangular laminate bilayer with length L , width W and thickness t as an example. The strain components ${}^p S_1$ and ${}^m S_1$ in the x direction of the magnetic and piezoelectric layers can be written as ${}^p S_1 = {}^p s_{11} {}^p T_1 + {}^p s_{12} {}^p T_2 + {}^p d_{31} E_3$ and ${}^m S_1 = {}^m s_{11} {}^m T_1 + {}^m s_{12} {}^m T_2 + {}^m q_{31} H_3$, respectively. Similarly, the strain components in the y direction can also be expressed. For simplicity, the quantities of strain components in the y -direction can be disregarded with the assumption of ${}^p T_1 \gg {}^p T_2$ and ${}^m T_1 \gg {}^m T_2$ in the vicinity of the electromechanical resonance.

Taking into account the expressions for the stresses ${}^p T_1$ and ${}^m T_1$, the differential equation for u_x can be derived. Under the assumption that the motion along the x axis is harmonic, the solution to the equation can be written as

$$u_x = A \cos(kx) + B \sin(kx), \quad (8.4.20)$$

where $k = \omega \sqrt{\bar{\rho} [\nu / {}^p s_{11} + (1-\nu) / {}^m s_{11}]^{-1}}$ and ω is the angular frequency. Constants A and B can be determined from the boundary conditions. With the assumption that the surfaces of the sample at $x=0$ and $x=L$ are free from the external stress, the boundary can be written in the form $\nu {}^p T_1 + (1-\nu) {}^m T_1 = 0$. Using the open circuit condition $\int_0^L D_3 dx = 0$, the magnetoelectric voltage coefficient can be derived as (Bichurin et al., 2003)

$$\alpha_{E,33} = \frac{2 {}^p d_{31} {}^m q_{31} {}^p s_{11} \nu (1-\nu) \tan(2kL)}{{}_s ({}^p d_{31}^2 - {}^p s_{11} {}^p \epsilon_{33}) kL - 2 {}^p d_{31}^2 \nu {}^m s_{11} \tan(2kL)}, \quad (8.4.21)$$

where $s_1 = \nu^m s_{11} + (1 - \nu)^p s_{11}$.

In the following, a NiFe₂O₄-PZT bilayer is taken as an example. Parameters used in the calculation are listed in Table 8.8 (Petrov et al., 2009).

Table 8.8 Parameters for PZT and nickel ferrite taken from (Petrov et al., 2009). Copyright (2008) by the American Institute of Physics

Material	s_{11} ($\times 10^{-12}$ m ² /N)	s_{12} ($\times 10^{-12}$ m ² /N)	ρ (kg/m ³)	q_{11} ($\times 10^{-12}$ m/A)	q_{31} ($\times 10^{-12}$ m/A)	d_{31} ($\times 10^{-12}$ m/V)	μ_{11}/μ_0	ϵ_{33}/ϵ_0
PZT	15.3	5	7500	—	—	−175	—	1750
NiFe ₂ O ₄	6.5	−2.4	5370	−680	70	—	80	—

Fig. 8.29 shows the frequency dependences of the longitudinal magnetoelectric voltage coefficients for the NiFe₂O₄-PZT bilayer. It is obvious that the coefficient at the resonance peaks caused by the oscillations is far higher than that at the low frequency. We should note that the calculated ME coefficient is much higher than that achieved in experiments, as shown in Fig. 8.32. But the calculated corresponding resonance frequencies agree with those observed in experiments very well.

Compared with the resonance frequency in PZT/Permendur (50/50), the resonance frequency of PZT-NiFe₂O₄ is several times higher. The corresponding ME coupling coefficient at resonance is also higher than that of PZT/Permendur (50/50). The material properties such as the piezoelectric coefficient, piezomagnetic coefficient, size, elastic properties and direction of applied fields, contributed to the resonance ME coupling effect and the corresponding frequency. Composites with designed resonance frequency are thus suggested for special applications.

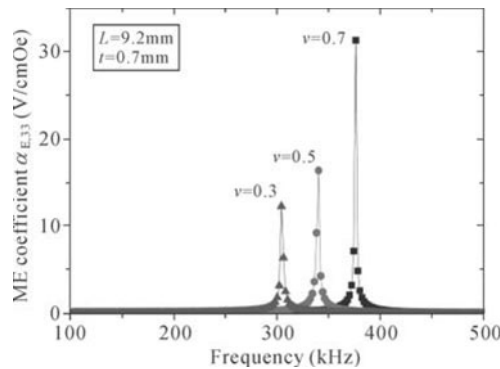


Fig. 8.29 Longitudinal ME coefficient vs. frequency with different volume fractions of PZT

8.4.3 *Magnetoelectric Coupling at Ferromagnetic Resonance*

Low-loss ferrimagnets, including yttrium iron garnet (YIG), lithium ferrite and

hexagonal ferrites, are one kind of material for wide use in microwave signal processing devices such as resonators, filters and phase shifters, due to low losses. The operating frequency is in the range of 1 – 70 GHz. However, devices based on such materials are usually slow, noisy and with large power consumption. An alternative way is to use ferrite-piezoelectric composites in which the magnetic frequency shift can be tuned by using an electric field. Obviously, composites with a large ME coupling effect are needed. It is known that most iron based ferromagnetic oxides show two high-frequency resonances in the magnetic susceptibility: one in the microwave region associated with the net ferromagnetic moment and a second resonance at a higher frequency corresponding to the antiferromagnetic alignment of sublattice moments (Bichurin et al., 2001; 2002). It is not surprising that the ME coupling effect may be unique in the vicinity of these two resonance regimes (Bichurin et al., 2001).

Fetisov and Srinivasan (2006) recently designed a PZT/YIG based microstripline resonator (Fig. 8.30), and large shift according to the electric field was obtained. For $E=10$ kV/cm across PZT, a frequency shift of 18 – 25 MHz was achieved. Moreover, upon reversal of the direction of E , the shift also reversed. The YIG-PZT bilayer structures could be a potential candidate for rapid, electric field tunable, passive microwave resonators and filters.

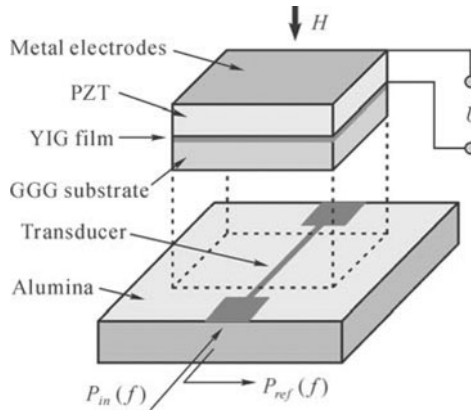


Fig. 8.30 Diagram showing a microstripline resonator with a bilayer of lead zirconate titanate (PZT) bonded to (111) yttrium iron garnet (YIG) film on gadolinium gallium garnet (GGG) substrate (Fetisov and Srinivasan, 2006). Copyright (2006) by the American Institute of Physics

The FMR frequency f_1 for in-plane H and f_2 for H , perpendicular to the plane, can be expressed in cgs units as

$$\begin{aligned} f_1 &= \gamma [H(H + 4\pi M_{\text{eff}})]^{1/2}, \\ f_2 &= \gamma (H - 4\pi M_{\text{eff}}), \end{aligned} \quad (8.4.22)$$

where γ is the gyromagnetic ratio and M_{eff} is the effective magnetization. For

YIG with $\gamma = 2.8$ MHz/Oe and $4\pi M = 1750$ G, the FMR frequency is in the range of 2.8 – 8.8 GHz. It should be noted that measurements could not be done below 2.5 GHz due to the broadening of FMR caused by the nonlinear effects, or above 9 GHz because of a drop in the efficiency of excitation of FMR with the stripline transducer (Chen et al., 2008).

Due to the ME coupling effect, the frequency shift can be tuned by the electric field. The in-plane resonance frequency under applied electric fields can be described as

$$f = \gamma[(H_0^{\parallel} + \delta H_E)(H_0^{\parallel} + \delta H_E + 4\pi M_{\text{eff}})]^{1/2}, \quad (8.4.23)$$

where δH_E is the electric field induced by the effective magnetic field and can be described by the ME coupling coefficient as $\delta H_E = \alpha E$.

A summary of magnetoelectric coupling parameters and relative frequencies for some laminated structures is listed in Table 8.9 (Chen et al., 2008).

Table 8.9 Summary of magnetoelectric coupling parameters and relative frequency for some laminated structures (Chen et al., 2008)

Structure	A (Oe cm/kV)	δH_E (Oe)	Δf (MHz)	E (kV/cm)	F (GHz)
YIG/PZT	0.39	7	40	18	5
YIG/PZT	0.88	8.75	25	10	2 – 10
NFO/PZT	1.1	330	925	300	9.3
LFO/PZT	0.39	117	—	300	9.3
LFO/PZT	0.5	40	—	80	9.3
YIG/GGG/PMN-PT	5.5	44	122	8	9.3
YIG/PMN-PT	50	150	420	30	9.3
FeGaB/Si/PMN-PT	3.4	28	900	8	2-3
Ni2MnGa/PMN-PT	41	230	370	6	9.5

PZT: lead zirconate titanate; NFO: nickel ferrite; LFO: lithium ferrite; YIG: yttrium iron garnet ferrite; and GGG: gadolinium gallium garnet

In fact, the ME interactions at FMR are also dependent on the applied electric field which can be either a DC or AC electric field. Further studies may include (a) coupling between microwave magnetic fields and DC electric fields, (b) coupling between microwave magnetic and electric fields, and (c) coupling when the FMR and EMR frequencies overlap, i.e., at MAR.

8.5 Remarks

Chapter 8 began with a brief summary of the history and the renaissance of magnetoelectric/multiferroic materials. The typical material systems and the

corresponding coupling mechanisms and theoretical explanations were discussed in detail, with respect to two major systems—single phase materials and composites.

For single phase multiferroic materials, the incompatibility between ferroelectricity and ferromagnetism was illustrated. That is mainly because the formation of long-range magnetic order induced by the partially filled inner shells (*d*- or *f*-levels) and the formation of ferroelectricity which needs empty *d*-state ions exclude each other. But not all experimental and theoretical results support the argument. Alternative mechanisms for the coexistence of the ferroelectric and (anti)ferromagnetic orders were suggested and can be roughly divided into two categories: the heterogenous source and the homologous source. The coupling in Bi perovskites (ferroelectricity induced by lone-pairs), hexagonal manganites (geometric ferroelectricity) and some systems of charge ordering multiferroics can be put into the former category; some other types of charge ordering multiferroics and spiral spin order induced multiferroics belong to the latter category. Because of the direct interaction in the homologous-source system, the magnetoelectric coupling is stronger than that in the heterogenous-source systems. However, because of the low Néel temperature of most magnetic atoms, the coupling effect usually exists at very low temperatures, which is one vital obstacle for possible applications.

As for the multiferroic composites, the coupling mechanism is quite easy to understand; that is, the elastic interaction plays an important role in the ME coupling in the ferroelectric (piezoelectric)/magnetic (magnetostrictive) composites. Large ME coefficients have been achieved in this kind of composite, especially at the resonance stage, which is very promising for device applications. Driven by the desire for miniaturized devices, further research will certainly be devoted to nanostructured materials. The interface effect, boundary conditions and possible dynamic coupling interaction should receive more attention.

Recently, there have been some other types of ME effect which have been found in the structure of the magnetic electrode and the ferroelectric ultrathin layer. Significant resistance manipulation of the structure can be induced by switching the polarization directions. These new findings may have wide potential applications in memory devices and sensors. Further studies may be needed to focus on the possibility of four resistance states in multiferroic tunneling junctions (MFTJs) with a pronounced difference in resistance and the possibility of controlling these resistances by an electric field (through ferroelectric polarization of the barrier) and by a magnetic field (through magnetization configuration of the electrodes).

References

- [1] Aguilar, R.V., Sushkov, A.B., Choi, Y.J., Cheong, S.W., Drew, H.D.: Spin phonon coupling in frustrated magnet CdCr_2O_4 . *Phys. Rev. B* **77**, 092412 (2008)

- [2] Anisimov, V.I., Aryasetiawan, F., Lichtenstein, A.I.: First-principles calculations of the electronic structure and spectra of strongly correlated systems: the LDA+ U method. *J. Phys. Condens. Matter.* **9**, 767-808 (1997)
- [3] Anisimov, V.I., Zaanen, J., Andersen, O.K.: Band theory and Mott insulators: Hubbard U instead of Stoner I. *Phys. Rev. B* **44**, 943-954 (1991)
- [4] Ascher, E., Rieder, H., Schmid, H., Stossel, H.: Some properties of ferromagnetoelectric nickel-iodine boracite, $\text{Ni}_3\text{B}_7\text{O}_{13}\text{I}$. *J. Appl. Phys.* **37**, 1404-1405 (1966)
- [5] Astrov, D.N.: The magnetoelectric effect in antiferromagnetics. *Sov. Phys. JETP* **11**, 708-709 (1960)
- [6] Atou, T., Chiba, H., Ohoyama, K., Yamaguchi, Y., Syono, Y.: Structure determination of ferromagnetic perovskite BiMnO_3 . *J. Solid State Chem.* **145**, 639-642 (1999)
- [7] Avellaneda, M., Harshé, G.: Magnetoelectric effect in piezoelectric/magnetostrictive multilayer (2-2) composites. *J. Intell. Mater. Syst. Struct.* **5**, 501-513 (1994)
- [8] Benveniste, Y.: A new approach to the application of Mori-Tanaka's theory in composite materials. *Mech. Mater.* **6**, 147-157 (1987)
- [9] Benveniste, Y.: Magnetoelectric effect in fibrous composites with piezoelectric and piezomagnetic phases. *Phys. Rev. B* **51**, 16424-16427 (1995)
- [10] Bertaut, E.F., Pauthenet, R., Mercier, M.: Propriétés magnétiques et structures du manganite d'yttrium. *Phys. Lett.* **7**, 110-111 (1963)
- [11] Bichurin, M.I., Filippov, D.A., Petrov, V.M., Laletsin, V.M., Paddubnaya, N., Srinivasan, G.: Resonance magnetoelectric effects in layered magnetostrictive-piezoelectric composites. *Phys. Rev. B* **68**, 132408 (2003)
- [12] Bichurin, M.I., Kornev, I.A., Petrov, V.M., Tatarenko, A.S., Kiliba, Y.V., Srinivasan, G.: Theory of magnetoelectric effects at microwave frequencies in a piezoelectric/magnetostrictive multilayer composite. *Phys. Rev. B* **64**, 94409 (2001)
- [13] Bichurin, M.I., Petrov, V.M., Kiliba, Y.V., Srinivasan, G.: Magnetic and magnetoelectric susceptibilities of a ferroelectric/ferromagnetic composite at microwave frequencies. *Phys. Rev. B* **66**, 134404 (2002)
- [14] Boomgaard, J.V.D., van Run, A., van Suchtelen, J.: Magnetoelectricity in piezoelectric-magnetostrictive composites. *Ferroelectrics* **10**, 295-298 (1976)
- [15] Brink, J., Khomskii, D.I.: Multiferroicity due to charge ordering. *J. Phys. Condens. Matter.* **20**, 434217 (2008)
- [16] Brinks, H.W., Fjellvg, H., Kjekshus, A.: Synthesis of metastable perovskite-type YMnO_3 and HoMnO_3 . *J. Solid State Chem.* **129**, 334-340 (1997)
- [17] Brixel, W., Rivera, J.P., Steiner, A., Schmid, H.: Magnetic field induced magnetoelectric effects (ME) H, in the perovskites Pb_2CoWO_6 and $\text{Pb}_2\text{FeTaO}_6$. *Ferroelectrics* **79**, 201-204 (1988)
- [18] Chaudhuri Roy, A., Ranjith, R., Krupanidhi, S.B., Mangalam, R.V.K., Sundaresan, A.: Interface dominated biferroic $\text{La}_{0.6}\text{Sr}_{0.4}\text{MnO}_{3/0.7}\text{Pb}(\text{Mg}_{1/3}\text{Nb}_{2/3})\text{O}_{3-0.3}\text{PbTiO}_3$ epitaxial superlattices. *Appl. Phys. Lett.* **90**, 122902 (2007)

- [19] Chen, Y., Wang, J., Liu, M., Lou, J., Sun, N.X., Vittoria, C., Harris, V.G.: Giant magnetoelectric coupling and E-field tunability in a laminated Ni_2MnGa /lead-magnesium-niobate-lead titanate multiferroic heterostructure. *Appl. Phys. Lett.* **93**, 112502 (2008)
- [20] Cheong, S.W., Mostovoy, M.: Multiferroics: a magnetic twist for ferroelectricity. *Nat. Mater.* **6**, 13-20 (2007)
- [21] Chiba, H., Atou, T., Syono, Y.: Magnetic and electrical properties of $\text{Bi}_{1-x}\text{Sr}_x\text{MnO}_3$: hole-doping effect on ferromagnetic perovskite BiMnO_3 . *J. Solid State Chem.* **132**, 139-143 (1997)
- [22] Choi, Y.J., Yi, H.T., Lee, S., Huang, Q., Kiryukhin, V., Cheong, S.W.: Ferroelectricity in an Ising chain magnet. *Phys. Rev. Lett.* **100**, 47601 (2008)
- [23] Cottam, M.G., Tilley, D.R., Zeks, B.: Theory of surface modes in ferroelectrics. *J. Phys. C Solid State Phys.* **17**, 1793-1823 (1984)
- [24] Curie, P.: Sur la symétrie dans les phénomènes physiques, symétrie d'un champ électrique et d'un champ magnétique. *Journal de Physique Théorique et Appliquée* **3**, 393-415 (1894)
- [25] Das, R.R., Kim, D.M., Baek, S.H., Eom, C.B., Zavaliche, F., Yang, S.Y., Ramesh, R., Chen, Y.B., Pan, X.Q., Ke, X.: Synthesis and ferroelectric properties of epitaxial BiFeO_3 thin films grown by sputtering. *Appl. Phys. Lett.* **88**, 242904 (2006)
- [26] Deng, C., Zhang, Y., Ma, J., Lin, Y.: Magnetoelectric effect in multiferroic heteroepitaxial BaTiO_3 - NiFe_2O_4 composite thin films. *Acta Mater.* **56**, 405-412 (2008)
- [27] Dong, S., Cheng, J., Li, J.F., Viehland, D.: Enhanced magnetoelectric effects in laminate composites of Terfenol-D/ $\text{Pb}(\text{Zr},\text{Ti})\text{O}_3$ under resonant drive. *Appl. Phys. Lett.* **83**, 4812-4814 (2003)
- [28] Dong, S., Zhai, J., Wang, N., Bai, F., Li, J., Viehland, D., Lograsso, T.A.: $\text{Fe-Ga/Pb}(\text{Mg}_{1/3}\text{Nb}_{2/3})\text{O}_3$ - PbTiO_3 magnetoelectric laminate composites. *Appl. Phys. Lett.* **87**, 2504-2506 (2005)
- [29] Du, S.Y., Wang, B.: *Micromechanics of Composite Materials*. Science Press, Beijing, China (1998)
- [30] Eerenstein, W., Mathur, N.D., Scott, J.F.: Multiferroic and magnetoelectric materials. *Nature* **442**, 759-765 (2006)
- [31] Eerenstein, W., Morrison, F.D., Dho, J., Blamire, M.G., Scott, J.F., Mathur, N.D.: Comment on "Epitaxial BiFeO_3 Multiferroic Thin Film Heterostructures". *Science* **307**, 1203 (2005)
- [32] Efremov, D.V., van den Brink, J., Khomskii, D.I.: Bond-versus site-centred ordering and possible ferroelectricity in manganites. *Nat. Mater.* **3**, 853-856 (2004)
- [33] Fetisov, Y., Srinivasan, G.: Electric field tuning characteristics of a ferrite-piezoelectric microwave resonator. *Appl. Phys. Lett.* **88**, 143503 (2006)
- [34] Fiebig, M.: Revival of the magnetoelectric effect. *J. Phys. Condens. Matter.* **38**, R123-R152 (2005)
- [35] Fischer, P., Polomska, M., Sosnowska, I., Szymanski, M.: Temperature dependence of the crystal and magnetic structures of BiFeO_3 . *J. Phys. C*

- Solid State Phys. **13**, 1931-1940 (1980)
- [36] Folen, V.J., Rado, G.T., Stalder, E.W.: Anisotropy of the magnetoelectric effect in Cr_2O_3 . Phys. Rev. Lett. **6**, 607-608 (1961)
 - [37] Fox, D.L., Tilley, D.R., Scott, J.F., Guggenheim, H.J.: Magnetoelectric phenomena in BaMnF_4 and $\text{BaMn}_{0.99}\text{Co}_{0.01}\text{F}_4$. Phys. Rev. B **21**, 2926-2936 (1980)
 - [38] Fuentes, L., Olivera, R., Garcia, M., Fuentes, M.: Meso and nano-magnetoelectricity: a review. Rev. Mex. Fis. **53**, 21-29 (2007)
 - [39] Getman, I.: Magnetoelectric composite materials: theoretical approach to determine their properties. Ferroelectrics **162**, 45-50 (1994)
 - [40] Harshe, G., Dougherty, J.P., Newnham, R.E.: Theoretical modeling of multilayer magnetoelectric composites. Int. J. Appl. Electromagn. Mater. **4**, 145-159 (1993a)
 - [41] Harshe, G., Dougherty, J.P., Newnham, R.E.: Theoretical modeling of 3-0/0-3 magnetoelectric composites. Int. J. Appl. Electromagn. Mater. **4**, 161-171 (1993b)
 - [42] Hatt, A.J., Spaldin, N.A.: Trilayer superlattices: A route to magnetoelectric multiferroics? Appl. Phys. Lett. **90**, 242916 (2007)
 - [43] Hill, N.A.: Why are there so few magnetic ferroelectrics? J. Phys. Chem. B **104**, 6694-6709 (2000)
 - [44] Hobbs, D., Kresse, G., Hafner, J.: Fully unconstrained noncollinear magnetism within the projector augmented-wave method. Phys. Rev. B **62**, 11556-11570 (2000)
 - [45] Hu, J.: Microscopic origin of magnetoelectric coupling in noncollinear multiferroics. Phys. Rev. Lett. **100**, 77202 (2008)
 - [46] Huang, G.Y., Wang, B.L., Mai, Y.W.: Effective properties of magnetoelectroelastic materials with aligned ellipsoidal voids. Mech. Res. Commun. **36**, 563-572 (2009)
 - [47] Huang, J.H., Chiu, Y.H., Liu, H.K.: Magneto-electro-elastic Eshelby tensors for a piezoelectric-piezomagnetic composite reinforced by ellipsoidal inclusions. J. Appl. Phys. **83**, 5364-5370 (2009)
 - [48] Huang, Z.J., Cao, Y., Sun, Y.Y., Xue, Y.Y., Chu, C.W.: Coupling between the ferroelectric and antiferromagnetic orders in YMnO_3 . Phys. Rev. B **56**, 2623-2626 (1997)
 - [49] Ikeda, N., Ohsumi, H., Ohwada, K., Ishii, K., Inami, T., Kakurai, K., Murakami, Y., Yoshii, K., Mori, S., Horibe, Y.: Ferroelectricity from iron valence ordering in the charge-frustrated system LuFe_2O_4 . Nature **436**, 1136-1138 (2005)
 - [50] Jang, H.W., Baek, S.H., Ortiz, D., Folkman, C.M., Das, R.R., Chu, Y.H., Shafer, P., Zhang, J.X., Choudhury, S., Vaithyanathan, V.: Strain-induced polarization rotation in epitaxial (001) BiFeO_3 thin films. Phys. Rev. Lett. **101**, 107602 (2008)
 - [51] Jiang, Q.H., Shen, Z.J., Zhou, J.P., Shi, Z., Nan, C.W.: Magnetoelectric composites of nickel ferrite and lead zirconate titanate prepared by spark plasma sintering. J. Eur. Ceram. Soc. **27**, 279-284 (2007)

- [52] Kato, K., Iida, S., Yanai, K., Mizushima, K.: Ferrimagnetic ferroelectricity of Fe_3O_4 . *J. Magn. Magn. Mater.* **31-34**, 783-784 (1983)
- [53] Katsura, H., Nagaosa, N., Balatsky, A.V.: Spin current and magnetoelectric effect in noncollinear magnets. *Phys. Rev. Lett.* **95**, 57205 (2005)
- [54] Kenzelmann, M., Harris, A.B., Jonas, S., Broholm, C., Schefer, J., Kim, S.B., Zhang, C.L., Cheong, S.W., Vajk, O.P., Lynn, J.W.: Magnetic inversion symmetry breaking and ferroelectricity in TbMnO_3 . *Phys. Rev. Lett.* **95**, 87206 (2005)
- [55] Khomskii, D: Magnetism and ferroelectricity; why do they so seldom coexist? (2001)
- [56] Khomskii, D.I.: Multiferroics: different ways to combine magnetism and ferroelectricity. *J. Magn. Magn. Mater.* **306**, 1-8 (2006)
- [57] Kimura, S., Hagiwara, M., Ueda, H., Narumi, Y., Kindo, K., Yashiro, H., Kashiwagi, T., Takagi, H.: Observation of higher-harmonic helical spin-resonance modes in the chromium spinel CdCr_2O_4 . *Phys. Rev. Lett.* **97**, 257202 (2006)
- [58] Kimura, T., Goto, T., Shintani, H., Ishizaka, K., Arima, T., Tokura, Y.: Magnetic control of ferroelectric polarization. *Nature* **426**, 55-58 (2003)
- [59] King-Smith, R.D., Vanderbilt, D.: Theory of polarization of crystalline solids. *Phys. Rev. B* **47**, 1651-1654 (1993)
- [60] Kiselev, S.V., Ozerov, R.P., Zhdanov, G.S.: Detection of magnetic order in ferroelectric BiFeO_3 by neutron diffraction (1963)
- [61] Laetsin, U., Padubnaya, N., Srinivasan, G., Devreugd, C.: Frequency dependence of magnetoelectric interactions in layered structures of ferromagnetic alloys and piezoelectric oxides. *Appl. Phys. A: Mater. Scie. Process.* **78**, 33-36 (2004)
- [62] Lawes, G., Harris, A.B., Kimura, T., Rogado, N., Cava, R.J., Aharony, A., Entin-Wohlman, O., Yildirim, T., Kenzelmann, M., Broholm, C.: Magnetically driven ferroelectric order in $\text{Ni}_3\text{V}_2\text{O}_8$. *Phys. Rev. Lett.* **95**, 87205 (2005)
- [63] Li, J.: Magnetoelctroelastic multi-inclusion and inhomogeneity problems and their applications in composite materials. *Int. J. Eng. Sci.* **38**, 1993-2011 (2000)
- [64] Li, J., Levin, I., Slutsker, J., Provenzano, V., Schenck, P.K., Ramesh, R., Ouyang, J., Roytburd, A.L.: Self-assembled multiferroic nanostructures in the $\text{CoFe}_2\text{O}_4\text{-PbTiO}_3$ system. *Appl. Phys. Lett.* **87**, 072909 (2005)
- [65] Li, J.Y., Dunn, M.L.: Micromechanics of magnetoelctroelastic composite materials: average fields and effective behavior. *J. of Intell. Materi. Syste. Struct.* **9**, 404-416 (1998)
- [66] Li, Q., Dong, S., Liu, J.M.: Multiferroic response and clamped domain structure in a two-dimensional spiral magnet: Monte Carlo simulation. *Phys. Rev. B* **77**, 054442 (2008)
- [67] Li, Y.J., Chen, X.M., Lin, Y.Q., Tang, Y.H.: Magnetoelectric effect of $\text{Ni}_{0.8}\text{Zn}_{0.2}\text{Fe}_2\text{O}_4/\text{Sr}_{0.5}\text{Ba}_{0.5}\text{Nb}_2\text{O}_6$ composites. *J. Eur. Ceram. Soc.* **26**, 2839-2844 (2006)

- [68] Liu, G., Nan, C.W., Sun, J.: Coupling interaction in nanostructured piezoelectric/magnetostrictive multiferroic complex films. *Acta Mater.* **54**, 917-926 (2006)
- [69] Lonkai, T., Tomuta, D.G., Hoffmann, J.U., Schneider, R., Hohlwein, D., Ihringer, J.: Magnetic two-dimensional short-range order in hexagonal manganites. *J. Appl. Phys.* **93**, 8191-8193 (2003)
- [70] Lopatin, S., Lopatina, I., Lisnevskaya, I.: Magnetoelectric PZT/ferrite composite material. *Ferroelectrics* **162**, 63-68 (1994)
- [71] Lorenz, B., Litvinchuk, A.P., Gospodinov, M.M., Chu, C.W.: Field-induced reentrant novel phase and a ferroelectric-magnetic order coupling in HoMnO_3 . *Phys. Rev. Lett.* **92**, 087204 (2004).
- [72] Lottermoser, T., Lonkai, T., Amann, U., Hohlwein, D., Ihringer, J., Fiebig, M.: Magnetic phase control by an electric field. *Nature* **430**, 541-544 (2004)
- [73] Lu, X.Y., Wang, B., Zheng, Y., Ryba, E.: Coupling interaction in 1-3-type multiferroic composite thin films. *Appl. Phys. Lett.* **90**, 133124 (2007)
- [74] Lu, X.Y., Wang, B., Zheng, Y., Ryba, E.: Critical phase transition temperatures of 1-3 type multiferroic composite thin films. *J. Phys. Appl. Phys.* **40**, 1614-1619 (2007)
- [75] Lu, X.Y., Wang, B., Zheng, Y., Ryba, E.: Phenomenological theory of 1-3 type multiferroic composite thin film: thickness effect. *J. Phys. Appl. Phys.* **42**, 015309 (2009)
- [76] MacDonald, A.H., Pickett, W.E., Koelling, D.D.: A linearised relativistic augmented-plane-wave method utilising approximate pure spin basis functions. *J. Phys. C Solid State Phys.* **13**, 2675-2683 (1980)
- [77] MacManus-Driscoll, J.L., Zerrer, P., Wang, H., Yang, H., Yoon, J., Fouchet, A., Yu, R., Blamire, M.G., Jia, Q.: Strain control and spontaneous phase ordering in vertical nanocomposite heteroepitaxial thin films. *Nat. Mater.* **7**, 314-320 (2008)
- [78] Martin, L.W., Chu, Y.H., Ramesh, R.: Advances in the growth and characterization of magnetic, ferroelectric, and multiferroic oxide thin films. *Mater. Sci. Eng. Rep.* **68**, 89-133 (2010)
- [79] Marzari, N., Vanderbilt, D.: Maximally localized generalized Wannier functions for composite energy bands. *Phys. Rev. B* **56**, 12847-12865 (1997)
- [80] Moreira dos Santos, A., Cheetham, A.K., Atou, T., Syono, Y., Yamaguchi, Y., Ohoyama, K., Chiba, H., Rao, C.N.R.: Orbital ordering as the determinant for ferromagnetism in biferroic BiMnO_3 . *Phys. Rev. B* **66**, 64425 (2002)
- [81] Moreira dos Santos, A., Parashar, S., Raju, A.R., Zhao, Y.S., Cheetham, A.K., Rao, C.N.R.: Evidence for the likely occurrence of magnetoferroelectricity in the simple perovskite, BiMnO_3 . *Solid State Commun.* **122**, 49-52 (2002)
- [82] Nakamura, M., Okuyama, D., Lee, J.S., Arima, T., Wakabayashi, Y., Kumai, R., Kawasaki, M., Tokura, Y.: Magnetically Tunable Metal—Insulator Superlattices. *Adv. Mater.*, **22**, 500 (2010)
- [83] Nan, C.W., Bichurin, M.I., Dong, S., Viehland, D., Srinivasan, G.: Multiferroic magnetoelectric composites: Historical perspective, status, and

- future directions. *J. Appl. Phys.* **103**, 031101-031135 (2008)
- [84] Nan, C.W.: Magnetoelectric effect in composites of piezoelectric and piezomagnetic phases. *Phys. Rev. B* **50**, 6082-6088 (1994)
- [85] Nan, C.W., Li, M., Huang, J.H.: Calculations of giant magnetoelectric effects in ferroic composites of rare-earth-iron alloys and ferroelectric polymers. *Phys. Rev. B* **63**, 144415 (2001)
- [86] Nan, C.W., Liu, G., Lin, Y., Chen, H.: Magnetic-field-induced electric polarization in multiferroic nanostructures. *Phys. Rev. Lett.* **94**, 197203 (2005)
- [87] Newnham, R.E., Skinner, D.P., Cross, L.E.: Connectivity and piezoelectric-pyroelectric composites. *Mater. Res. Bull.* **13**, 525-536 (1978)
- [88] Norifumi, F., Tadashi, I., Takeshi, Y., Taichiro, I.: Epitaxially grown YMnO_3 film: New candidate for nonvolatile memory devices. *Appl. Phys. Lett.* **69**, 1011-1013 (1996)
- [89] Okuyama, D., Nakamura, M., Wakabayashi, Y., Itoh, H., Kumai, R., Yamada, H., Taguchi, Y., Arima, T., Kawasaki, M. and Tokura, Y.: Epitaxial-strain effect on charge/orbital order in $\text{Pr}_{0.5}\text{Ca}_{0.5}\text{MnO}_3$ films. *Appl. Phys. Lett.*, **95**, 152502 (2009)
- [90] Palai, R., Katiyar, R.S., Schmid, H., Tissot, P., Clark, S.J., Robertson, J., Redfern, S.A.T., Catalan, G., Scott, J.F.: β phase and γ - β metal-insulator transition in multiferroic BiFeO_3 . *Phys. Rev. B* **77**, 14110 (2008).
- [91] Panchapakesan, U., Pollock, C.A., Chen, X.M.: The effect of high glucose and PPAR- $\{\gamma\}$ agonists on PPAR- $\{\gamma\}$ expression and function in HK-2 cells. *Am. J. Phys-Renal Physiol.* **287**, F528-F534 (2004)
- [92] Park, S., Choi, Y.J., Zhang, C.L., Cheong, S.W.: Ferroelectricity in an $S=1/2$ Chain Cuprate. *Phys. Rev. Lett.* **98**, 057601 (2007)
- [93] Petrov, V., Srinivasan, G., Bichurin, M., Galkina, T.: Theory of magnetoelectric effect for bending modes in magnetostrictive-piezoelectric bilayers. *J. Appl. Phys.* **105**, 063911 (2009)
- [94] Petrov, V.M., Bichurin, M.I., Zibtsev, V.V., Mandal, S.K., Srinivasan, G.: Flexural deformation and bending mode of magnetoelectric nanobilayer. *J. Appl. Phys.* **106**, 113901 (2009)
- [95] Picozzi, S., Ederer, C.: First principles studies of multiferroic materials. *J. Phys. Condens. Matter* **21**, 303201 (2009)
- [96] Prellier, W., Singh, M.P., Murugavel, P.: The single-phase multiferroic oxides: from bulk to thin film. *J. Phys. Condens. Matter* **17**, R803-R832 (2005)
- [97] Qi, X., Dho, J., Tomov, R., Blamire, M.G., MacManus-Driscoll, J.L.: Greatly reduced leakage current and conduction mechanism in aliovalent-ion-doped BiFeO_3 . *Appl. Phys. Lett.* **86**, 062903 (2005)
- [98] Qiao, L., Bi, X.: Domain configuration and phase transition for BaTiO_3 thin films on tensile Si substrates. *J. Cryst. Growth* **310**, 5327-5330 (2008)
- [99] Qu, B., Zhong, W., Prince, R.: Interfacial coupling in ferroelectric superlattices. *Phys. Rev. B* **55**, 11218-11224 (1997)
- [100] Qu, W., Wlodarski, W., Meyer, J.U.: Comparative study on micromorphology

- and humidity sensitive properties of thin-film and thick-film humidity sensors based on semiconducting MnWO_4 . *Sens. Actuat. B Chem.* **64**, 76-82 (2000)
- [101] Ramesh, R., Spaldin, N.A.: Multiferroics: progress and prospects in thin films. *Nat Mater* **6**, 21-29 (2007)
- [102] Ramirez, A.P.: Colossal magnetoresistance. *J. Phys. Condens. Matter* **9**, 8171-8199 (1997)
- [103] Ren, S.Q., Weng, L.Q., Song, S.H., Li, F., Wan, J.G., Zeng, M.: $\text{BaTiO}_3/\text{CoFe}_2\text{O}_4$ particulate composites with large high frequency magnetoelectric response. *J. Mater. Sci.* **40**, 4375-4378 (2005)
- [104] Resta, R.: Macroscopic polarization in crystalline dielectrics: the geometric phase approach. *Rev. Mod. Phys.* **66**, 899-915 (1994)
- [105] Röntgen, W.C.: Ueber die durch Bewegung eines im homogenen electrischen Felde befindlichen Dielectricums hervorgerufene electrodynamische Kraft. *Ann. Phys.-berlin* **271**, 264-270 (1888)
- [106] Ryu, J., Priya, S., Uchino, K., Kim, H.E.: Magnetolectric effect in composites of magnetostrictive and piezoelectric materials. *J. Electroceram.* **8**, 107-119 (2002)
- [107] Salvador, P.A., Doan, T.D., Mercey, B., Raveau, B.: Stabilization of YMnO_3 in a perovskite structure as a thin film. *Chem. Mater.* **10**, 2592-2595 (1998)
- [108] Sergienko, I.A., Dagotto, E.: Role of the Dzyaloshinskii-Moriya interaction in multiferroic perovskites. *Phys. Rev. B* **73**, 94434 (2006)
- [109] Seshadri, R., Hill, N.A.: Visualizing the role of Bi 6s “lone pairs” in the off-center distortion in ferromagnetic BiMnO_3 . *Chem. Mater.* **13**, 2892-2899 (2001)
- [110] Shishidou, T., Mikamo, N., Uratani, Y., Ishii, F., Oguchi, T.: First-principles study on the electronic structure of bismuth transition-metal oxides. *J. Phys. Condens. Matter*, **16**, S5677-S5683 (2004)
- [111] Singh, M.K., Ryu, S., Jang, H.M.: Polarized Raman scattering of multiferroic BiFeO_3 thin films with pseudo-tetragonal symmetry. *Phys. Rev. B* **72**, 132101 (2005)
- [112] Singh, M.P., Prellier, W.: Oxide superlattices for multiferroics: opportunities, issues, and challenges. *Phil. Mag. Lett.* **87**, 211-222 (2007)
- [113] Slutsker, J., Levin, I., Li, J., Artemev, A., Roytburd, A.L.: Effect of elastic interactions on the self-assembly of multiferroic nanostructures in epitaxial films. *Phys. Rev. B* **73**, 184127 (2006)
- [114] Slutsker, J., Tan, Z., Roytburd, A.L., Levin, I.: Thermodynamic aspects of epitaxial self-assembly and magnetoelectric response in multiferroic nanostructures. *J. Mater. Res.* **22**, 2087-2095 (2007)
- [115] Smolenskii, G.A., Ioffe, V.A.: Communications Nr. 71 du Colloque International de Magnétisme, Grenoble (1958)
- [116] Smolensky, G.A., Isupov, V.A., Krainik, N.N., Agranovskaya, A.I.: Concerning the coexistence of the ferroelectric and ferrimagnetic states. *Isv. Akad. Nauk SSSR, Ser Fiz* **25**, 1333 (1961)

- [117] Sosnowska, I., Neumaier, T.P., Steichele, E.: Spiral magnetic ordering in bismuth ferrite. *J. of Phys. C Solid State Phys.* **15**, 4835-4846 (1982)
- [118] Srinivasan, G., DeVreugd, C.P., Flattery, C.S., Laletsin, V.M., Paddubnaya, N.: Magnetoelectric interactions in hot-pressed nickel zinc ferrite and lead zirconate titanate composites. *Appl. Phys. Lett.* **85**, 2550-2552 (2004)
- [119] Stein, S., Wuttig, M., Viehland, D., Quandt, E.: Magnetoelectric effect in sputtered composites. *J. Appl. Phys.* **97**, 10Q301 (2005)
- [120] Subramanian, M.A., He, T., Chen, J., Rogado, N.S., Calvarese, T.G., Sleight, A.W.: Giant room-temperature magnetodielectric response in the electronic ferroelectric LuFe_2O_4 . *Adv. Mater.* **18**, 1737-1739 (2006)
- [121] Sugawara, F., Iiida, S., Syono, Y., Akimoto, S.: Magnetic properties and crystal distortions of BiMnO_3 and BiCrO_3 . *J. Phys. Soc. Jpn.* **25**, 1553-1558 (1968)
- [122] Teague, J.R., Gerson, R., James, W.J.: Dielectric hysteresis in single crystal BiFeO_3 . *Solid State Commun.* **8**, 1073-1074 (1970)
- [123] Thiele, C., Drr, K., Bilani, O., Rdel, J., Schultz, L.: Influence of strain on the magnetization and magnetoelectric effect in $\text{La}_{0.7}\text{A}_{0.3}\text{MnO}_3/\text{PMN-PT}$ (001) (A=Sr, Ca). *Phys. Rev. B* **75**, 54408 (2007)
- [124] Timoshenko, S., Young, D.H.: *Vibration problems in engineering*. Van Nostrana (1955)
- [125] Tokura, Y., Nagaosa, N.: Orbital physics in transition-metal oxides. *Science* **288**, 462-468 (2000)
- [126] van den Boomgaard, J., Born, R.A.J.: A sintered magnetoelectric composite material $\text{BaTiO}_3\text{-Ni}(\text{Co}, \text{Mn})\text{Fe}_{34}$. *J. Mater. Sci.* **13**, 1538-1548 (1978)
- [127] van den Boomgaard, J., Terrell, D.R., Born, R.A.J., Giller, H.: An in situ grown eutectic magnetoelectric composite material. Part I: Composition and unidirectional solidification. *J. Mater. Sci* **9**, 1705-1709 (1974)
- [128] van Suchtelen, J.: Product properties: a new application of composite materials. *Philips Res. Rep.* **27**, 28-37 (1972)
- [129] Verwey, E.J.W., Haayman, P.W.: Electronic conductivity and transition point of magnetite. *Physica* **8**, 979-987 (1941)
- [130] Waintal, A., Chenavas, J.: Transformation sous haute pression de la forme hexagonale de $\text{MnT}'\text{O}_3$ (T'=Ho, Er, Tm, Yb, Lu) en une forme perovskite. *Mater. Res. Bull.* **2**, 819-822 (1967)
- [131] Wan, J., Zhang, H., Wang, X., Pan, D., Liu, J., Wang, G.: Magnetoelectric CoFe_2O_4 -lead zirconate titanate thick films prepared by a polyvinylpyrrolidone-assisted sol-gel method. *Appl. Phys. Lett.* **89**, 122914 (2006)
- [132] Wang, B., Woo, C.H.: Magnetoelectric effects due to elastic coupling in ferroelectric/ferromagnetic multilayers. *J. Appl. Phys.* **103**, 124107 (2008)
- [133] Wang, J., Neaton, J.B., Zheng, H., Nagarajan, V., Ogale, S.B., Liu, B., Viehland, D., Vaithyanathan, V., Schlom, D.G., Waghmare, U.V.: Epitaxial BiFeO_3 multiferroic thin film heterostructures. *Science* **299**, 1719-1722 (2003)
- [134] Wang, K.F., Liu, J.M., Ren, Z.F.: Multiferroicity: the coupling between

- magnetic and polarization orders. *Adv. Phys.* **58**, 321-448 (2009)
- [135] Weng, L., Fu, Y., Song, S., Tang, J., Li, J.: Synthesis of lead zirconate titanate-cobalt ferrite magnetoelectric particulate composites via an ethylenediaminetetraacetic acid-citrate gel process. *Scripta Mater.* **56**, 465-468 (2007)
- [136] Wilson, H.: On the electric effect of rotating a dielectric in a magnetic field. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character* **204**, 121-137 (1905)
- [137] Wu, T., Zurbuchen, M.A., Saha, S., Wang, R.V., Streiffer, S.K., Mitchell, J.F.: Observation of magnetoelectric effect in epitaxial ferroelectric film/manganite crystal heterostructures. *Phys. Rev. B* **73**, 134416 (2006)
- [138] Xu, G., Hiraka, H., Shirane, G., Li, J., Wang, J., Viehland, D.: Low symmetry phase in (001) BiFeO₃ epitaxial constrained thin films. *Appl. Phys. Lett.* **86**, 182905 (2005)
- [139] Yakel, H.L.: On the structures of some compounds of the perovskite type. *Acta Crystall.* **8**, 394-398 (1955)
- [140] Yakel, H.L., Koehler, W.C., Bertaut, E.F., Forrat, E.F.: On the crystal structure of the manganese (III) trioxides of the heavy lanthanides and yttrium. *Acta Crystall.* **16**, 957-962 (1963)
- [141] Yamasaki, Y., Miyasaka, S., Kaneko, Y., He, J., Arima, T., Tokura, Y.: Magnetic reversal of the ferroelectric polarization in a multiferroic spinel oxide. *Phys. Rev. Lett.* **96**, 207204 (2006).
- [142] Yan, L., Li, J., Viehland, D.: Deposition conditions and electrical properties of relaxor ferroelectric Pb(Fe_{1/2}Nb_{1/2})O₃ thin films prepared by pulsed laser deposition. *J. Appl. Phys.* **101**, 104107 (2007)
- [143] Yang, C.H., Koo, T., Lee, S.H., Song, C., Lee, K.B., Jeong, Y.: Orbital ordering and enhanced magnetic frustration of strained BiMnO₃ thin films. *Europhys. Lett.* **74**, 348 (2006)
- [144] Yang, J.J., Zhao, Y.G., Tian, H.F., Luo, L.B., Zhang, H.Y., He, Y.J., Luo, H.S.: Electric field manipulation of magnetization at room temperature in multiferroic CoFe₂O₄/Pb(Mg_{1/3}Nb_{2/3})_{0.7}Ti_{0.3}O₃ heterostructures. *Appl. Phys. Lett.* **94**, 212504 (2009)
- [145] Zgür, ü., Alivov, Y., Morko, H.: Microwave ferrites, part 2: passive components and electrical tuning. *J. Mater. Sci. Mater. Electron.* **20**, 911-952 (2009)
- [146] Zhai, J., Xing, Z., Dong, S., Li, J., Viehland, D.: Magnetoelectric laminate composites: an overview. *J. Am. Ceram. Soc.* **91**, 351-358 (2008)
- [147] Zhai, J., Xing, Z., Dong, S., Li, J., Viehland, D.: Thermal noise cancellation in symmetric magnetoelectric bimorph laminates. *Appl. Phys. Lett.* **93**, 072906 (2008)
- [148] Zhang, J.X., Li, Y.L., Schlom, D.G., Chen, L.Q., Zavaliche, F., Ramesh, R., Jia, Q.X.: Phase-field model for epitaxial ferroelectric and magnetic nanocomposite thin films. *Appl. Phys. Lett.* **90**, 052909 (2007)
- [149] Zhang, N., Liang, D., Schneider, T., Srinivasan, G.: Is the magnetoelectric

- coupling in stickup bilayers linear? *J. Appl. Phys.* **101**, 083902-083905 (2007)
- [150] Zhao, K., Dai, Y.R., Wan, J.G., Zhang, Z.F., Zhu, J.S.: The study of magnetic field-induced strain of $\text{Ni}_2\text{MnGa/Pb}(\text{Zr}_{0.52}\text{Ti}_{0.48})\text{O}_3$ composite. *Mater. Sci. Eng. A* **438**, 1019-1021 (2006)
- [151] Zheng, H., Wang, J., Lofland, S.E., Ma, Z., Mohaddes-Ardabili, L., Zhao, T., Salamanca-Riba, L., Shinde, S.R., Ogale, S.B., Bai, F.: Multiferroic $\text{BaTiO}_3\text{-CoFe}_2\text{O}_4$ nanostructures. *Science* **303**, 661-663 (2004)
- [152] Zheng, H., Wang, J., Mohaddes-Ardabili, L., Wuttig, M., Salamanca-Riba, L., Schlom, D.G., Ramesh, R.: Three-dimensional heteroepitaxy in self-assembled $\text{BaTiO}_3\text{-CoFe}_2\text{O}_4$ nanostructures. *Appl. Phys. Lett.* **85**, 2035-2037 (2004)
- [153] Zhong, X.L., Wang, J.B., Liao, M., Huang, G.J., Xie, S.H., Zhou, Y.C., Qiao, Y., He, J.P.: Multiferroic nanoparticulate $\text{Bi}_{3.15}\text{Nd}_{0.85}\text{Ti}_3\text{O}_{12}\text{-CoFe}_2\text{O}_4$ composite thin films prepared by a chemical solution deposition technique. *Appl. Phys. Lett.* **90**, 2903 (2007)

Dielectric Breakdown of Microelectronic and Nanoelectronic Devices

9.1 Introduction

As great progress has been achieved in micro-fabrication and micro-processing technologies, the size of electronic devices has been scaled down to sub-micrometers. This is especially embodied in the Si metal-oxide-semiconductor (MOS)-based devices, e.g., chips and memories, obeying the well known Moore's law, with the number of transistors on an integrated circuit doubling approximately every two years. To minimize the short-channel effect, the thickness of dielectric oxide in a MOS device is approaching a few nanometers (Stathis and DiMaria, 1999; Frank et al., 2001). In recent years, some basic functional components of many microelectronic devices, such as dielectric oxides in a MOS device, have been scaled down to the nanoscale in one or more dimensions.

Actually, it would be more appropriate to call them nanoelectronic devices.

The technological evolution leads to great enhancement in the performance of the devices. A smaller device has a higher electric field and thus operates at a higher speed. For a perfect single device, smaller size also means lower power consumption. However, it also pushes some basic functional components of devices to their physical limits, causing a reliability problem. As a consequence, device designers need to make a tradeoff between performance and reliability. Generally speaking, there are several crucial issues relating to the physical limits of microelectronic and nanoelectronic devices. They are (1) the critical size for maintaining the bulk property, (2) dielectric breakdown, (3) power dissipation, (4) thermal noise, (5) the Heisenberg uncertainty principle, etc. (Chiabrera et al., 1989). These would become bottlenecks for further scaling. In tackling these problems, one of the most formidable challenges is the unpredictability of the dielectric breakdown. Take MOS field-effect transistors (MOSFETs) as an

example. Such devices make use of the insulating property of dielectric oxide to modulate the transport of carriers. When a breakdown occurs in the dielectric oxide, probably due to external electric fields, it becomes conductive so as to cause the failure of the device. Dielectric breakdown is a crucial issue directly related to the lifetime of electronic devices, their performance and operating range. Therefore, many theoretical and experimental works have been devoted to this field, aiming to find out the exact mechanism of dielectric breakdown and improve the performance of electronic devices.

A MOS device is one of the most common electronic devices in microelectronics. Its basic structure is the metal-oxide-semiconductor structure, which is widely used in logic circuits, dynamic random access memory (DRAM) circuits, electrically erasable programmable read-only memory (EEPROM), FLASH memory, etc. In a MOS structure, silicon dioxide (SiO_2) or silicon oxynitride (SiO_xN_y) layers are usually used as dielectric films due to their outstanding insulating properties and convenient fabrication process. They are called gate oxides as in MOSFETs. With continual scaling down, the silicon oxide is approaching its physical limits, manifested by increasing breakdown frequency and a larger leakage of current. Recently, more and more papers have focused on finding appropriate materials with higher dielectric constants, so-called high- k dielectrics, such as $\text{HfO}_2/\text{HfSiO}_x$ (Zafar et al., 2002; Umezawa et al., 2005; Okada et al., 2007), Al_2O_3 (Park et al., 2000; Zafar et al., 2002), TiO_2 (Kadoshimaa et al., 2003), ZrO_2 (Gusev et al., 2001) and LaAlO_3 (Luo et al., 2009), to replace the silicon oxide as the dielectric oxide. Nevertheless, most dielectric breakdown studies still focus on the silicon oxide breakdown problem nowadays, especially on SiO_2 films. To have an overall understanding of the history of the dielectric breakdown problem of the MOS structure, interested readers are referred to the papers by Solomon (1977), Green et al. (2001), Dumin (2001), Azizi and Yiannacouras (2003), Ghetti (2004), Lombardo et al. (2005), etc.

Owing to the quite wide range of research areas of dielectric breakdown issues, we discuss only dielectric breakdown of MOS devices and contribute the main content of this chapter to dealing with the dielectric breakdown problems of SiO_2 . Most of the theories and methods mentioned here are still applicable to dielectric breakdown problems of other structures and materials. Therefore, in the following, if not mentioned, the dielectric oxide or gate oxide refers to SiO_2 . Some reliability issues of high- k materials (e.g., LaAlO_3) and ferroelectric materials (e.g., PbTiO_3) are also discussed due to their increasing importance. Moreover, we mainly tackle the breakdown of thin film structures.

Due to its complexity, especially for difficulty in direct and real-time observation, although many models have been proposed, such as Klein and Solomon (1976) for an impact ionization breakdown model, McPherson and Baglee (1985a; 1985b) for a thermal chemical breakdown model, Chen et al. (1985; 1986) for a hole injection breakdown model, Cartier et al. (1995) and Suñé and Wu (2001; 2004) for a hydrogen release breakdown model and Cheung (1999; 2001) for a unified model, the mechanism of dielectric breakdown in MOS devices is still far from clear. Nevertheless, people have found that there were four

basic characteristics of dielectric breakdown in MOS devices. (1) The first one is that dielectric breakdown is a complicated evolution process with the electric field, thermal field and mechanical stress field intercoupling with each other. (2) The second characteristic is its high sensitivity to external fields. Usually, the time to breakdown depends on the field intensity exponentially. Due to such high sensitivity, a small variation in fields could cause a large variation in the time to breakdown (T_{BD}). Nevertheless, it also indicates great controllability of the performance of the device. Modeling the impact of external fields on dielectric breakdown and finding the so-called field accelerating factors is one of the key research issues in this field. (3) Thirdly, dielectric breakdown is most likely a defect-participating process. Unlike the high field time-zero breakdown of defect-free samples, most dielectric breakdown events occur at much lower fields where the generation and redistribution of defects play an important role. The microscopic origin of these defects is an intensively debated issue. (4) Finally, dielectric breakdown is a statistical event. For example, two similar devices fabricated and tested in the same conditions show a different time to breakdown. This is mainly due to the high sensitivity of dielectric breakdown to fields and the inhomogeneity of the samples. What statistical distributions the dielectric breakdown phenomena should follow is still not clear.

People have paid attention to the gate oxide breakdown problem since the 1960's due to the wide application of MOS devices. The main motivation was to develop quantitative theories to predict the reliability of thin oxide film devices. It has been found that oxide breakdown under not too high stress was a consequent event of the following process, so-called time dependent breakdown (TDBD). (1) As external stress was applied, defects were gradually generated in the oxide and the leakage current increased. (2) The density of defects increased by enough to form a conductive percolation passage connecting the electrodes and led to a sudden jump in current; i.e., breakdown occurred. (3) If the current was large enough the oxide inevitably wore out. Therefore, to describe the whole process of breakdown, we need to model each sub-process, including current leakage, defect generation, defect percolation and the final breakdown of the oxide. These processes do not seem to occur one after another with a well defined time boundary; they are coupled with each other. For example, a hot carrier of tunneling current could generate defects; such defects could help tunneling (DiMaria and Cattier, 1995; Takagi et al., 1999). Many models have been proposed for such processes.

As the dielectric film is scaled down to sub 1.5 nm, as has already happened in MOS devices, tunneling current becomes significant even in normal operating conditions. There are at least three reasons for fully understanding tunneling phenomena in thin oxide MOS devices. Firstly, a significant tunneling current in MOS devices results in large stand-by power consumption. Secondly, traditional capacitance versus voltage (CV) measurements in thin oxide MOS devices is severely disturbed by the large tunneling current. In such a case, we have to conduct tunneling measurements instead of traditional CV measurements for extracting device physical parameters, e.g., oxide thickness, doping levels, etc.

Furthermore, tunneling current is directly related to the reliability of MOS devices, as it provides a large number of carriers to cause defect generation. Therefore, it is necessary to accurately understand this phenomenon.

Theoretical models of tunneling had been established by many researchers (Kane, 1961; Bardeen, 1961; Harrison, 1961; Prange, 1963). Traditional Fowler-Nordheim models (Fowler and Nordheim, 1928; Lenzlinger and Snow, 1969) usually implied a metal gate and a triangular barrier. They were not accurate enough to model tunneling of thin oxide MOS devices. Subsequent studies proposed tunneling models using different tunneling current components and the quantum effects both in the barrier and in the accumulation or inversion layers (Stern and Howard, 1967; Baraff and Appelbaum, 1972; Pals, 1972; Ando, 1976; Suñé et al., 1991; Ghetti et al., 2000). Modeling tunneling current is the basis for the current-induced-breakdown models. Due to its importance in breakdown phenomena, an individual section is devoted to the tunneling mechanism in this chapter. Another important issue about tunneling in dielectrics is its controllability through changing factors such as thickness, mechanical stress and polarization, which is also discussed in this section.

Defect generation is also called trap generation. It is common to consider that such a process is responsible for oxide degradation. Compared with the tunneling process, although many models have been proposed, the exact mechanism of defect generation is still not quite clear. Whether defect generation is mainly caused by an electric field or hot carriers is still under debate. This is mainly due to the lack of experimental observations of defects on the microscopic scale although first-principle calculations have shown that oxygen vacancy or hydrogen related defects may be the best candidates (Blöchl and Stathis, 1999a; 1999b; Luo et al., 2009). Actually this is the key problem of the dielectric breakdown phenomena.

Among various defect generation models, the three most popular models are the anode hydrogen release model (AHR model) (Cartier et al., 1995; Suñé and Wu, 2001; 2004), anode hole injection model (AHI model or $1/E$ model) (Chen et al., 1985; 1986; Lee et al., 1987; Hu and Lu, 1999) and the thermochemical model (E model) (McPherson and Baglee, 1985a; 1985b; McPherson, 1986; McPherson and Mogul, 1997; 1998). All of them are able to explain certain experimental results; nevertheless, all undergo some unsolved contradictions. The AHR model and AHI model are quite similar except for the different emphasis of the roles of hydrogen and holes in the defect generation. They both attribute defect generation to carriers (electrons and/or holes). The E model considers the field-enhanced-bond-breakage as the cause of the breakdown and current plays a less important role. Therefore, the AHR model and AHI model can be categorized as the current-induced-breakdown models while the E model can be categorized as the field-induced-breakdown model. In this chapter, these models will be introduced in detail. We are not going to justify which model is right or wrong but try to express the ideas of all these models to the readers. What must be pointed out is that, although it is quite common to accept that dielectric breakdown is a process with electric, thermal and mechanical stress fields intercoupling with each other, the

effects of mechanical stress and the thermal field have attracted poor attention. A comprehensive defect generation model is still lacking. Nevertheless, a thermal breakdown model and a mechanical-stress-enhanced breakdown model will also be given.

Percolation models for dielectric breakdown developed by Suñé and his coworkers (Suñé et al., 1990; Suñé and Wu, 2001), Degraeve et al. (1995; 1998; 2000) and Stathis (1999) could reproduce the statistical features of breakdown. Such models could determine the critical defect density, the slope of Weibull distribution and their dependence on oxide thickness. There were only two parameters in the simplest percolation model, i.e., the oxide thickness and the effective defect size. Results also indicated that not all the defects were responsible for triggering breakdown. Although percolation models apparently succeeded in many aspects, there remain several questions for further research. Firstly, there are still no direct experiments which show the existence of a percolation passage prior to breakdown. Secondly, microscopic analyses of defects are needed to determine the effective size of defects rather than by data fitting. Finally, how these defects are distributed in the oxide is still ambiguous. An individual description of the statistical aspects of oxide breakdown is given in this chapter.

By using mainly transmission electron microscopy (TEM) and emission microscopy (EMMI), the damage from breakdown in a MOS structure has been investigated (Lombardo et al., 1998; Radhakrishnan et al., 2001; Tung et al., 2002; Tang et al., 2004). It was found that the degree of damage depended on the gate current/voltage compliance level during the voltage/current stress. Polarity-dependent dielectric breakdown-induced epitaxy (DBIE) was observed even at the earliest state of breakdown. Rapture passages consisting of cavities were also found through the gate and Si substrate, which indicated oxide melting and evaporation in this zone during breakdown. In severe cases, not only the oxide but also the gate, channel, source and drain were substantially damaged. All of this damage indicated that heat dispersion played a vital role. This description is given in the sixth section.

As space is limited, the experimental methods of dielectric breakdown are not included in this chapter. Interested readers are referred to these papers (Ghetti et al., 2000; Green et al., 2001; Heh et al., 2003; Wu et al., 2002). In summary, this chapter is arranged as follows. In the next section, some basic concepts of dielectric breakdown will be given. In the third section, different tunneling mechanisms, including electron tunneling, hole tunneling and trap assisted tunneling, will be presented. The microscopic origin of defects from first-principle calculation will be also included in this part. In the fourth part we will illustrate the main ideas of the AHI model, AHR model and the E model. Some other breakdown models will also be briefly mentioned in this part. The intention of this section is to give readers the picture that dielectric breakdown is a complicated process coupling with electric, mechanical and thermal fields. Following that, the fifth section is devoted to the statistical and percolation models of breakdown. The sixth section is about some important experimental observations of the damage

after breakdown. Finally, some remarks will be given to summarize this topic. Some imperative problems and promising applications of dielectric breakdown in the future will be pointed out there.

9.2 Basic Concepts

In this section some basic concepts in dielectric breakdown are introduced, including MOS structure, tunneling modes, breakdown modes, defect generation, stress induced leakage current (SILC), holes generation and energetics of defects. The content in this part is quite easy to understand; therefore, readers familiar with the breakdown issue can skip over it.

MOS structure and MOSFET represent the basic structure and the basic device to illustrate the dielectric breakdown throughout the chapter. Therefore, an introduction to MOS structure is given first. Devices operate in different operating modes under different stress conditions, e.g., accumulation and inversion. Such modes determine the carrier source of the tunnel current and therefore have a great impact on breakdown. In this section we just briefly introduce such tunneling modes, and detailed tunneling models and formulations are given in the next section. Defect generation is an important process in low stress/long term dielectric breakdown. In this section, we also introduce some physical quantities in describing defect generation. Those defect generation models are given in the fourth section of this chapter.

9.2.1 MOS Structure

Metal–oxide–semiconductor (MOS) structure, a typical representative of metal–insulator–semiconductor structure (MIS), is widely used in microelectronic and nanoelectronic devices such as MOSFETs. The oxide is usually silicon dioxide due to its easy fabrication and high dielectric performance. Other oxides may also be used, such as SiO_xN_y and the high- k materials HfO_2 , TiO_2 and LaAlO_3 . In the following, if not mentioned, the oxide refers to SiO_2 . Such a MOS structure is obtained by growing a layer of silicon dioxide (SiO_2) on top of a silicon substrate and depositing a layer of metal or polycrystalline silicon (SiO_2 as the gate oxide, poly-Si gate is more commonly used nowadays than metal gate). In order to further reduce the sheet resistance of the gate electrode, usually a thin metal silicide layer, such as Co, Ti, W and Ni-silicide, is deposited upon the poly-Si.

A MOSFET is based on a MOS capacitance between a body electrode and a gate electrode. Using the gate electrode together with source and drain electrodes, it can modulate charge concentration below the oxide. Taking an n-MOSFET for illustration, as shown in Fig. 9.1, such a structure usually consists of a gate electrode, a SiO_2 gate oxide, a p-doped Si substrate acting as a body electrode and

two highly n-doped regions (n^+) using a diffusion technique. The two doped regions are called drain and source according to their voltage. A p-MOSFET is formed when the Si substrate is n-doped and the source and drain regions are p^+ -doped. There are four terminals connecting to the external circuit, i.e. the gate electrode, base electrode, source electrode and drain electrode. Voltage with various polarities and magnitudes can be applied to modulate the operating modes of the transistor. A complementary MOS (CMOS) structure is formed by integrating an n-MOSFET and a p-MOSFET together as a basic structure.

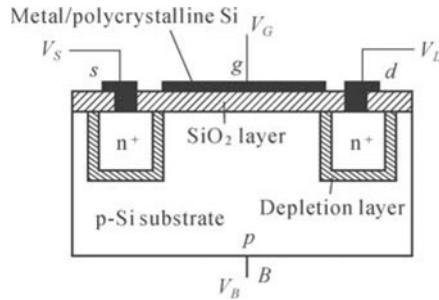


Fig. 9.1 Schematic picture of cross section of an n-MOSFET

As depicted in Fig. 9.2, when there is no applied voltage at the source, drain and substrate, i.e., they are connected to the earth and a positive voltage ($V_G > 0$) at the gate, a depletion layer with immobile negative charges appears near the SiO_2 -Si interface as the holes in the p-Si substrate are pushed out of this region by the positive electric field. When V_G increases, the electric field near the SiO_2 -Si interface is so strong that minority carriers in the p-Si substrate, i.e., electrons, are attracted to form an inversion layer where the density of the minority carriers exceeds that of the majority carriers. To get an inversion layer, the electric field near the interface must be sufficiently strong to bend the bands so much that the conduction (valence) band edge comes below (above) the Fermi level for an n-MOSFET (p-MOSFET). This inversion layer forms a conductive channel between the drain and source.

When an n-MOSFET undergoes a negative voltage at the gate ($V_G < 0$), the majority carriers holes will be attracted by the negative electric field near the SiO_2 -Si interface. A thin accumulation layer where the density of the majority carriers is much higher than that in the bulk will be formed. Although this is not the normal operational mode of a MOSFET, people often use this mode to investigate the tunneling of majority carriers. A similar accumulation layer of electrons can be formed in a p-MOSFET when a positive enough gate voltage is applied.

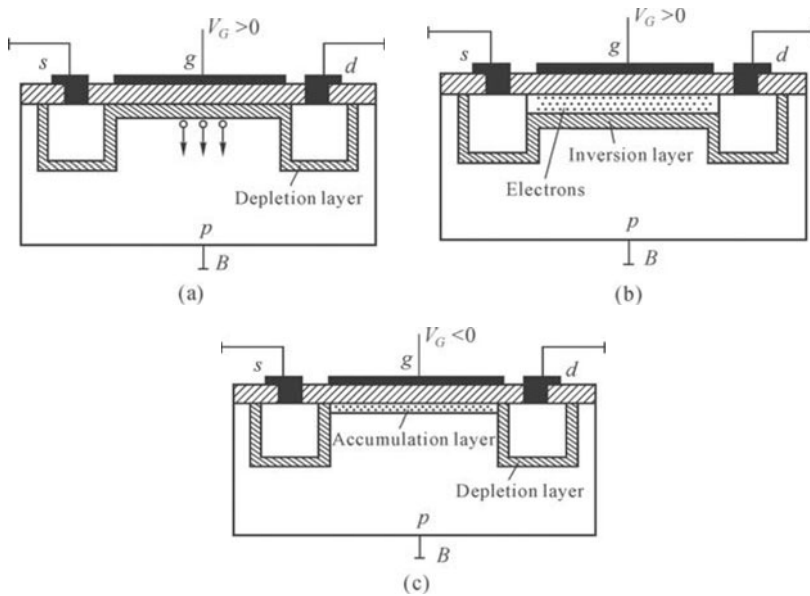


Fig. 9.2 Schematic illustration of the forming of (a) a depletion layer, (b) inversion layer and (c) accumulation layer when an n-MOSFET undergoes different voltage at the gate with the other three terminals connected to the earth

9.2.2 Different Tunneling Modes

As the gate voltage is large and the oxide film is thin enough, the tunneling current would be dominant across the MOS structure. As depicted in Fig. 9.3 and Fig. 9.4, there are four basic tunneling modes according to the different operational modes in the MOSFET structure. They are holes tunneling from the accumulation layer, electrons tunneling from the inversion layer, holes tunneling from the inversion layer and electrons tunneling from the accumulation layer, respectively. Hole tunneling is also called valence band electrons tunneling.

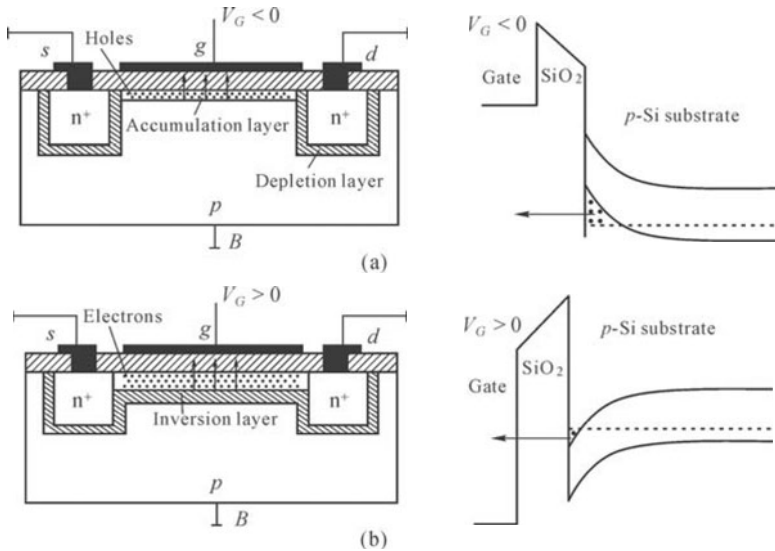


Fig. 9.3 Tunneling modes in an n -MOSFET. (a) Holes tunneling from accumulation layer; (b) Electrons tunneling from inversion layer

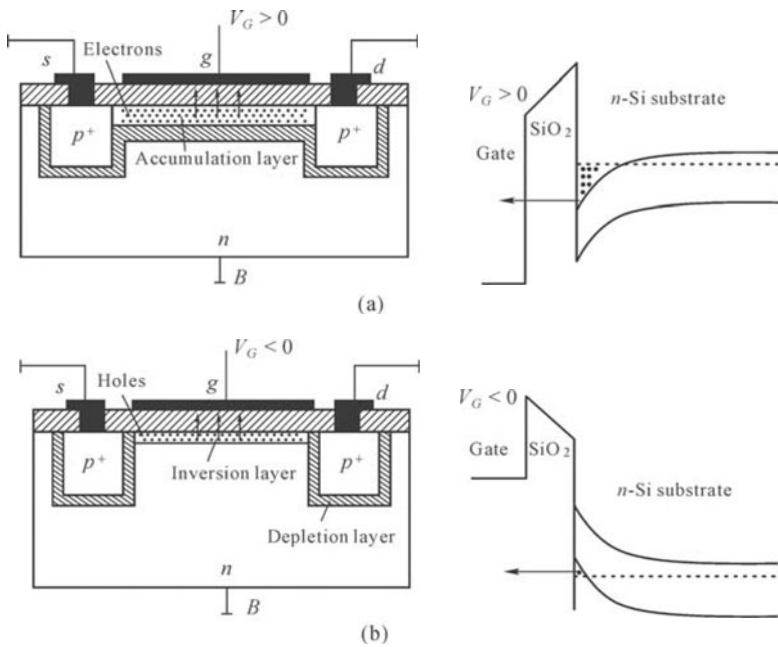


Fig. 9.4 Tunneling modes in a p -MOSFET. (a) Holes tunneling from accumulation layer; (b) Electrons tunneling from inversion layer

In general, tunneling can be classified into direct tunneling (DT) and Fowler-Nordheim tunneling (FNT). Direct tunneling occurs in a low field regime, where electrons (holes) tunnel throughout the whole oxide directly from one electrode to the other. For Fowler-Nordheim tunneling, the electric field is high enough to bend the tunneling barrier into a triangular shape, and electrons (holes) tunnel into the oxide conduction (valence) band before entering the anode (cathode) contact.

Another important tunneling mode intimately related to dielectric breakdown is trap assisted tunneling (TAT). In TAT, traps play the role of stepping stones which assist the electrons (holes) to tunnel into or through the oxide. Since stressing the gate oxide results in traps generation, this tunneling component probably accounts for the obvious increased tunneling current through the gate oxide after electric stress. Fig. 9.5 schematically depicts the three tunneling modes. For cursory evaluation, as the barrier height of the n-silicon and SiO_2 is about 3 eV, for thin oxide ($t_{\text{ox}} < 3$ nm) at gate voltage lower than 3 V, the tunneling is dominated by DT. For thicker oxide, DT is very unlikely to happen and the tunneling is dominated by FNT or TAT. A theory of tunneling, with consideration of the quantum effects in the inversion and accumulation layers, will be given in detail in Section 9.3.

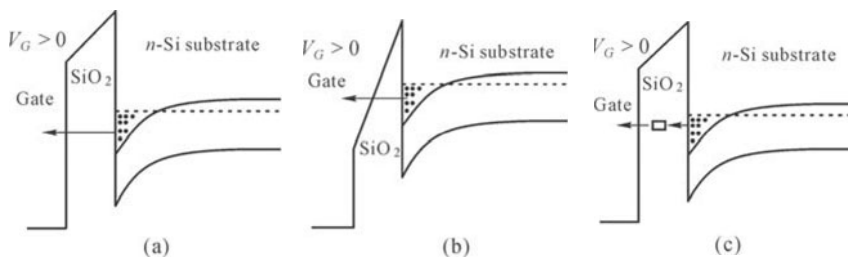


Fig. 9.5 Schematic illustration of (a) direct tunneling (b) FN tunneling and (c) trap-assisted tunneling in a MOS structure with an n-Si substrate

9.2.3 Dielectric Breakdown Modes

When a dielectric, such as gate oxide, receives an external voltage or current, it will gradually or suddenly (according to the magnitude of stress and the strength of the dielectric) lose its insulating property, manifesting an increasing leakage current. Such a phenomenon is called a dielectric breakdown. While the electric field acts as a leading factor, other fields such as a thermal field, mechanical stress, etc., together with the electric field, are coupled with each other in the breakdown process. We call these external fields stress fields and the device experiences external fields as under stress.

A dielectric breakdown is generally attributed to the defects generation and percolation during the stress (Fig. 9.6). At the early stage, defects are generated sporadically within the bulk or at the interface (Fig. 9.6a). They also called traps

as they can trap electrons or holes. When the oxide is under stress for a longer time, traps become more and more and some of them connected with each other to form clusters (Fig. 9.6b). At the same time, the leakage current continues to increase. When one or some of the clusters form a passage connecting both the electrodes, this conductive passage leads to a sudden increase in current and breakdown occurs (Fig. 9.6c). We call this the first breakdown as other clusters may continue to form more conductive passages and lead to a second breakdown. If the leakage current flowing through the passage is so large that the local thermal heat released by it cannot be promptly balanced, the local temperature of the passage is so high as to melt and evaporate the material in and near the passage; thus the so-called hard breakdown occurs (Fig. 9.6d).

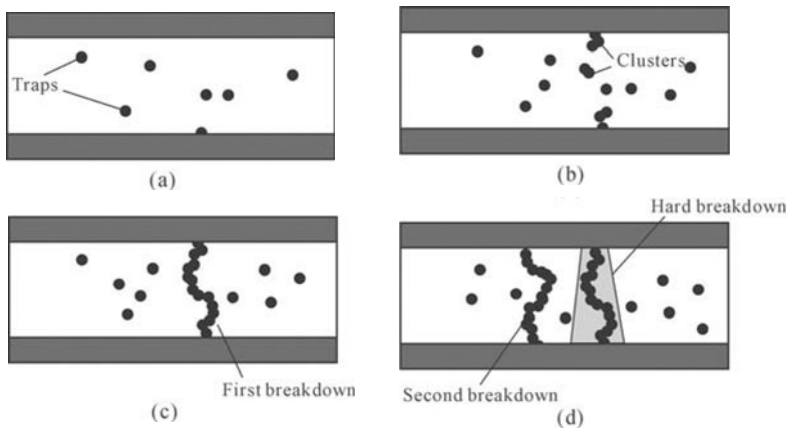


Fig. 9.6 Schematic pictures of different stages of dielectric breakdown. (a) Traps are created sporadically at the early stage or already exist before stress; (b) Cluster forms as traps connect with each other; (c) A conductive passage forms and causes the first breakdown; (d) More passages form and hard breakdown occurs if the heat flow is so large

According to whether the defects are externally introduced or intrinsically generated, dielectric breakdown can be categorized into extrinsic breakdown and intrinsic breakdown. Defects can be introduced at the stage of fabrication, such as from particulates, oxide thinning, stacking faults, metal impurities, precipitates, hydrogen species, etc. We call these defects extrinsic defects. With advanced fabrication technologies and better shielding, extrinsic breakdown due to such defects can be suppressed to the large extent. Defects can also be generated intrinsically by holes or electrons, probably by impacting the lattice and forming some broken bonds, and thus is essentially intrinsic and inevitable.

Researchers may naturally focus on intrinsic breakdown. However, the definition of intrinsic and extrinsic breakdown makes more literal sense than practical sense. In experiments, we can never be sure that whether the observed breakdown is actually intrinsic or not. In the current fabrication technology, hydrogen is a common impurity for both Si and SiO₂ and is introduced during the fabrication stage in order to reduce the density of interface states by passivating

dangling bonds. It is unreasonable to ignore the effect of hydrogen when almost all MOS devices are fabricated by this process. Actually, many experiments indicate that hydrogen plays an important role in defect generation of the silicon oxide film (Gale et al., 1983; Feigl et al., 1984; Buchanan et al., 1993). Therefore, the hydrogen participating breakdown seems difficult to avoid. This is the basic motivation for the establishment of the AHR model. Compared with the AHR model, the AHI model attributes the injection of holes to the cause of defect generation and the E model considers intrinsic bond breakage as the culprit of breakdown. Therefore, both of them are considered as intrinsic breakdown models.

Dielectric breakdown is also divided into time-dependent dielectric breakdown (TDDB) and time-zero dielectric breakdown (TZDB). TDDB takes a period of time before the oxide breaks down after the stress is applied, while TZDB refers to breakdown that suddenly occurs just after the stress is applied. TDDB testing is often used as an experimental method to evaluate the quality of MOS devices, since the testing condition of this kind of experiment is closer to the normal operating condition of the devices. Since, at the operating conditions, the time to breakdown, T_{BD} , of a dielectric film meeting the industrial requirement is about ten years, it is impractical to conduct such a long term experiment. TDDB testing is a kind of accelerating experiment with quite higher stress than the normal operating condition. To find exactly the field accelerating factors is a crucial concern of TDDB. TZDB occurs at an even higher stress condition and such a kind of testing is often used to investigate the intrinsic electric strength of the dielectric films. Fig. 9.7 shows the result of a TDDB testing spanning a period of five years (Wu et al., 2002), indicating the strong sensitivity of stress conditions.

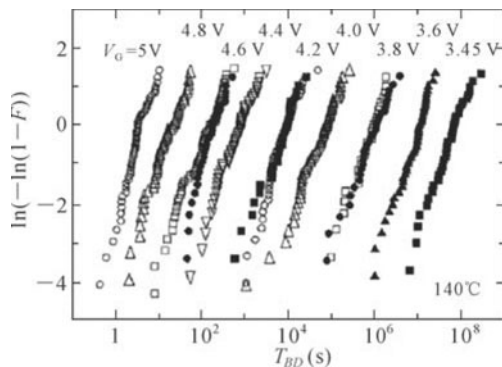
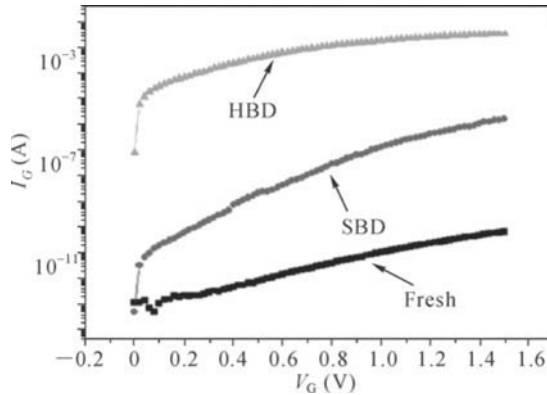


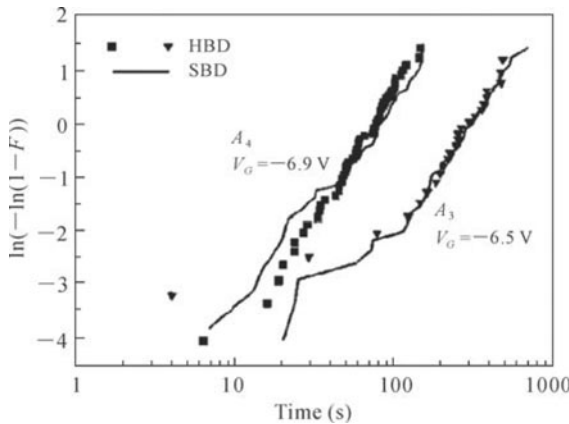
Fig. 9.7 Distributions of various stress voltage at the same temperature from a long term TDDB testing, where F is the cumulative failure probability. Reprinted with permission from Wu et al. (2002). Copyright (2002) by IEEE

Many researchers also classify dielectric breakdown as soft breakdown (SBD) and hard breakdown (HBD) according to the damage to the insulating material caused by the breakdown event. Soft breakdown, also called pre-breakdown, quasi breakdown, early breakdown or stress-induced-leakage current (SILC) B mode

(Ghetti, 2004), is characterized by a significant increase in leakage current and its reversible damage to the oxide. Soft breakdown often occurs in ultra thin films and/or at low stress. On the contrary, hard breakdown, which often occurs in thicker films and/or at higher stress, has an even steeper increase in current and causes complete irreversible damage to the oxide. One of the distinguishing characteristics between soft breakdown and hard breakdown is the current-voltage dependence after breakdown. The IV curve of the device after soft breakdown can be described by power law, i.e., $I = G_0 V^\delta$, whereas that after hard breakdown manifests an ohmic dependence, as shown in Fig. 9.8a.



(a)



(b)

Fig. 9.8 Comparison of some characteristics of HBD and SBD. (a) Different dependences of gate current on gate voltage before and after breakdown for a 2-nm gate oxide in inversion mode. Reprinted with permission from Tang et al. (2004). Copyright (2004) by IEEE. (b) First breakdown event statistical distributions with different device area and stress. Reprinted with permission from Suñé et al. (2000). Copyright (2000) by IEEE

TDDDB and TZDB, or soft breakdown and hard breakdown, may be considered as limits of each other. However, we must keep in mind that the mechanisms may be quite different, since they take place under different stress conditions resulting in quite different energy distributions of the carriers. Carriers having different energy may trigger different energy dependent processes (DiMaria et al., 1992). Nevertheless, it is now generally accepted that soft and hard breakdown originate from the same precursor defects, and can be described by the same Weibull statistics (Suñé et al., 2000) as depicted in Fig. 9.8b. Alam et al. (2002a; 2002b) proposed a model based on the assumption that the severity of breakdown depended on the power dissipation rather than on any physical difference between precursor defects and explained some experimental results quite well.

9.2.4 Defect Generation

9.2.4.1 Defect Generation Rate

Breakdown is considered a process consisting of defect generation due to injected carriers or bond breakage. For current-induced breakdown models, it is convenient to define the defect generation rate, i.e., density of the generated defect per injected carrier, to describe defect generation as

$$P_g = \frac{dN_d}{dN_c}, \quad (9.2.1)$$

where N_d and N_c are the density of the defect and injected carrier, respectively. Both of them are time dependent. P_g is also called defect generation probability. Using the relationship $N_c = Q_c / q$, where Q_c is the total charge of the injected carriers and q is the charge of a single carrier, Eq. (9.2.1) can be rewritten as

$$P_g = \frac{q dN_d}{dQ_c}, \quad (9.2.2)$$

When breakdown occurs, we often define three quantities, i.e., the charge of carrier to breakdown, Q_{BD} , the density of defect to breakdown, N_{BD} , and the time to breakdown, T_{BD} . N_{BD} is also called the critical density of defects. Q_{BD} may be a more suitable quantity to measure than T_{BD} , since it relates breakdown to the injected current. They are related by

$$Q_{BD} = \int_0^{t_{BD}} J dt, \quad (9.2.3)$$

where J is the current density during the stress.

It is easy to see that Q_{BD} is related to N_{BD} by the defect generation rate P_g from Eq. (9.2.2). When P_g is independent of time, we can obtain the following simple relation of N_{BD} and Q_{BD} by integrating Eq. (9.2.2):

$$N_{BD} = P_g Q_{BD} / q, \quad (9.2.4)$$

or more generally

$$N_{BD} = \frac{1}{q} \int_0^{Q_{BD}} P_g dQ. \quad (9.2.5)$$

It is observed that in a large variety of oxide thicknesses stressed over a wide range of voltage, Q_{BD} is inversely related to the initial defect generation rate as indicated by Eq. (9.2.4) or Eq. (9.2.5) (DiMaria et al., 1993; Dumin et al., 1995; Stathis and DiMaria, 1998). These experimental results support the current-induced breakdown models. It should be pointed out that it is commonly observed that the relationship of N_{BD} and Q_{BD} is nonlinear, typically sigmoidal (DiMaria and Stathis, 2001), which indicates the defect generation rate is a function of the injected current.

For the thermochemical model (E model), defect generation is considered as an electric field-enhanced reaction to bond breakage. It is natural to define a reaction rate constant, k , as the defect generation rate, i.e., density of generated defect per second. The reaction rate depends on the field-modulated enthalpy of activation, $\Delta H(E)$, as follows:

$$k = \nu_0 \exp \left[-\frac{\Delta H(E)}{k_B T} \right], \quad (9.2.6)$$

where ν_0 is a characteristic collision (interaction) frequency, k_B is the Boltzmann constant, E is the electric field and T is the temperature.

9.2.4.2 The Energy Source of Defect Generation

The energy of the applied field is mainly carried by the electrons. Electrons gain energy and release some of the energy through interaction with the lattice and interface before entering into the anode. The released energy can excite phonons. If the energy is high enough, it can also break bonds or generate secondary electrons and holes by impact ionization. Other processes may also occur by making use of such energy. We call all energy releasing processes using a joint

name as energy dissipation. These processes are probably responsible for breakdown. Therefore, breakdown is regarded as an energy driven process. The energy stored in a perfect MOS structure without leakage current undergoing stress V is

$$E = \int V dQ, \quad (9.2.7)$$

where Q is equal to the displacement charge Q_{disp} , which is stored at the electrodes. This energy is completely reversible since it equals the ideal capacity of the MOS structure.

When there is leakage current, the energy is dissipated by the irreversible process of charge transport and recombination. In this case, the total charge is equal to the leakage component and the displacement component. This amount of dissipated energy is denoted by E_{leak} . The residual energy E_{disp} is the reversible one stored in the displacement charge. E_{leak} is the energy that can be exploited by defect generation and E_{disp} is the stored energy which would be released in the breakdown transient. Therefore, the energy exploited by defect generation is

$$E_{\text{leak}} = E_{\text{total}} - E_{\text{disp}} = E_{\text{total}} - \frac{1}{2} Q_{\text{disp}} V_{\text{BD}} = \int V dQ - \frac{1}{2} Q_{\text{disp}} V_{\text{BD}}, \quad (9.2.8)$$

where V_{BD} is the voltage just before breakdown. Fig. 9.9 schematically depicts the energy components of a MOS structure during stress at constant voltage.

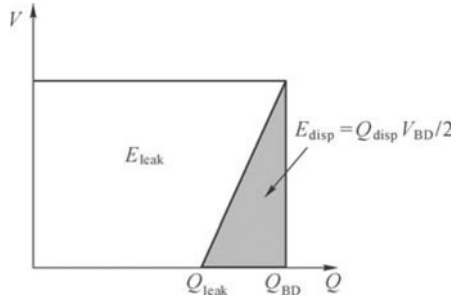


Fig. 9.9 Schematic illustration of energy components of a MOS structure during stress at constant voltage

9.2.5 Basic Statistical Concepts of Dielectric Breakdown

A dielectric breakdown is a statistical event. Devices prepared under the same conditions usually have different T_{BD} , Q_{BD} and N_{BD} . Theoretically, if we test enough samples in the stress condition in the same way as in the normal operating

condition, we can get a quite exact distribution of T_{BD} . As already seen in Fig. 9.7 and Fig. 9.8b, we often plot the cumulative failure rate as a function of stress time to reflect such a distribution. For a given cumulative failure rate, F , we can define a $T_{BD}(F)$. For example, $T_{BD}(0.63)$ means that after stressing the samples for a period of $T_{BD}(0.63)$, 63% of samples have broken down. Then we can easily predict the probability of when devices fabricated under the same conditions will undergo breakdown.

However, it is almost impossible to do such experiments since, under operating conditions, it may take about ten years to observe failures for those samples meeting the industrial requirements. Furthermore, it is impractical to test an enormous number of samples. What is usually done is to test a finite number ($\sim 10^2$) of samples under higher stress conditions and/or higher temperature. We call such experiments voltage and temperature accelerating experiments. Due to the relatively small amount of samples, a higher cumulative failure rate F is often used to define the T_{BD} , denoted as T_{test} . To get the T_{BD} at a much lower failure rate at operating conditions, denoted as T_{real} , we need to make a percentile scaling. Moreover, we usually test samples with a larger area than those of the real devices, such as chips. Since a larger area means a larger number of weak zones, area scaling is necessary.

If the thickness of tested samples is different from that of a real device, thickness scaling is also needed. Generally speaking, we need to use five procedures to use the test data to predict T_{real} . They are voltage scaling (projection), temperature scaling (projection), percentile scaling (projection), area scaling (projection) and thickness scaling (projection). We call these procedures reliability projections.

Often we choose the Weibull or log-normal distribution ‘a priori’ to describe the statistical behavior of dielectric breakdown. That means we use the distribution function to fit the data and then project it to the practical operating conditions by using the above five procedures. As breakdown is more likely a weakest-spot problem of extreme value statistics, Weibull distribution has been found to be more suitable than log-normal distribution, which can give rise to over-optimistic projections. The probability density function of a Weibull distribution is defined as (Weibull, 1951)

$$f(t; \alpha, \beta) = \frac{\beta}{\alpha} t^{\beta-1} \exp\left(-\frac{t^\beta}{\alpha}\right), \quad (9.2.9)$$

where t is the variable which can be time, position and whatever else, with α and β being two parameters. As here in a dielectric breakdown, we denote t as stressing time, the respective cumulative distribution function is

$$F(t; \alpha, \beta) = 1 - \exp\left(-\frac{t^\beta}{\alpha}\right). \quad (9.2.10)$$

It is convenient to introduce the so-called Weibull number

$$W \equiv \ln[-\ln(1-F)] = \beta \ln t - \ln \alpha. \quad (9.2.11)$$

If we plot W versus $\ln t$, we will get a straight line with a slope of β and intercept $-\ln \alpha$. Experiments have shown that the slope β is related to the oxide thickness, and the intercept $-\ln \alpha$ is related to the oxide area and stress (see, e.g., Wu et al., 1999). Assuming that the breakdown obeys the Weibull distribution, it is easy to see that percentile projection can be described by the following relation:

$$\frac{T_{\text{real}}}{T_{\text{test}}} = \left[\frac{\ln(1-F_{\text{test}})}{\ln(1-F_{\text{real}})} \right]^{1/\beta}, \quad (9.2.12)$$

where F_{test} and F_{real} are the cumulative failure rate of the test sample and the real device, respectively. The area projection formula can be demonstrated (Stathis, 1999) as

$$\frac{T_{\text{real}}}{T_{\text{test}}} = \left(\frac{A_{\text{test}}}{A_{\text{real}}} \right)^{1/\beta} \quad (9.2.13)$$

where A_{test} and A_{real} are the area of the test sample and real device, respectively.

As for voltage, temperature and thickness projection, the projection relations are still not clear. Actually, finding the voltage/temperature accelerating factors and the thickness dependence of T_{BD} is one of the key issues of dielectric breakdown. Fig. 9.10 schematically illustrates the process for reliability projections using the Weibull distribution a priori. What we must emphasize is that such a projection is very sensitive to the distribution parameters. Therefore, this can imply quite a large error in the prediction. Actually, this is called the projection problem.

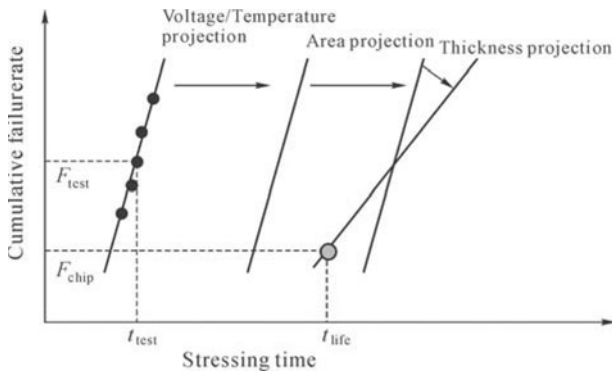


Fig. 9.10 Schematic illustration of projection procedure

9.2.6 Stress Induced Leakage Current

It was first observed by Maserjian and Zamani (1982) in experiments that the leakage current increased when the gate oxide film was under stress or after stress, as compared with the leakage current before stress. We call this increasing component of leakage current the stress-induced-leakage current (SILC). To conduct SILC experiments, we often measure the leakage current before stress J_0 , then stress the gate oxide in the high-field regime where Fowler-Nordheim tunneling is usually dominant, and measure the leakage current at low-field direct-tunneling regime J_s . The normalized current increase is defined by

$$\frac{\Delta J}{J_0} = \frac{J_s - J_0}{J_0}. \quad (9.2.14)$$

SILC is considered to be caused by defect generation during the stress. These generated defects can act as stepping stones to assist the tunneling of electrons and increase the leakage current. This process is the so-called trap-assisted tunneling. Compared with other possible defect component(s), such as positively-charged sites near the anode, the neutral electron traps are more likely responsible for SILC, as shown by DiMaria (1995).

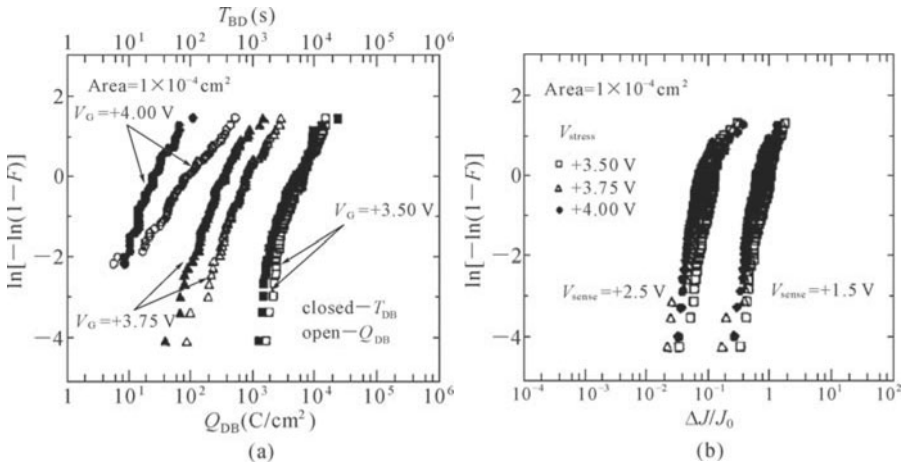


Fig. 9.11 Comparison of failure distributions of (a) and (b) at various stress conditions. Reprinted with permission from Buchanan et al. (1997). Copyright (1997) by Elsevier

SILC experiments are often implemented to measure the density of interface states and traps together with the capacitance-voltage (CV) method. Compared with Q_{BD} and T_{BD} , whose distributions are dependent on the magnitude and polarity of the bias, the distribution of $\Delta J/J_0$ at the breakdown point is found to be relatively independent of the stress condition (as shown in Fig. 9.11). Since measure-

ment of Q_{BD} or T_{BD} represents the charge or time needed to trigger breakdown, their voltage dependence indicates they are not only related to the critical defect density but also to the rate of defect generation. On the other hand, $\Delta J / J_0$ is independent of the rate of defect generation, thus providing a better measure of the critical defect density (Buchanan et al., 1997). A quantitative model of SILC based on the trap-assisted tunneling mechanism will be presented in the fourth section of this chapter.

9.2.7 Holes Generation

Experiments have shown that not only the anode but also the cathode region is damaged after breakdown. Electrons are probably not responsible for the damage to the cathode, since they have low energy at the cathode. This indicates that positive mobile charges might be released by electrons in the bulk oxide or at the interface, and then travel back to the cathode with high enough energy to damage it. Based on this supposition and the fact that holes can be generated by a high field through either impact ionization in the bulk oxide or injection from the anode, people consider holes as one of the likely candidates responsible for oxide breakdown.

9.2.7.1 Impact Ionization in the Oxide

Secondary electrons and holes can be generated in the bulk insulator by hot carrier current through impact ionization. For an oxide layer sandwiched by two electrodes, assuming that the injected carriers are electrons with a current density of J (it can be easily extended to the case of injection of holes), the density of holes is p , and then evolution of p is described by the following equation (Klein, 1972; Klein and Solomon, 1976):

$$\frac{dp}{dt} = \frac{J}{q} \alpha - \frac{J}{q} \sigma p - \frac{\mu_p E p}{l}, \quad (9.2.15)$$

where the terms on the right of the equation represent ionization, recombination and electric drift, respectively, α is the coefficient of ionization per length, σ is the recombination cross section, l is the thickness of the oxide, q is the unit charge of the electron, μ_p is the mobility of holes and E is the electric field in the oxide.

Generally speaking, α , σ , μ_p are field dependent coefficients. Usually, σ is assumed as a constant, with α and μ_p taking the following forms:

$$\alpha = \alpha_0 \exp(-H/E), \quad (9.2.16)$$

$$\mu_p = CE \exp(-D/E), \quad (9.2.17)$$

where α_0 , H , C , D are constants.

Early models of dielectric breakdown of SiO_2 emphasized the roles of impact ionization. These models proposed that hot current could induce impact ionization and generated a large amount of electron-hole pairs. The holes would be left behind in the bulk to form space charges due to their low mobility or the out-of-balance of ionization and recombination. These space charges enhance the electric field in the oxide and conversely increase the injection of hot current and trigger more ionization. Finally, as the current loses its stability, breakdown occurs. The electric field affected by the space charge is described by Poisson's equation. For example, for one dimension we have

$$\frac{\partial E}{\partial x} = \frac{qp}{\epsilon_r \epsilon_0}, \quad (9.2.18)$$

where ϵ_0 and ϵ_r are the permittivity of vacuum and the relative dielectric constant of the oxide, respectively. The injected current is usually assumed to be an FN injection, thus

$$J = AE_c^2 \exp\left(-\frac{B}{E_c}\right), \quad (9.2.19)$$

where A and B are constants and E_c is the electric field at the cathode.

Models emphasizing the terms of recombination in Eq. (9.2.15) are usually called impact-recombination models (IR models) (DiStefano and Shatzkes, 1974; 1975; 1976; Shatzkes and Av-Ron, 1976; Klein and Solomon, 1976), and those emphasizing electric drift are called impact-drift models (ID models) (O'Dwyer, 1969; Klein, 1972). Impact ionization in the bulk SiO_2 band gap occurs only when the energy of the hot electron is larger than the band gap energy (~ 9 eV). However, when the oxide film becomes thinner, hot electrons are rare and this mechanism seems insignificant. Thus it causes speculation about the applicability of such models in a thin film/low field regime.

9.2.7.2 Holes Injection from Anode

It has been observed that injected electrons could introduce a positive charge and fast interface state in the oxide since the early 1970's. ID and IR models have been proposed to explain the experimental observations; however, they are not applicable in a low field regime. For injected carriers with lower energy, the hole

current is still observed. It is attributed to the so-called anode-hole-injection (AHI) mechanism, where holes come from the anode rather than the bulk oxide.

For MOSFETs, the anode is usually a p-silicon or n-silicon substrate where holes can accumulate at the interface. For example, for a p-MOSFET, when V_G is negative enough (see Fig. 9.3a), the density of holes can be much higher than in bulk silicon as they are attracted by the negative field near the interface, forming the accumulation layer. Holes can tunnel from the accumulation layer if the oxide is thin enough.

Holes injected from the anode can also be caused by impact ionization in the anode by hot electrons entering the anode (as depicted in Fig. 9.12a where a p-MOSFET undergoes a negative gate voltage.). The electrons enter the Si substrate with a minimum energy of $q\phi_b$, where ϕ_b is the potential drop from SiO_2 to Si substrate. They lose energy by exciting phonons in both the oxide and the substrate and by impact ionization in the substrate.

Such holes can be detected by so-called charge separation experiments. It is explained in Fig. 9.12b. Due to the transverse electric field of the depletion region and the longitudinal electric field induced by V_G , the generated hole-electron pairs are separated. If the tunneling of holes from the substrate to the n^+ poly Si gate is negligible (V_G is not too large and the oxide is not too thin), most holes drift to the drain and source, and the generated electrons and the original injected electrons together drift to the electrode of the substrate.

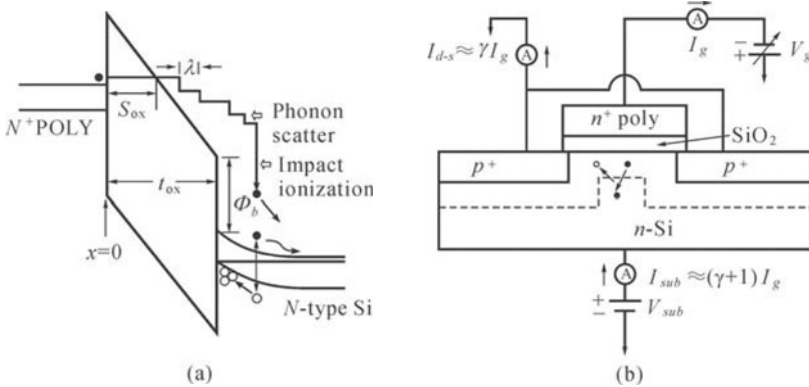


Fig. 9.12 Scheme of (a) impact ionization in silicon substrate and (b) the charge separation technique to measure the quantum yield. Reprinted with permission from Chang et al. (1985). Copyright (1985) by the American Institute of Physics

Denoting $\gamma(E)$ as the quantum yield of impact ionization, i.e., the number of electron-hole pairs generated by a hot carrier of energy E , then the average quantum yield can be determined as the ratio of the current of the source and drain terminals I_{ds} to the current of the gate terminal I_g , i.e.,

$$\langle \gamma \rangle = \frac{I_{ds}}{I_g}. \quad (9.2.20)$$

If the oxide film is thick, electrons injected into the oxide (by FN injection) would experience phonon scattering before they enter the substrate. Due to the randomizing effect, the change of energy distribution of electrons is difficult to know. Therefore it is difficult to determine the quantum yield at a given energy for thick films. However, if the oxide film is thin enough, electrons cross the oxide by direct tunneling without scattering in the conduction band of the oxide. The energy distribution does not change. Therefore, we can determine $\gamma(E)$ more exactly in thin film.

Another plausible source of hole current at low field was that proposed by Fischetti (1985a; 1985b). He assumed that the electrons gain an appropriate energy in the oxide and lost the energy at the anode to excite surface plasma oscillations. These collective oscillations decay into hot electron-hole pairs. Some of the holes may tunnel back to the oxide and get trapped there.

9.2.7.3 Separating Holes From Electrons

In order to investigate the effect of holes on breakdown, it is necessary to separate the current of holes from the current of electrons. Fig. 9.12b shows such kinds of charge separation techniques. However, in that case the electron current is still dominant across the oxide, since electrons have barriers of about 3 or 4 eV while holes have 5 or 6 eV. Due to different interfacial energy barriers of the carriers, it is more difficult to inject holes than electrons. One method suitable for ultrathin oxide film is to use the so-called electron-tunneling-blocking technique (DiMaria and Stathis, 2001) as shown in Fig. 9.13. However, this kind of injection can occur only on p-channel field-effect-transistors (p-FETs).

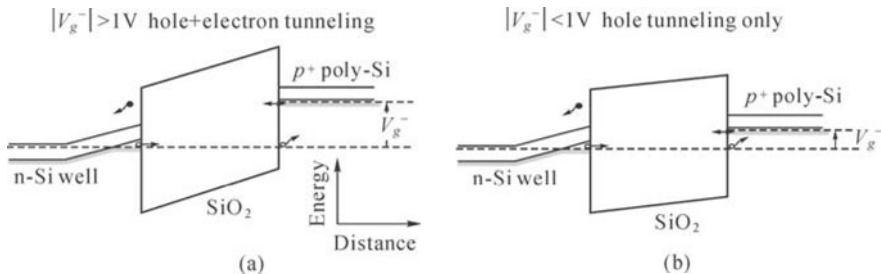


Fig. 9.13 Schematic energy band for an ultrathin gate oxide structure showing tunneling and injection of thermal carriers from appropriate bands biased under different inversion conditions. (a) Both holes and electrons can tunnel through the oxide. (b) Electron tunneling is blocked since the conduction band of the well is higher than the valence-band of poly-Si. Reprinted with permission from DiMaria and Stathis (2001). Copyright (2001) by the American Institute of Physics

9.2.8 Energetics of Defects

By using the first-principles calculation and a knowledge of thermodynamics, the energetics of various defects in the oxide layer can be analyzed. Systematic analysis of the energetics of possible defects in oxide can provide us with an overview of which defect is the most promising one responsible for dielectric breakdown. In general, we must answer the following two questions. (1) In what conditions is the existence of a defect energetically favorable in the oxide? (2) In what conditions, is the defect-participating-breakdown process, such as trapping and detrapping electrons or holes, energetically favorable? To answer these questions, we usually calculate the formation energy and the reaction energy of the defect. These energies depend on the surrounding environments and stress conditions, e.g., a reductive or oxidative environment with or without stress. As we will see in the following, only those defects that have suitable formation energy and reaction energy are probably the ones responsible for dielectric breakdown.

9.2.8.1 Formation Energy

Defect formation energy usually refers to the energy needed to form a defect in the thermodynamic equilibrium. It can be calculated from the total energy of the super cell with defects, following the approach described in previous reports (Zhang and Northrup, 1991; Tanaka et al., 2003). Such an approach treats the atomic system as a grand canonical ensemble where atoms can be removed from, or added into, the system with definite chemical potentials.

For a defect α with charge q in the oxide, its formation energy $E_f(\alpha, q)$ depends on the atomic chemical potentials and also varies with the electron chemical potential, i.e., the Fermi energy. It is given by

$$E_f(\alpha, q) = E_{sc}(\alpha, q) - \sum_i n_i \mu_i + q(\varepsilon_F + E_{\text{VBM}}). \quad (9.2.21)$$

In this expression, $E_{sc}(\alpha, q)$ represents the total energy of the defective super cell, n_i is the atomic number of atom i in the defective super cell and μ_i is the corresponding chemical potential of atom i . At $T = 0$, the entropy contribution is neglected, and the Fermi energy ε_F is measured from the top of the valence band. E_{VBM} is the one-electron energy of the valence-band maximum (VBM). Eq. (9.2.21) can also be expressed as follows:

$$E_f(\alpha, q) = E_{sc}(\alpha, q) - (E_{\text{perf}} - \sum_j n_j \mu_j) + q(\varepsilon_F + E_{\text{VBM}}), \quad (9.2.22)$$

where E_{perf} is the total energy of the perfect super cell, and n_j and μ_j are the atomic number and chemical potential of atom j which is removed from the perfect super cell to form a defect α , respectively.

It is expected that the E_{VBM} of a defect containing super cell be different from that of the perfect super cell due to the band-gap distortion caused by the defect, as defects can distort the band structure significantly (Laks et al., 1992; García et al., 1995; Pöykkö et al., 1996; Mattila et al., 1998). Therefore, it is necessary to account for this distortion and line up a band structure of the defective super cell with the perfect super cell. It is assumed that the potentials of atoms in the perfect super cell are similar to those far from a defect in the defective super cell. Then the shift of the VBM in a defective super cell can be obtained by using a macroscopic averaging technique (Baldereschi et al., 1988; Peressi et al., 1998), which calculates the average electrostatic potential difference between the defect-containing super cell and the perfect super cell. In the finite-size super cell calculation, these treatments can compensate for the error caused by Coulomb interaction between the charged defects and the background image charge. The E_{VBM} is computed as

$$E_{\text{VBM}} = E_{\text{VBM}}^{\text{perfect}} + V_{\text{av}}^{\text{defect}} - V_{\text{av}}^{\text{perfect}}, \quad (9.2.23)$$

where $E_{\text{VBM}}^{\text{perfect}}$ is the one-electron energy of the VBM of the perfect super cell, and $V_{\text{av}}^{\text{defect}}$ and $V_{\text{av}}^{\text{perfect}}$ are the average potentials of the defective super cell and perfect super cell, respectively.

$E_{\text{VBM}}^{\text{perfect}}$ can be calculated as

$$E_{\text{VBM}}^{\text{perfect}} = E_{\text{perf}}(0) - E_{\text{perf}}(+1), \quad (9.2.24)$$

where $E_{\text{perf}}(q)$ is the total energy of a perfect super cell with charge q .

Fermi energy ε_{F} varies in the range of band gap value E_{g} . E_{g} is defined as the energy difference between the highest occupied and lowest unoccupied one-electron states. Therefore, it is the energy difference between E_{VBM} and the one-electron energy of the conduction-band minimum E_{CBM} , i.e.,

$$E_{\text{g}} = E_{\text{CBM}}^{\text{perfect}} - E_{\text{VBM}}^{\text{perfect}} = [E_{\text{perf}}(-1) - E_{\text{perf}}(0)] - [E_{\text{perf}}(0) - E_{\text{perf}}(+1)]. \quad (9.2.25)$$

The chemical potentials of the atomic species cannot be obtained directly. We can determine them from various equilibrium conditions of the multi-component systems. Here we would take the high- k material LaAlO_3 as an example (Luo et al., 2009). It is assumed that LaAlO_3 is always stable, and the chemical potentials of the three atomic species can vary to satisfy the following conditions:

$$\mu_{\text{La}} + \mu_{\text{Al}} + 3\mu_{\text{O}} = \mu_{\text{LaAlO}_3}, \quad (9.2.26)$$

and

$$\mu_{\text{La}} \leq 0, \quad \mu_{\text{Al}} \leq 0, \quad \mu_{\text{O}} \leq 0, \quad (9.2.27)$$

where μ_{LaAlO_3} is the generalized free formation energy per formula unit of the perfect LaAlO_3 . Fig. 9.14 shows the phase diagram of the ternary La-Al-O system, which is derived from the combination of the chemical potentials of Al, Al_2O_3 , La, La_2O_3 , and O_2 . The vertices of the big triangle correspond to three dominant component chemical potentials and the compounds are placed on the edge according to their stoichiometric proportion. Points A–E denote the five possible chemical-potential combinations and the intersections of these five points (dark color zone) define the thermodynamical equilibrium region of the ternary system under different chemical potentials. At point A, for example, LaAlO_3 is in equilibrium with Al_2O_3 and O_2 ; thus the atomic chemical potential at this point is also fulfilled by the following relation, i.e.,

$$\text{Point A:} \quad 2\mu_{\text{Al}} + 3\mu_{\text{O}} = \mu_{\text{Al}_2\text{O}_3}, \quad \mu_{\text{O}} = \mu_{\text{O}(\text{bulk})}. \quad (9.2.28)$$

Similarly, for other points, we have the following relations, i.e.,

$$\text{Point B:} \quad 2\mu_{\text{La}} + 3\mu_{\text{O}} = \mu_{\text{La}_2\text{O}_3}, \quad \mu_{\text{O}} = \mu_{\text{O}(\text{bulk})}, \quad (9.2.29)$$

$$\text{Point C:} \quad 2\mu_{\text{La}} + 3\mu_{\text{O}} = \mu_{\text{La}_2\text{O}_3}, \quad \mu_{\text{La}} = \mu_{\text{La}(\text{bulk})}, \quad (9.2.30)$$

$$\text{Point D:} \quad \mu_{\text{Al}} = \mu_{\text{Al}(\text{bulk})}, \quad \mu_{\text{La}} = \mu_{\text{La}(\text{bulk})}, \quad (9.2.31)$$

$$\text{Point E:} \quad 2\mu_{\text{Al}} + 3\mu_{\text{O}} = \mu_{\text{Al}_2\text{O}_3}, \quad \mu_{\text{Al}} = \mu_{\text{Al}(\text{bulk})}. \quad (9.2.32)$$

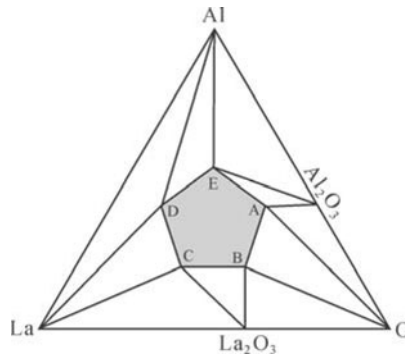


Fig. 9.14 The schematic phase diagram of the La-O-Al ternary system. Reprinted with permission from Luo et al. (2009). Copyright (2009) by American Physical Society

Table 9.1 shows the calculated results of the formation energies of oxygen related defects in LaAlO_3 . In order to derive the formation energy in the dilute limit, the effects of super cell size convergence on the formation energy are investigated.

Table 9.1 Defect-formation energies of various isolated defects in LaAlO_3 in different super cell sizes (20, 40, 90, and 135 atoms). The chemical potential is fixed at oxidizing conditions in the calculations. Reprinted with permission from Luo et al. (2009). Copyright (2009) by American Physical Society

Defect	Super cell size and formation energies (eV)				Defect	Super cell size and formation energies (eV)			
	20	40	90	135		20	40	90	135
O_i^0	6.409	4.601	4.437	4.404	V_O^0	6.727	6.464	6.408	6.545
O_i^{1-}	6.620	4.837	4.471	4.464	V_O^{1+}	6.593	6.544	6.562	6.475
O_i^{2-}	3.621	4.696	4.431	4.392	V_O^{2+}	3.141	1.805	0.996	1.084

9.2.8.2 Charge State Levels

Defects play important roles in the degradation of oxide as they can act as stepping stones to assist electrons (holes) to tunnel through the oxide. Such a trap-assisted tunneling (TAT) is implemented by trapping and detrapping electrons or holes. During tunneling, the energy level of defects will change as their charge states have changed. We usually use the charge-state level to describe the energy level of defects. It represents the stability of the charge state of the defects. In equilibrium, the charge-state level of a defect as a function of the Fermi level is called the thermodynamic charge-state level (Blöchl and Stathis, 1999a; 1999b; Blöchl, 2000). For example, when a neutral defect is charged by an electron from an electrode, the reaction is



The energy change in the reaction is

$$\Delta E = E(\alpha, -1) - E(\alpha, 0) - \varepsilon_\text{F} , \quad (9.2.34)$$

where the Fermi energy ε_F of the electrode defines the energy of the electron in the reservoir. $E(\alpha, -1)$ and $E(\alpha, 0)$ represent the energy of defective super cells with α^- and α^0 , respectively. Then the thermodynamic charge-state level is defined as

$$\varepsilon_{0/-} = E_{\text{relax}}(\alpha, -1) - E_{\text{relax}}(\alpha, 0). \quad (9.2.35)$$

The sub-index “relax” indicates that the energy must be determined with the structure individually relaxed. It is easy to see that the thermodynamic charge-state level can be regarded as reaction energy. The defect is most probably uncharged when the charge-state level is above the Fermi level and is most probably charged when the charge-state level is below the Fermi level.

The thermodynamic charge state level is not appropriate for describing non-equilibrium situations, such as tunneling. In such cases, we would like to use a switching charge-state level rather than a thermodynamic charge-state level. Since electron tunneling is much faster than atomic motion while atomic motion is much faster when compared with the lifetime of the electron in the defect, i.e.

$$t_{\text{tunnel}} \ll t_{\text{atom}} \ll t_{\text{reside}}, \quad (9.2.36)$$

the energy level variation must be computed for the structure of the initial charge state without relaxing the equilibrium structure. For a reaction described by Eq. (9.2.33), the switching charge-state level which is dependent on its initial charge state is defined as

$$\varepsilon_{\text{sw}}(0/-) = E_{\text{frozen}}(\alpha, -1) - E_{\text{frozen}}(\alpha, 0), \quad (9.2.37)$$

where the sub-index “frozen” means that we should compute the energy in the initial atomic configuration. In general, we have $\varepsilon_{\text{sw}}(0/-) \neq \varepsilon_{\text{sw}}(-/0) \neq \varepsilon_{0/-}$.

9.2.8.3 Electron Affinity and Hole Affinity

The charge-state level described above is appropriate to describe TAT in thin oxide films since it assumes that the electrodes are the reservoirs of electrons/holes. For a trapping process where the electrons/holes come from the conduction/valence band, it is more suitable to use another two quantities, i.e., the electron affinity and the hole affinity of the defect. The electron affinity of the defect is considered as the energy gain when the electron from the bottom of the conduction band is trapped at the defect. Assuming the initial and final states of defect α with charge q and $q-1$, respectively, the energy gain is equal to the difference between the initial state (sum of the energy of the perfect super cell with an additional electron and the energy of a defective super cell with charge q) and the final state (sum of the energy of the neutral perfect super cell and the energy of a defective super cell with charge $q-1$). Therefore, the electron affinity $\chi_e(\alpha, q)$ can be written as (Foster et al., 2001)

$$\chi_e(\alpha, q) = E_{\text{perf}}(-) + E(\alpha, q) - E_{\text{perf}}(0) - E(\alpha, q-1) + \kappa, \quad (9.2.38)$$

where $E_{\text{perf}}(-)$, $E_{\text{perf}}(0)$, $E_{\text{perf}}(0)$ and $E(\alpha, q-1)$ are defined as before. κ is the difference between the band gap from the DFT calculation and the experiment measurement, i.e.

$$\kappa = E_{\text{g}}^{\text{exp}} - E_{\text{g}}^{\text{DFT}}, \quad (9.2.39)$$

which is used to correct the DFT calculation error.

Similarly, the hole affinity of a defect is defined as the energy gain when a hole is trapped from the top of the valence band as follows:

$$\chi_h(\alpha, q) = E_{\text{perf}}(+) + E(\alpha, q) - E_{\text{perf}}(0) - E(\alpha, q+1). \quad (9.2.40)$$

Moreover, similar to the charge-state level, we can compute energies of the final state with the atomic configuration frozen as the initial one or relaxed. Such kind of affinity is called “vertical” and “relaxed” affinity, respectively.

9.3 Mechanism Analysis of Tunneling Phenomena in Thin Oxide Film

With continual scaling of an electronic device, the thickness of the dielectric oxide in a metal-insulator-metal capacitor is approaching a few nanometers. One of the formidable challenges in drastic downscaling is the large tunneling current through oxide layers. This large leakage current contributes a disruptive signal in the switchable circuit and makes the traditional CV measurements unreliable. Furthermore, the tunneling current also plays an important role in oxide degradation as the hot tunneling carriers trigger defect generation.

To reduce the leakage current, new materials such as high- k materials, strained Si, together with the corresponding metal gate technology, are introduced to replace the traditional silicon oxide structure. Since the high- k materials have larger dielectric constants, replacing the SiO_2 oxide layer with high- k materials would allow the MOS device to increase its capacitance without significantly reducing the thickness of the dielectric layer. As the thickness decreases to a few nanometers, neither the traditional SiO_2 nor the new high- k material layers under the external field could avoid the problem of large leakage current. Thus, it is a formidable task to understand the microscopic origin of the large leakage current and find the corresponding methods to tackle it.

Experimentally, many factors, such as the configuration of the insulator, active interface defects and the contacting electrodes, can influence the leakage current of the MOS devices. After intensive research activities relating to the tunneling

issue, the theory of tunneling became well established and some models also took account of the quantum effects in accumulation and inversion layers. In this section, tunneling models of the MOS structure, including inversion and accumulation situations, are introduced. Such models form the basis for further modeling of defect generation, especially for current-induced breakdown models and also provide an alternative method to determine the parameters of the oxide, such as barrier height and thickness.

9.3.1 Self-Consistent Schrödinger's and Poisson's Equations

As already shown in Fig. 9.3b and Fig. 9.4b, an inversion layer forms when the electric field at the interface is sufficiently strong to bend the energy bands so much that the conduction (valence) band edge comes below (above) the Fermi level. The strong electric field at the interface produces a potential well with a width comparable to the carrier wavelength in the perpendicular direction of the interface. Thus near the interface the energy level separates into sub-bands, which form the so-called 2D electron gas (this view was first postulated by Schrieffer (1957) and was confirmed in experiments by Fowler et al. (1966)). Similarly, 2D electron gas also forms at the accumulation layer (Fig. 9.3a and Fig. 9.4a). Such sub-bands effectively affect the tunneling feature through the oxide and other properties of the electronic transport, such as the mobility of electrons (Fang and Howard, 1966). Thus a careful treatment of such a quantum mechanical effect is demanded in analysis. Readers are referred to the excellent review paper written by Ando et al. (1982).

9.3.1.1 Effective-Mass Approximation

The idea of effective-mass approximation is that the effect of the periodic lattice potential on the behavior of electrons can be described by an effective mass tensor. That is to say, for those electrons around the minimum of the conduction band, they behave as if they have mass m^* . Neglecting the nonparabolicity and coupling between different conduction-band minima, the kinetic energy operator can be written as

$$T = \frac{1}{2} \sum_{i,j} w_{ij} p_i p_j, \quad (9.3.1)$$

where w_{ij} are the elements of the reciprocal effective mass tensor and

$p_j = -i\hbar \left(\frac{\partial}{\partial x_j} \right)$. As we are dealing with a MOS structure which is described in the

coordinate system with $z = 0$ plane located at the SiO_2/Si interface, w_{ij} and p_j are defined in the same coordinate system. In the local coordinate system where the coordinate axis is the principal axis of the bulk constant-energy ellipsoid, w_{ij} is reduced to $w'_{kk} = 1/m'_k$, where m'_k is the principal effective mass. Defining $\|a\|$ as the transformation matrix, we have

$$p_j = \sum_k a_{jk} p'_k, \quad (9.3.2)$$

$$w_{ij} = \sum_k a_{ik} a_{jk} w'_{kk}, \quad (9.3.3)$$

where a_{ij} are the elements of $\|a\|$. Using Eqs. (9.3.2) and (9.3.3) we can deal with situations of different orientations of the interface or surface, such as (001), (110) and (111) surfaces of Si, the bulk constant-energy ellipsoid and the external fields.

Under effective mass approximation, the electron wave function can be expressed as the product of the Bloch function at the bottom of the conduction band valley $u_\alpha(\mathbf{R})$ and an envelope function $F(x, y, z)$, i.e.

$$\psi(x, y, z) = F(x, y, z) u_\alpha(\mathbf{R}), \quad (9.3.4)$$

where $u_\alpha(\mathbf{R})$ is the Bloch function noted by the wave vector $\mathbf{k}_\alpha$ at the energy minimum. Then, Schrödinger's equation is reduced to the effective mass equation

$$[T - e\phi - E]F = 0, \quad (9.3.5)$$

where E is the energy and ϕ is the potential profile excluding the periodic atomic part.

By assuming that the potential $\phi(z)$ depends only on the perpendicular distance from the interface z , the envelope function can be further expressed as

$$F(x, y, z) = \xi(z) \exp(ik_1 x + ik_2 y), \quad (9.3.6)$$

where k_1 and k_2 are wave vectors in the k_x - k_y plane measured relative to the projection of $\mathbf{k}_\alpha$ on this plane. Substituting Eq. (9.3.6) into Eq. (9.3.5), and together with Eq. (9.3.1), we can get the only z -dependent equation, i.e.

$$-\frac{1}{2} w_{33} \hbar^2 \frac{d^2 \xi}{dz^2} - \hbar^2 (w_{13} k_1 + w_{23} k_2) \frac{d\xi}{dz} - [e\phi(z) + E'] \xi(z) = 0, \quad (9.3.7)$$

with the energy E' dependent on the z -direction motion, which satisfies

$$E = E' + \frac{1}{2} \hbar^2 (w_{11} k_1^2 + 2w_{12} k_1 k_2 + w_{22} k_2^2). \quad (9.3.8)$$

By defining

$$\xi(z) = \zeta(z) \exp[-iz(w_{13} k_1 + w_{23} k_2) / w_{33}], \quad (9.3.9)$$

and substituting it into Eq. (9.3.7), the equation is further reduced to

$$\frac{d^2 \zeta_i}{dz^2} + \frac{2m_z}{\hbar^2} [E''_i + e\phi(z)] \zeta_i(z) = 0. \quad (9.3.10)$$

Here we have denoted $m_z = w_{33}^{-1}$. As we can see, m_z describes the motion perpendicular to the surface. Subscript i has been introduced to label the solutions and E''_i satisfies

$$E_i(k_1, k_2) = E''_i + \frac{1}{2} \hbar^2 \left[(w_{11} - \frac{w_{13}^2}{w_{33}}) k_1^2 + 2(w_{12} - \frac{w_{13} w_{23}}{w_{33}}) k_1 k_2 + (w_{22} - \frac{w_{23}^2}{w_{33}}) k_2^2 \right]. \quad (9.3.11)$$

For each i , Eq. (9.3.11) represents a parabolic sub-band with a minimum energy E''_i . From Eq. (9.3.11), for each sub-band we can define the principal effective masses, m_x and m_y . Then Eq. (9.3.11) can be expressed as

$$E_i(k_x, k_y) = E''_i + \frac{1}{2} \hbar^2 \left(\frac{k_x^2}{m_x} + \frac{k_y^2}{m_y} \right). \quad (9.3.12)$$

Obviously, m_x and m_y depend on the orientation of the constant-energy ellipsoids in the bulk with respect to the surface. Fig. 9.15 shows the orientations, degenerate levels (indicated by concentric ellipses) of the constant-energy ellipses and Brillouin zones for the (001), (111) and (110) surfaces of Si. Table 9.2 also lists the effective masses for the three surface orientations of Si.

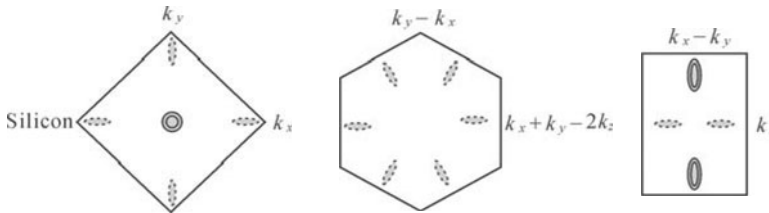


Fig. 9.15 Schematic representation of the constant-energy ellipses and Brillouin zones for the (001), (111) and (110) surfaces of Si. The solid and dash lines indicate sub-bands with higher and lower energies. Taken from Stern and Howard (1967). Copyright (1967) by American Physical Society

Table 9.2 Effective masses for the three surface orientations of Si. The principal effective masses in the ellipsoids are m_x , m_y and m_z . m_x , m_y are principal effective masses of constant-energy ellipse in the surface, and m_z is the effective mass perpendicular to the surface. n_v is the degeneracy of each set of ellipses. After Stern and Howard (1967)

Surface orientation	m_x	m_y	m_z	n_v
{100}	m_l	m_l	m_l	2
	m_t	m_t	m_t	4
{110}	m_t	$(m_t+m_l)/2$	$2m_t m_l / (m_t+m_l)$	4
	m_l	m_l	m_t	2
{111}	m_t	$(m_t+2m_l)/3$	$3m_t m_l / (m_t+2m_l)$	6

With appropriate boundary conditions and approximation, Eq. (9.3.10) can be solved either analytically or numerically. The simplest boundary condition is that ζ_i goes to zero at the interface and in the bulk Si substrate, i.e. $\zeta_i(z=0)=0$ and $\zeta_i(z \rightarrow \infty)=0$.

9.3.1.2 Poisson's Equation

To be consistent, the potential $\phi(z)$ in Eq. (9.3.10) must be determined by Poisson's equation

$$\frac{d^2 \phi}{dz^2} = -\frac{\rho}{\kappa_{sc} \epsilon_0}, \quad (9.3.13)$$

where κ_{sc} is the static dielectric constant of Si substrate and ρ is the charge density, which consists of the contributions from the fixed space charges, the bulk mobile charges (i.e., the bulk extended electron states), the induced charges bounded in the space charge layer (i.e., the 2D electron gas) and the image charges at the semiconductor-insulator interface, i.e.

$$\rho = \rho_{\text{fixed}} + \rho_{\text{mobile}} + \rho_{\text{bound}} + \rho_{\text{image}}. \quad (9.3.14)$$

The corresponding boundary conditions are

$$\begin{aligned} \phi|_{z=0} &= 0, \\ \kappa_{\text{ox}} \left(\frac{d\phi}{dz} \right) \Big|_{z=0^-} &= \kappa_{\text{sc}} \left(\frac{d\phi}{dz} \right) \Big|_{z=0^+}, \end{aligned} \quad (9.3.15)$$

where κ_{ox} and κ_{sc} are the static dielectric constants of the oxide and Si substrate,

respectively.

The induced charge density $\rho_{\text{bound}}(z)$ is the sum of the contributions of all of the sub-bands, i.e.,

$$\rho_{\text{bound}}(z) = -e \sum_i N_i |\zeta_i(z)|^2, \quad (9.3.16)$$

where $\zeta_i(z)$ is normalized electron state solved by Eq. (9.3.10). N_i is the number of electrons of the i th sub-band, the so-called sub-band occupancy. It is computed as

$$N_i = \int_{E_i}^{\infty} f(E) D_i(E) dE, \quad (9.3.17)$$

where

$$f(E) = \frac{1}{1 + \exp\left(\frac{E - \varepsilon_F}{k_B T}\right)} \quad (9.3.18)$$

is the Fermi-Dirac distribution function where k_B , T and ε_F are the Boltzmann constant, temperature and Fermi energy, respectively. $D_i(E)$ is the density of states of the i th sub-band, and it is computed as

$$D_i(E) = \frac{1}{\pi \hbar^2} n_{vi} m_{di} H(E - E''_i), \quad (9.3.19)$$

where n_{vi} is the degeneracy of the i th sub-band whose lowest energy is E''_i , $m_{di} = (m_{li} m_{2i})^{1/2}$ is the so-called density-of-states effective mass, and $H(x) = 1$ if $x > 0$, otherwise $H(x) = 0$. The total number of carriers in all the sub-bands is

$$N_{\text{bound}} = \sum_i N_i. \quad (9.3.20)$$

So far, we have not considered the contributions of the fixed charges ρ_{fixed} , the bulk mobile charges ρ_{mobile} or the image charges ρ_{image} to the potential $\phi(z)$, written as V_{fixed} , V_{mobile} and V_{image} . The fixed charges come from the ionized impurities, whose density must be determined by the positions of the Fermi energy and the impurities level. Nevertheless, for some situations ρ_{fixed} can be computed by simple approximations. For example, in the inversion layer we usually assume that all impurities are ionized in the depletion zone (Stern and Howard, 1967; Stern, 1972). Another example is that for a double-doping accumulation layer at very low temperature (Stern, 1974), the always ionized minority impurities mainly

contribute to the fixed charges. The V_{image} arises from the difference between the dielectric constants of the semiconductor and the insulator (Ando et al., 1982), and can be computed as

$$V(z) = \frac{\kappa_{\text{sc}} - \kappa_{\text{ox}}}{\kappa_{\text{sc}} + \kappa_{\text{ox}}} \frac{e^2}{16\pi\epsilon_0\kappa_{\text{sc}}z}. \quad (9.3.21)$$

The ρ_{mobile} term should be treated carefully as it is quite different in the inversion layer and in the accumulation layer. This is schematically illustrated in Fig. 9.16. For the inversion layer in a p-Si substrate, as the bulk mobile electrons are rare and well separated from the bound states by the depletion layer, ρ_{mobile} and V_{mobile} are negligible. However, for the accumulation layer in an n-Si substrate, the bound and mobile electrons occupy the same region of space. In such a case, only at quite low temperature is the effect of the mobile electrons negligible. Therefore, both bound and extended states should be considered in the accumulation layer.

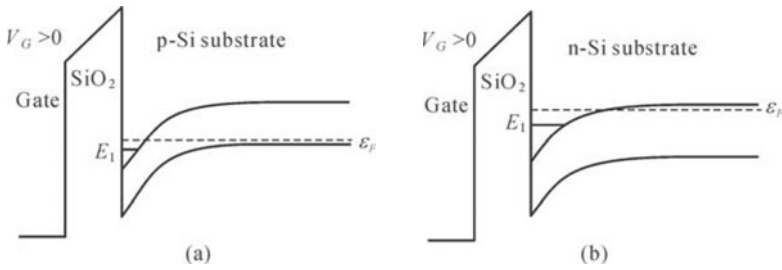


Fig. 9.16 Schematic picture of relative positions of the bound states (denoted by E_1 as the lowest sub-band) and the extended states (denoted as the Fermi energy). (a) Inversion layer in a p-Si substrate, where the Fermi energy near the valence band is above the E_1 sub-band. (b) Accumulation layer in n-Si substrate, where the Fermi energy near the conduction band is above the E_1 sub-band

9.3.1.3 Inversion Layer

Here we consider a double-doping p-Si substrate in the inversion mode. Supposing that all donors and acceptors are ionized in the depletion layer, the density of fixed charges is

$$\rho_{\text{fixed}} = e(N_D - N_A), \quad (9.3.22)$$

and the thickness of the depletion layer is approximately given by

$$d_{\text{depl}} = \left(\frac{2\kappa_{\text{sc}}\epsilon_0\phi_0}{\rho_{\text{fixed}}} \right)^{\frac{1}{2}}, \quad (9.3.23)$$

where ϕ_0 is the absolute value of the band bending at the interface with respect to the bulk. By solving the Poisson equation of ρ_{fixed} , V_{fixed} can be easily calculated as

$$V_{\text{fixed}} = \frac{e\rho_{\text{fixed}}d_{\text{depl}}}{k_{\text{sc}}\epsilon_0} z \left(1 - \frac{z}{2d_{\text{depl}}}\right) = \frac{eN_{\text{fixed}}}{k_{\text{sc}}\epsilon_0} z \left(1 - \frac{z}{2d_{\text{depl}}}\right). \quad (9.3.24)$$

For approximation, ϕ_0 can be replaced by E_g / e , where $E_g \approx (E_c - E_F)_b$ is the band gap of the bulk. Subscript 'b' denotes the bulk. After considering the corrections (Stern, 1972; Ando et al., 1982) from the substrate bias ϕ_{sb} between the source or drain and the substrate, the relative position between the lowest sub-band edge and the Fermi energy at the surface, and the correction of the abrupt termination of the depletion layer, the band bending ϕ_0 can be written in the form

$$\phi_0 = [(E_c - E_F)_b + E_F - k_B T] / e + \phi_{\text{sb}} - eN_{\text{bound}}z_{\text{av}} / \kappa_{\text{sc}}\epsilon_0, \quad (9.3.25)$$

where the last term is the potential drop associated with the inversion layer bounded electrons, with z_{av} as the average penetration of the charges.

The first treatment of the quantum effects in the inversion layer was conducted by Stern and Howard (1967) by solving Schrödinger's equation. In their work, the lowest electric sub-band was concerned. Later, Stern published his self-consistent solution of Schrödinger's and Poisson's equations using the approximation of the triangular potential well for this problem (Stern, 1972). In this case the solution of Schrödinger's equation is the Airy function. When the temperature is low enough and only the lowest sub-band is occupied, a variational method could also be used (Fang and Howard, 1966; Stern, 1972). However, when the electric field at the interface is strong enough, the approximation of the triangular potential well is not appropriate, and the equations should be solved numerically.

9.3.1.4 Accumulation Layer

The self-consistent solution of Eqs. (9.3.10) and (9.3.13) is more complicated in the accumulation layer since the contribution of bulk mobile charges cannot be ignored if the temperature is not too low. Only when the temperature is low enough, are the electrons in the bulk due to the ionization of the donors negligible. For this reason, most early studies dealing with the quantum effect of the accumulation layer had been restricted to very low temperatures (Stern, 1974; Baraff and Appelbaum, 1972; Pals, 1972; Ando, 1976).

A more general and simpler treatment of the accumulation layer was given by Suñé et al. (1991). The basic idea of their approach was to treat the bound states and the extended states on the same footing, i.e. self-consistently solving both

kinds of states in the same Schrödinger's and Poisson's equations. This was done by placing the MOS structure in a three-dimensional quantum box, as shown in Fig. 9.17. One should bear in mind that the $z = 0$ plane has been moved to the bulk substrate.

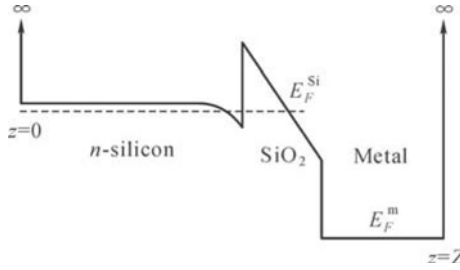


Fig. 9.17 Schematic picture of typical electrostatic potential energy in a MOS structure in accumulation mode. $z = 0$ and $z = Z$ denote the finite boundary of the semiconductor and metal. The potentials there are assumed to be infinite

Although different layers exist in the box, the electron wave function ψ is still assumed to be written as the product of a Bloch function and envelope function as Eq. (9.3.4). Due to the potential discontinuity, the Bloch function has a z dependent factor $\theta(z)$. By assuming that the envelope function $F(x, y, z)$ satisfies the periodic boundary conditions in the x and y directions, it takes the form

$$F(x, y, z) = \frac{1}{\sqrt{XY}} \exp(ik_x x + ik_y y) \xi(z), \quad (9.3.26)$$

where X and Y are the transversal dimensions of the semiconductor. k_x and k_y take values of $\frac{2\pi}{X}r$ and $\frac{2\pi}{Y}l$, respectively, and r and l are quantum numbers.

The three-dimensional Schrödinger equation is reduced to a one-dimensional equation of the envelope function $\xi(z)$ as Eq. (9.3.10), which is rewritten here

$$\frac{d^2 \xi_i}{dz^2} + \frac{2m_z}{\hbar^2} [E''_i + e\phi(z)] \xi_i(z) = 0, \quad (9.3.27)$$

and the total energy of the electron state is

$$E_i = E''_i + \frac{\hbar^2}{2} \left(\frac{1}{m_x} k_x^2 + \frac{1}{m_y} k_y^2 \right). \quad (9.3.28)$$

Eq. (9.3.27) should be solved for the whole MOS structure, with the

changeable effective mass m_z in accordance with the layer where Schrödinger's equation is to be integrated. As for the Si substrate with different orientations, m_z takes values listed in Table 9.2; in the oxide, an isotropic effective mass can be assumed, e.g. $m_z = m_0$; and in the metal, the free electron mass $m_z = m_0$ is considered.

The Poisson equation must be solved in the total quantum box. Eq. (9.3.13) is slightly modified as

$$\frac{d^2\phi}{dz^2} = -\frac{\rho(z)}{\kappa\epsilon_0}, \quad (9.3.29)$$

where κ also takes a value according to the layer under consideration. In the finite box, the energy spectrum of extended states ($E''_i > 0$) is discrete, so that they can be calculated in the same way as bound states ($E''_i < 0$); i.e., we need not treat them separately. Therefore, the charge density $\rho(z)$ including bound states, mobile charges and fixed charges is given by

$$\rho(z) = e(N_d^+ - \sum_i n_{vi} \sum_{E_z} N_i), \quad (9.3.30)$$

where N_d^+ is the density of the fixed ionized charge (in the oxide layer, $N_d^+ = 0$), n_{vi} is the valley degeneracy, and N_i is the carrier concentration of the i th subband, which can be calculated by Eq. (9.3.17).

9.3.2 Transmission Coefficient

9.3.2.1 WKB Approximation

A transmission coefficient, also called tunneling coefficient, transmission probability or tunneling probability, describes either the amplitude or the intensity of a transmitted wave relative to the incident wave. In the case of current tunneling, $T = |J_{\text{tran}}|/|J_{\text{in}}|$ with J_{in} and J_{tran} being the current densities of the incident electronic wave and transmitted electronic wave, respectively.

There are several methods for calculating the transmission coefficient across potential barriers. The WKB (Wentzel-Kramers-Brillouin) approximation (Merzbacher, 1970) is most convenient when the potential varies slowly. Using the traditional WKB approximation, the tunneling coefficient of electrons with energy E entering into the oxide is

$$T_{\text{WKB}}(E) \approx \exp \left[-2 \int_0^{T_t} \kappa(E) dz \right], \quad (9.3.31)$$

where $z = 0$ and $z = T_t$ are the classical turning points and $\kappa(E)$ is the wave vector within the oxide, which can be approximated by

$$\kappa(E) = \frac{2\pi}{h} \sqrt{2m_{\text{ox}}(q\Phi(z) - E)}, \quad (9.3.32)$$

where $\Phi(z)$ is the potential barrier profile, m_{ox} is the effective mass of the electron in the oxide, q is the electronic charge and h is the Planck constant.

However, the WKB approximation fails when the potential profile varies abruptly. It also cannot tackle well the reflection of the wave function from the potential discontinuities. To solve this problem, a modified WKB method was proposed (Register et al., 1999). In this method, the transmission coefficient can be written in the form of

$$T = T_R \cdot T_{\text{WKB}}, \quad (9.3.33)$$

where T_R is the correction factor, taking into account the reflections from all boundaries within the oxide. For a poly-Si gate/oxide/Si structure, it is expressed as

$$T_R = \frac{4v_{\text{Si}}(E)v_{\text{ox}}(\phi_{\text{cat}})}{v_{\text{Si}}^2(E) + v_{\text{ox}}^2(\phi_{\text{cat}})} \times \frac{4v_{\text{Si}}(E + qV_{\text{ox}})v_{\text{ox}}(\phi_{\text{an}})}{v_{\text{Si}}^2(E + qV_{\text{ox}}) + v_{\text{ox}}^2(\phi_{\text{an}})}, \quad (9.3.34)$$

where $v_{\text{Si}}(E)$ and $v_{\text{Si}}(E + qV_{\text{ox}})$ are the group velocities of the electrons entering and leaving the oxide, respectively, with V_{ox} being the oxide potential drop and $v_{\text{ox}}(\phi_{\text{cat}})$ and $v_{\text{ox}}(\phi_{\text{an}})$ the magnitudes of the purely imaginary “group velocity” of electrons at the cathode side and the anode side within the oxide, respectively; E is the energy of the electron in the sub-band, and ϕ_{cat} and ϕ_{an} are the net barrier heights for the electron at the cathode and anode interfaces, respectively.

9.3.2.2 Piecewise-Linear Potential Approximation

Another method for computing the transmission coefficient is to solve Schrödinger's equations for a given potential barrier. For two kinds of barriers we can easily solve the equation, i.e., the rectangular barrier and the triangular barrier. The solutions of Schrödinger's equations for these two kinds of barrier are the plane function and the Airy function, respectively. We can make use of these two known

solutions to solve any potential barrier approximately. The procedure is to first approximate the potential barrier by a piecewise-linear potential or multistep potential, solve Schrödinger's equations at each piecewise interval and connect them at the interface using the continuity conditions.

By assuming x as the direction of tunneling, for the piecewise-linear potential approximation (Lui and Fukuma, 1986), the potential profile $V(x)$ is approximated by the N intervals piecewise linear function as depicted in Fig. 9.18, i.e.

$$\begin{cases} V(x) = V_i - F_i(x - a_i), \\ F_i = -(V_{i+1} - V_i) / (a_{i+1} - a_i), \end{cases} \quad (9.3.35)$$

where $i = 0, 1, 2, \dots, N-1$ for $a_i < x < a_{i+1}$ with V_i and a_i being constants.

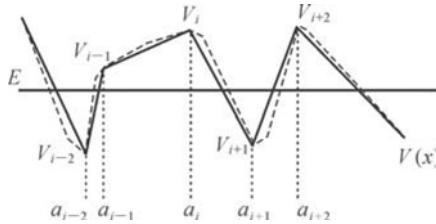


Fig. 9.18 Schematic of piecewise-linear potential approximation. The dashed line is the real potential

By assuming that the potential at the left incident zone and the right out zone is flat, the incident wave and the out wave are plane waves, i.e.

$$u_0(x) = A \exp[ik_L(x - a_0)] + B \exp[-ik_L(x - a_0)], \quad \text{for } x < a_0, \quad (9.3.36a)$$

$$u_{N+1}(x) = H \exp[ik_R(x - a_N)] + G \exp[-ik_R(x - a_N)], \quad \text{for } x > a_N, \quad (9.3.36b)$$

where A, B, H and G are the coefficients to be determined; the wave vectors of the waves depend on the electron energy E and the potential as

$$\begin{cases} k_L = [2m^*(E - V_L) / \hbar^2]^{1/2}, \\ k_R = [2m^*(E - V_R) / \hbar^2]^{1/2}, \end{cases} \quad (9.3.37)$$

where $\hbar$ is the reduced Planck constant and m^* is the electron mass, which is assumed to be uniform through the barrier. The solution of Schrödinger's equation across the interval (a_i, a_{i+1}) ($i = 0, 1, 2, \dots, N-1$) is given as

$$u_i(x) = C_i^{(1)} Ai(z_i) + C_i^{(2)} Bi(z_i), \quad (9.3.38)$$

with

$$\begin{cases} z_i = r_i(c_i - x), \\ r_i = -(2m^* F_i / \hbar^2)^{1/3}, \\ c_i = a_i + (V_i - E) / F_i, \end{cases} \quad (9.3.39)$$

where Ai and Bi are the Airy functions. $C_i^{(1)}$ and $C_i^{(2)}$ are coefficients to be determined. At the interface at a_i , the continuity conditions of the wave function are

$$u_i(a_i) = u_{i+1}(a_i), \quad (9.3.40a)$$

$$\frac{du_i(a_i)}{dx} = \frac{du_{i+1}(a_i)}{dx}. \quad (9.3.40b)$$

Denoting $\mathbf{C}_i = (C_i^{(1)} \ C_i^{(2)})^T$, with Eqs. (9.3.38) and (9.3.40), we can obtain the relationship of coefficients $\mathbf{C}_i$ and $\mathbf{C}_{i+1}$ as follows:

$$\mathbf{M}_i(z_{i,i})\mathbf{C}_i = \mathbf{M}_{i+1}(z_{i+1,i})\mathbf{C}_{i+1}, \quad (9.3.41)$$

where

$$\mathbf{M}_i(z) = \begin{pmatrix} Ai(z) & Bi(z) \\ -r_i Ai'(z) & -r_i Bi'(z) \end{pmatrix}, \quad (9.3.42)$$

$$z_{i,j} = r_i(c_i - a_j),$$

for $i, j = 0, 1, 2, \dots, N-1$. $\mathbf{M}_i(z)$ is called the transfer matrix.

Similarly, using the continuity conditions at the left interface and right interface, we can obtain

$$\mathbf{M}_R(A \ B)^T = \mathbf{M}_0(z_{0,0})\mathbf{C}_0, \quad (9.3.43a)$$

$$\mathbf{M}_N(z_{N,N})\mathbf{C}_N = \mathbf{M}_L(H \ G)^T, \quad (9.3.43b)$$

where $\mathbf{M}_R$ and $\mathbf{M}_L$ are defined as

$$\mathbf{M}_R = \begin{pmatrix} 1 & 1 \\ ik_R & -ik_R \end{pmatrix}, \quad \mathbf{M}_L = \begin{pmatrix} 1 & 1 \\ ik_L & -ik_L \end{pmatrix}. \quad (9.3.44)$$

For N piecewise-linear potential intervals, Eqs. (9.3.41) and (9.3.43) provide us $2(N+1)$ equations containing $2N$ $\mathbf{C}_i$ coefficients and coefficients A, B, H and G . To exactly determine these coefficients, the normalization condition of the wave function is also needed. Making use of these two equations, we can eliminate the C coefficients and are left only with equations for A, B, H and G as follows:

$$\mathbf{M}_R(A \ B)^T = \mathbf{M}\mathbf{M}_L(H \ G)^T, \quad (9.3.45)$$

with

$$\mathbf{M} = \mathbf{M}_1(z_{1,0})\mathbf{M}_1^{-1}(z_{1,1})\mathbf{M}_2(z_{2,1})\cdots\mathbf{M}_i^{-1}(z_{i,j})\mathbf{M}_{i+1}(z_{i+1,j})\cdots\mathbf{M}_N(z_{N,N-1})\mathbf{M}_N^{-1}(z_{N,N}). \quad (9.3.46)$$

If there is no incidental particle coming from the right, i.e., $G = 0$, the transmission coefficient is

$$T(E) = \frac{|H|^2 k_R}{|A|^2 k_L}. \quad (9.3.47)$$

9.3.2.3 Multi-Step Potential Approximation

The calculation of the piecewise-linear potential approximation is quite complicated in some situations, due to the Airy function. Ando and Itoh (1987) suggested a simpler method using multi-step functions to replace the piecewise linear functions. This method is depicted in Fig. 9.19.

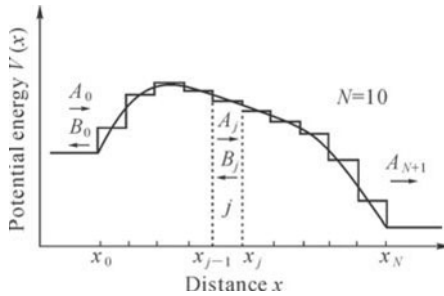


Fig. 9.19 Schematic of multi-step potential approximation. After Ando and Itoh (1987) with permission of J. Appl. Phys.

Considering the x dependent potential $V(x)$, effective electron mass $m^*(x)$ and the dielectric constants $\varepsilon(x)$, they are approximated by multi-step functions as follows:

$$\begin{cases} V(x) = V_j = V[(x_{j-1} + x_j) / 2] \\ m^*(x) = m_j^* = m^*[(x_{j-1} + x_j) / 2] \\ \varepsilon(x) = \varepsilon_j = \varepsilon[(x_{j-1} + x_j) / 2] \end{cases} \quad (9.3.48)$$

for $x_{j-1} < x < x_j$ ($j = 1, 2, \dots, N, N+1$).

As in the interval $x_{j-1} < x < x_j$, the approximated potential is constant, the wave function ψ_j in the j -th region is the plane wave, i.e.

$$\psi_j(x) = A_j \exp(ik_j x) + B_j \exp(-ik_j x), \quad (9.3.49)$$

where

$$k_j = \sqrt{2m_j^*(E - V_j)} / \hbar, \quad (9.3.50)$$

and $\hbar$ is the reduced Planck constant.

Using the continuity conditions of $\psi_j(x)$ and $(1/m_j^*)\left(\frac{d\psi_j}{dx}\right)^{-1}$ at each interface, A_j and B_j can be determined by

$$\begin{pmatrix} A_j \\ B_j \end{pmatrix} = \prod_{l=0}^{j-1} \mathbf{M}_l \begin{pmatrix} A_0 \\ B_0 \end{pmatrix}, \quad (9.3.51)$$

where

$$\mathbf{M}_l = \frac{1}{2} \begin{bmatrix} (1 + S_l) \exp[-i(k_{l+1} - k_l)]x_l & (1 - S_l) \exp[-i(k_{l+1} + k_l)]x_l \\ (1 - S_l) \exp[i(k_{l+1} + k_l)]x_l & (1 + S_l) \exp[i(k_{l+1} - k_l)]x_l \end{bmatrix}, \quad (9.3.52)$$

and S_l is defined as

$$S_l = \frac{m_{l+1}^*}{m_l^*} \frac{k_l}{k_{l+1}}. \quad (9.3.53)$$

Defining matrix $\mathbf{M}$ as the product of $\mathbf{M}_l$

$$\mathbf{M} = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} = \prod_{l=0}^N \mathbf{M}_l, \quad (9.3.54)$$

and taking $A_0 = 1$ and $B_{N+1} = 0$, the transmission coefficient is

$$T(E) = \frac{m_0^*}{m_{N+1}^*} \frac{k_{N+1}}{k_0} |A_{N+1}|^2. \quad (9.3.55)$$

9.3.3 Tunneling Current Components

9.3.3.1 Fowler-Nordheim Tunneling Current

For a MOS structure with thick oxide under high field, as shown in Fig. 9.5b, if there are no space charges in the oxide, electrons are injected from the cathode into the conduction band of the oxide. This kind of tunneling is called FN tunneling. The tunneling barrier is approximated by a triangular type, i.e.

$$\Phi(z) = \Phi_B - E_c z, \quad (9.3.56)$$

where Φ_B and E_c are the barrier height and electric field at the oxide/cathode interface, and z is the tunneling direction.

Substituting Eq. (9.3.56) into Eq. (9.3.32) and combining with Eq. (9.3.31), we can obtain the tunneling coefficient of a given energy E

$$T(E) = \exp\left(-\frac{B}{E_c}\right), \quad (9.3.57)$$

where

$$B = \frac{8\pi\sqrt{2m_{\text{ox}}}\Phi_B^3}{3qh}. \quad (9.3.58)$$

Then the Fowler-Nordheim tunneling current is given as

$$J = AE_c^2 \exp\left(-\frac{B}{E_c}\right). \quad (9.3.59)$$

Specifically, A depends on the effective mass of the electron in the oxide and barrier height at the cathode, viz.

$$A = \frac{q^3(m_0/m_{\text{ox}})}{8\pi h\Phi_B}. \quad (9.3.60)$$

9.3.3.2 Direct Tunneling Current

As shown in Fig. 9.5a, when the oxide is thin enough and the stress is not high to bend the conduction band of the oxide enough, the dominant tunneling current is the DT current. By considering the quantum effect in the Si substrate, for direct tunneling from the inversion layer, the tunneling current comes from the bound

states. In the accumulation case, the contribution from the extended states also needs to be considered. The tunneling current due to the extended states is computed by

$$J_{\text{extended}} = \frac{e}{8\pi^3\hbar} \int_0^\infty dk_z \int_0^\infty dk_l [f(E - E_{\text{fL}}) - f(E - E_{\text{fR}})] T(E) \frac{\partial E}{\partial k_z}, \quad (9.3.61)$$

where $dk_l = dk_x dk_y$, $f(E - E_{\text{fL}})$ and $f(E - E_{\text{fR}})$ are the Fermi distribution functions at the left electrode and the right electrode, and E is the total energy of the extended state.

The tunneling current from the bound states is the contribution of each sub-band in all the valleys, i.e.

$$J_{\text{bound}} = \sum_i n_{vi} \sum_{E''_i} J(E''_i), \quad (9.3.62)$$

where $J(E''_i)$ is the tunneling current from the i th sub-band. For bound states, $J(E''_i)$ depends on the lifetime of the state $\tau_i(E)$ (Rana et al., 1996) as

$$J(E''_i) = \frac{N_i}{\tau_i(E)}. \quad (9.3.63)$$

N_i is the i th sub-band occupancy which can be computed by Eq. (9.3.17).

The life time $\tau_i(E)$ is approximately given by

$$\tau_i(E) = \frac{\int_0^{T_i^i} \sqrt{2m_z / [E''_i - E_c(z)]} dz}{T(E)}, \quad (9.3.64)$$

where T_i^i is the classical turning point of the i th sub-band, and $E_c(z)$ is the bottom of the conduction band of Si substrate; $T(E)$, which can be computed by Eqs. (9.3.33), (9.3.47) or (9.3.55), is the classical transmission probability of electrons with energy E . Eqs. (9.3.61)–(9.3.64) can also be applied to FN tunneling.

9.3.3.3 Trap-Assisted Tunneling

The stress-induced-leakage-current (SILC) is attributed to trap-assisted tunneling. Those traps act as stepping stones, assisting the electron to tunnel through the oxide and in turn generate more traps. Therefore, an appropriate physical model of trap-assisted tunneling is very important for gate oxide reliability. Based on many

SILC experiments, it is now commonly accepted that trap-assisted tunneling should include both elastic tunneling and inelastic tunneling (Takagi et al., 1999). That is to say, electrons could lose part of their energy to relax the traps during the inelastic tunneling, as depicted in Fig. 9.20.

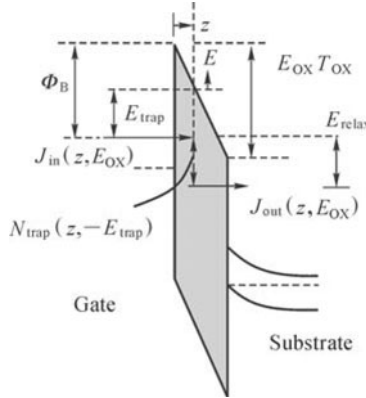


Fig. 9.20 Schematic band diagram of the tunneling model for SILC

Following the work done by Takagi et al. (1999), we consider the trap-assisted tunneling current at position z in the oxide and where the barrier is approximated to be linear. As for inelastic tunneling, traps are assumed to relax during the tunneling. The concentration of traps is $N_{\text{trap}}(z, E)$, with E being the energy level of the traps relative to the bottom of the conduction band. $\rho(z, E)$ and σ are the electron occupancy and capture cross section of the traps located at z with energy level E . The tunnel-in current into these traps, $J_{\text{in}}(z, E_{\text{OX}})$, and the tunnel-out current from the traps, $J_{\text{out}}(z, E_{\text{OX}})$, can be represented by

$$J_{\text{in}}(z, E_{\text{OX}}) = \sigma N_{\text{trap}}(z, z \cdot E_{\text{OX}} - \phi_B) [1 - \rho(z, z \cdot E_{\text{OX}} - \phi_B)] \cdot J(\phi_B, z, E_{\text{OX}}), \quad (9.3.65)$$

$$J_{\text{out}}(z, E_{\text{OX}}) = \sigma N_{\text{trap}}(z, z \cdot E_{\text{OX}} - \phi_B) \rho(z, z \cdot E_{\text{OX}} - \phi_B) \cdot J(\phi_B^*, T_{\text{OX}} - z, E_{\text{OX}}), \quad (9.3.66)$$

where

$$\phi_B^* = \phi_B - z \cdot E_{\text{OX}} + E_{\text{relax}} / q. \quad (9.3.67)$$

Here, $J(\phi, z, E_{\text{OX}})$ is the uniform tunneling current density under the condition that the barrier height between the cathode and SiO_2 is ϕ , the tunneling distance is z and the electric field across the insulator is E_{OX} . T_{OX} is the total thickness of the oxide. The barrier height for the tunneling out electrons ϕ_B^* is higher than that in elastic tunneling by E_{relax} / q .

$J_{\text{in}}(z, E_{\text{OX}})$ should be balanced with $J_{\text{out}}(z, E_{\text{OX}})$ at a steady state. Then we determine $\rho(z, E)$ and can obtain the tunneling current at z as

$$J'(z, E_{\text{OX}}) = J_{\text{in}}(z, E_{\text{OX}}) = J_{\text{out}}(z, E_{\text{OX}}) = \sigma N_{\text{trap}}(z, z \cdot E_{\text{OX}} - \phi_{\text{B}})A(z), \quad (9.3.68)$$

with

$$A(z) = \frac{J(\phi_{\text{B}}^*, T_{\text{OX}} - z, E_{\text{OX}})J(\phi_{\text{B}}, z, E_{\text{OX}})}{J(\phi_{\text{B}}^*, T_{\text{OX}} - z, E_{\text{OX}}) + J(\phi_{\text{B}}, z, E_{\text{OX}})}. \quad (9.3.69)$$

It is easy to see that when z_i satisfies $J(\phi_{\text{B}}', T_{\text{OX}} - z_i, E_{\text{OX}}) = J(\phi_{\text{B}}, z_i, E_{\text{OX}})$, $A(z)$ reaches its maximum. As $A(z)$ is actually a sharp peak function of z , we can make the following approximation:

$$A(z) \approx \frac{t_{\text{eff}} J(\phi_{\text{B}}, z_i, E_{\text{OX}})}{2} \delta(z - z_i), \quad (9.3.70)$$

where t_{eff} is called the effective tunneling thickness and $\delta(z)$ is the delta function.

As a consequence, the total tunneling current through SiO_2 can be described by

$$J_{\text{tot}}(E_{\text{OX}}) = \int_0^{T_{\text{OX}}} J'(z, E_{\text{OX}}) dz = \frac{t_{\text{eff}} \sigma N_{\text{trap}}(z_i, z_i \cdot E_{\text{OX}} - \phi_{\text{B}})}{2} J(\phi_{\text{B}}, z_i, E_{\text{OX}}). \quad (9.3.71)$$

It is easy to see that Eq. (9.3.71) can also be used to compute elastic trap-assisted tunneling by setting $E_{\text{relax}} = 0$ in Eq. (9.3.67).

9.3.4 Microscopic Investigation of Defects from First-Principles Calculation

As we have shown, defects play important roles in oxide degradation. They can trap charges and act as stepping stones to assist tunneling, thus leading to the increase in leakage current. Despite such importance, people seldom observed the microstructures of these defects directly in experiments. Usually we detect interface states to infer the existence of the defects. As the great advance in computer science, the first-principles calculation has become an efficient method for us to investigate the microscopic origin of these defects. By computing and comparing the energy levels of defects as well as investigating their influence on band structure and conductance etc., we can infer which defect is most responsible for oxide degradation. We have illustrated some concepts of the energetics of defects in Section 9.2. Here we would like to use a first-principles calculation

example to show how defects affect the band structure and conductance of the oxide.

Oxygen vacancies have been identified as an important factor in the formation of electron traps in much experimental research (Imai et al., 1988; Agnello et al., 2002; Stathis, 2002). Recent electron energy loss spectrometry analysis (Li et al., 2008) showed that the neutral oxygen deficiency within a breakdown path can be as high as 50%, and the vacancies spread out radially from the center of the percolation path. Therefore, it is significant to investigate oxygen vacancies. Fig. 9.21 shows our model of the Al–SiO₂–Al system, which can be divided into the left electrode, the scattering region and the right electrode.

The effect of the oxygen deficiency is simulated in Fig. 9.21b by removing two oxygen atoms in the bulk network of SiO₂. Combining the nonequilibrium Green's function technique with the self-consistent pseudopotential method in the framework of DFT and LDA approximation, the density of state (DOS) of the perfect and the oxygen deficient silica, i.e., SiO_x ($x < 2$), can be computed. The results are shown in Fig. 9.21c and Fig. 9.21d. There are additional defect states above the valence band maximum for the SiO_{1.95}, and the conduction band minimum (CBM) is lowered by 0.6 eV. With higher V_O concentration, the atomic structure is severely distorted in SiO_{1.66}, resulting in more defect states in the band gap and further lowering the CBM.

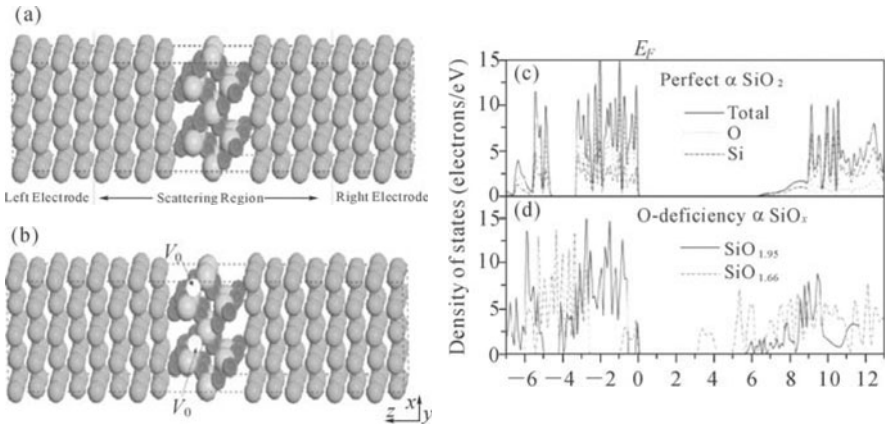


Fig. 9.21 Schematic diagram of Al–SiO₂–Al system (a) with and (b) without oxygen vacancies. (c) Calculated results of the density of state (DOS) of the perfect and (d) the oxygen deficient silica. The Fermi level is set to zero. Reprinted with permission from Luo et al. (2009). Copyright (2009) by the American Institute of Physics

The leakage current in a nanoscale device is explored with the Landauer–Büttiker formula:

$$J = \frac{2e}{h} \int T(E) [n_F(E - \mu_R) - n_F(E - \mu_L)] dE, \quad (9.3.72)$$

where μ_R and μ_L represent the chemical potential of the left and right electrode, respectively, n_F is the Fermi distribution function and $T(E)$ is the transmission probability of electrons passing through the barrier. For zero temperature calculation, the above formula can be written as

$$J = \frac{2e}{h} \int_{\mu_L}^{\mu_R} T(E) dE. \quad (9.3.73)$$

In Fig. 9.22a, we show the transmission coefficient at the zero bias for the perfect and V_O containing silica junctions. It is clear to see that the transmission probability of electrons is strongly enhanced by the introduction of the V_O defect structure. The transmission peaks in the oxygen deficient junction are correlated with the active electrical defect states induced by the vacancies. The current as a function of the applied voltage bias for both the perfect and oxygen deficient systems is shown in Fig. 9.22b. The leakage currents show a steeper increase for the oxygen deficient system than for the perfect one as the applied voltage increases.

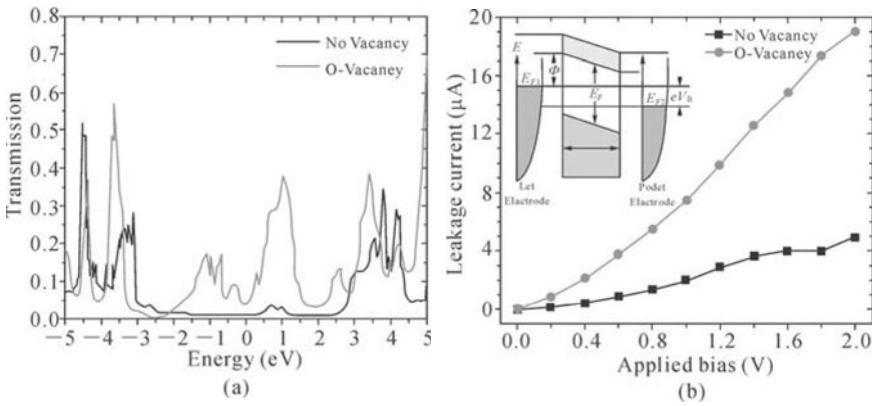


Fig. 9.22 Calculated (a) transmission coefficient at the zero bias and (b) current as a function of the applied voltage bias for both the perfect and oxygen deficient systems. Reprinted with permission from Luo et al. (2009). Copyright (2009) by the American Institute of Physics

9.3.5 Manipulating Tunneling by Applied Strains

Given the importance of the enormous strain field in the dielectric breakdown process (Jeffery et al., 1998; Pey et al., 2002; Choi et al., 2009), it is interesting to look at the strain effects on the electronic states. Fig. 9.23 shows the DOS information obtained from simulation with 2% tensile strain along the (001) plane of SiO_2 . Similar to the effect of oxygen deficiency, the CBM is lowered by about

0.8 eV as compared to the unstrained structure. The DFT simulations also show that the tensile stress in SiO_2 gives a relatively high conduction band DOS, which indicates large modification of interfacial carrier mobility at the conduction band edges. It is foreseeable from this result that applied strain can effectively affect the tunneling property of SiO_2 .

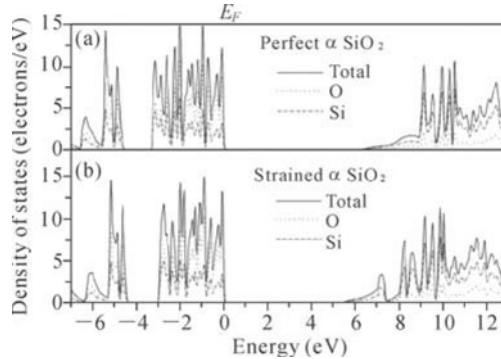


Fig. 9.23 First-principles result of the dependence of the total DOS on the electron energy for (a) the unstrained and (b) tensile strained SiO_2 . The Fermi level is set to zero. Reprinted with permission from Luo et al. (2009). Copyright (2009) by the American Institute of Physics

Another example of control of tunneling by applied strain is some kind of indirect effect. It is well known that mechanical strain largely affects the magnitude of polarization, critical thickness and order of transition of ferroelectric materials. A ferroelectric tunnel junction (FTJ) exploits the physics of spontaneous polarization dependent tunneling in ferroelectric barriers. Experimental and theoretical work indicated that the tunneling electro-resistance (TER) across the FE barriers can vary significantly in asymmetric FTJ when the polarization is reversed (Zhuravlev et al., 2005; Zheng and Woo, 2009). It inspires us that we can use applied strain to manipulate the polarization state and then control the tunneling of FTJ.

Combining the nonequilibrium Green functions with density-functional theory, here we investigate the effect of applied strain on the electron transport of a ferroelectric tunnel junction (FTJ) with PbTiO_3 thin film sandwiched between Pt electrodes (Luo et al., 2010). It was found that the four unit cells unstrained PbTiO_3 tunneling barrier tended to be in a paraelectric (PE) state and switched to a ferroelectric (FE) state with increasing compressive strain u_m to -1.12% . The calculated transmission spectra for the highly strained ($u_m = -2.57\%$) FE state junction and unstrained PE state junction are given in Fig. 9.24a. The transmission spectra differed from each other, suggesting the changes in electron states in the two different strained junctions.

The overall transmission of the unstrained PE junction is higher than that of the strained FE junction, especially for the energy window just above the Fermi level. That means the transmission peaks in the unstrained PE state junction will contribute to its conductance at small voltage bias. According to quantum

mechanics, the conductance of the above systems can be expressed by the sum of zero voltage transmission probability at the Fermi energy E_F . Within strain engineering, a significant TER ratio of 28 can be obtained through the manipulation of the PE/FE transition. To quantify the sensitivity of the TER with the applied mechanical forces, we have defined the giant piezoelectric resistance (GPR) effect, i.e.

$$\text{GPR} = \frac{R_{\text{strained}} - R_{\text{unstrained}}}{\min(R_{\text{strained}}, R_{\text{unstrained}})} \quad (9.3.74)$$

The tunneling current is calculated with the Landauer–Büttiker formula (9.3.72). As shown in Fig. 9.24b, the dependence of tunneling current and GPR on the applied voltage is calculated at a strain of -2.57% . In contrast to the low resistance in the unstrained PE junction, the current of the strained FE junction changes significantly, with an increase in resistance by about two orders of magnitude. The corresponding GPR ratio can reach 9000%, which is an important result that has exciting technological implications.

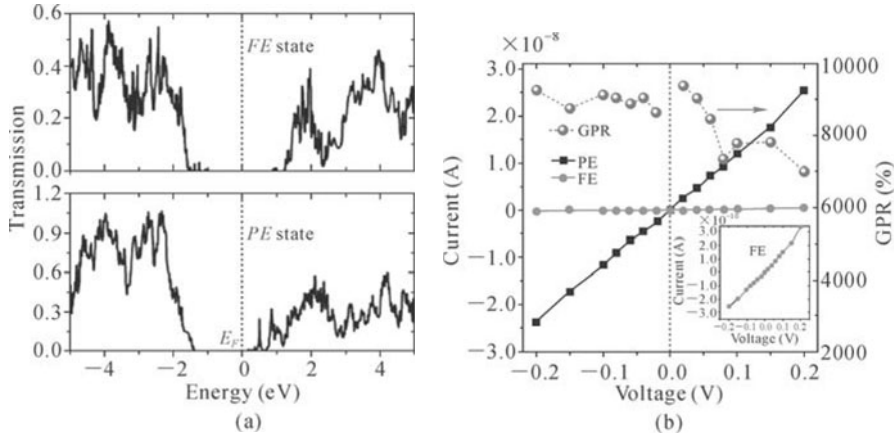


Fig. 9.24 (a) Dependence of the transmission spectra on the energy for junctions with strain induced FE state barrier (the upper one) and unstrained PE state barrier (the lower one). The Fermi level is set to zero. (b) Tunneling current as a function of the applied bias for the PE and FE junctions and their resistance ratios are depicted by the GPR. Reprinted with permission from Luo et al. (2010). Copyright (2010) by the American Institute of Physics

9.4 Degradation Models in Gate Oxide Films

Quite a number of dielectric breakdown models tackling degradation in gate oxide have been proposed since the 1960s and early 1970s when the MOS structure became widely used in semiconductors. The early models of oxide breakdown

were based on impact ionization. The basic idea of these models is that impact ionization generates holes and leads to positive charges in the bulk oxide. The enhanced electric field near the cathode due to those charges causes an increase in electron injection and in turn triggers more impact ionization. Positive charges are generated by the imbalance in holes generation and drifting out of the positive charges (O'Dwyer, 1969; Klein, 1972) or of holes generation and recombination (DiStefano and Shatzkes, 1974; 1975; 1976; Shatzkes and Av-Ron, 1976; Klein and Solomon, 1976), which were called the ID model and IR model, respectively. Later on, models of oxide breakdown which were attributed to electron trapping were proposed (Harari, 1978; Meyer and Crook, 1983).

Around the mid 1980s, three models became prevalent. They were the thermochemical model (*E* model), anode hole injection model (AHI model) and the anode hydrogen release model (AHR model). The thermochemical model was proposed by McPherson and Baglee (1985a; 1985b) based on the idea that defect generation was the reaction of bond breakage. Almost at the same time, Chen and his co-workers proposed the so-called AHI model, emphasizing the role of holes in those early impact ionization models but without specifying the origin of hole impact ionization (Chen et al., 1985; 1986). The idea of the AHR model is that hydrogen species can be released by hot electrons at the interface of the anode and travel back to the cathode to generate defects. Both experimental work (see, e.g., Gale et al., 1983; Feigl et al., 1984; Buchanan, 1993; Salvatore et al., 2005) and first-principles calculations (see, e.g., Blöchl and Stathis, 1999a; 1999b; Blöchl, 2000) provide evidence of hydrogen participating defect generation. However, it was not until 2001, when Suñé and Wu (2001) attempted to quantitatively model the process of hydrogen participating defect generation. Up until now these degradation models still exhibit robust vitality in specific regimes. Actually, there was a trend to incorporate them into a unified model (Okada and Yoneda, 1999; Hu and Lu, 1999; McPherson et al., 2000; Cheung, 2001). Such a unified model can resolve the contradiction between the *E* model and $1/E$ model. In the remaining part of this section, the three mentioned degradation models and those tackling the thermal and mechanical effects on oxide degradation are introduced.

9.4.1 Anode Hole Injection Model

This was first proposed by Chen et al. (1985; 1986) and modified by Lee et al. (1988). They attributed hole trapping as the cause of thin oxide breakdown. In the original version of this model, the holes were believed to be generated by impact ionization in the oxide. However, subsequent research did not specify that the holes came from impact ionization. Nevertheless, to illustrate the main idea of this model, in the following we still consider that the holes are from impact ionization. As the holes get trapped near the cathode, the local field is enhanced as depicted in Fig. 9.25. This enhanced local field increases the injection current of electrons from the cathode and in turn more holes are generated and get trapped near the

cathode until breakdown occurs. Initially, the potential barrier in the oxide under stress is approximated as a triangular one described by Eq. (9.3.56) and the electron injection is FN tunneling described by Eq. (9.3.59). The existence of positive charges by trapped holes will change the shape of the barrier into a non-triangular one. By assuming that an effective electric field E_{eff} can be used to take account of the barrier change, the dependence of tunneling current on the electric field is still described by Eq. (9.3.59), i.e.

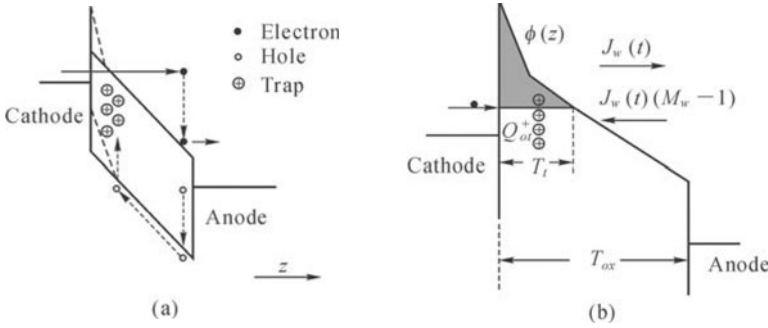


Fig. 9.25 Schematic illustration of hole induced breakdown model

$$J_w = AE_{\text{eff}}^2 \exp\left(-\frac{B}{E_{\text{eff}}}\right), \quad (9.4.1)$$

where the effective electric field is

$$E_{\text{eff}} = \frac{\Phi_B^{3/2}}{\frac{3}{2} \int_0^{T_i} \sqrt{\Phi'(z)} dz}. \quad (9.4.2)$$

For a given density of positive charges (trapped holes) Q_{ot}^+ , the non-triangular $\Phi'(z)$ can be calculated by using the Gauss law (Chen et al., 1985). The density of positive charges is given by

$$Q_{\text{ot}}^+ = \eta \int_0^{T_i} J_w(t)(M_w - 1) dt, \quad (9.4.3)$$

where η is the trapping efficiency and M_w is the multiplication factor related to the impact ionization coefficient $\alpha(E_{\text{eff}})$ as

$$M_w = \frac{1}{1 - \int_{T_i}^{T_{\text{ox}}} \alpha(E_{\text{eff}}) dz} \approx 1 + \int_{T_i}^{T_{\text{ox}}} \alpha(E_{\text{eff}}) dz. \quad (9.4.4)$$

The impact ionization coefficient $\alpha(E_{\text{eff}})$ depends on the effective field as follows:

$$\alpha(E_{\text{eff}}) = \alpha_0 \exp(-H / E_{\text{eff}}), \quad (9.4.5)$$

with H and α_0 being constants.

By assuming that the area of the wearing out region of the oxide is A_w and that of the residual region is A_r , the total current through the oxide can be calculated as

$$J_{\text{total}} = \frac{1}{A_w + A_r} (A_w J_w + A_r J_r), \quad (9.4.6)$$

where J_r is computed by using Eq. (9.3.59).

A critical density of trapped holes is assumed to be that needed to trigger the breakdown. Then the time to the breakdown t_{BD} is inversely proportional to J_w according to Eq. (9.4.3), i.e.

$$t_{\text{BD}} \approx \frac{c}{J_w (M_w - 1)} \approx \tau_0 \exp\left(\frac{B + H}{E_{\text{eff}}}\right) = \tau_0 \exp\left(\frac{G}{E_{\text{eff}}}\right), \quad (9.4.7)$$

where G , c and τ_0 are constants. From Eq. (9.4.7) we can see t_{BD} is exponentially dependent on $1/E_{\text{eff}}$. For simplicity, we may approximate E_{eff} as the average electric field across the oxide. That is why this model is also called the $1/E$ model.

To include the effect of extrinsic factors on oxide breakdown, such as particulate contamination, crystalline defects in the substrate, localized thin regions or surface roughness, Lee et al. (1988) proposed using the “effective oxide thinning” factor, ΔT_{ox} . Then the intrinsic breakdown model is modified to account for the statistical variation of the breakdown.

$$t_{\text{BD}} = \tau_0 \exp\left[\frac{G(T_{\text{ox}} - \Delta T_{\text{ox}})}{V_{\text{ox}}}\right] = \tau_0 \exp\left[\frac{GT_{\text{eff}}}{V_{\text{ox}}}\right]. \quad (9.4.8)$$

In subsequent works, the origins of holes have been extended to include hole injection from the anode to explain the experimental results of low field breakdown. The anode hole injection theory was developed by Fischetti (1985). These injected holes are considered to be responsible for the oxide breakdown at a low electric field where impact ionization in the bulk oxide is unlikely to occur. That is why the model is called the anode hole injection model.

However, we must bear in mind that even after taking into consideration all these origins of holes, the AHI model explains experimental data much better at a high electric field than at a low field regime. Several modified hole induced

breakdown models were proposed to extend the model to a low field regime. Readers are referred to Alam et al. (1999; 2000; 2002).

9.4.2 Thermochemical Model

The thermochemical model is also called the “E model” as it predicts an exponential dependence of t_{BD} on the electric field. The basic idea of the E model is that the degradation of dielectrics is a field-enhanced-reaction process of bond breakage. The reaction is driven by a free energy difference between the initial bond state and the broken bond state. It was first proposed by McPherson and Baglee (1985) as a phenomenological model without considering the microscopic aspects of bond breakage. Ten years later, McPherson and his co-workers gave a description of the molecular origin and physics of bond breakage in SiO_2 (McPherson et al., 1997; McPherson and Mogul, 1997; 1998).

The structure of SiO_2 is depicted in Fig. 9.26a. The SiO_4 tetrahedrons are linked with each other by the oxygen atoms to form the network of SiO_2 . While the bond angle of O-Si-O is fixed at 109° , the angle of Si-O-Si can change between 120° – 180° . Under local stress, the bond strength Si-O-Si is severely weakened as the bond angle deviates from 150° , and oxygen vacancies, denoted as the $\text{O}_3\text{Si-Si=O}_3$ molecule, tend to form as depicted in Fig. 9.26b.

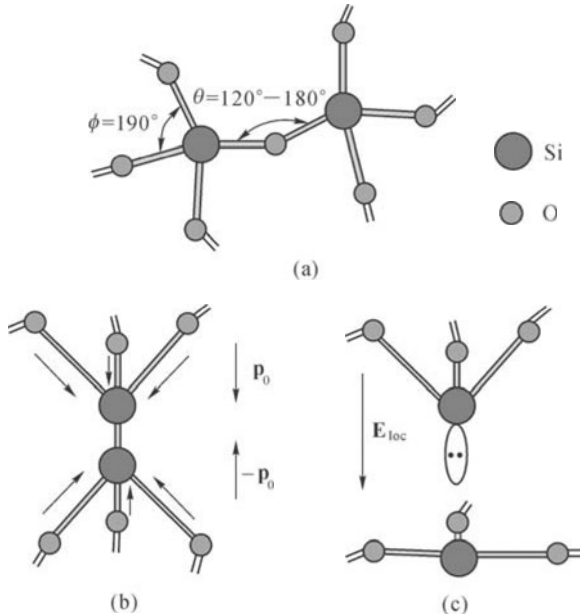


Fig. 9.26 Schematic picture of (a) the structure of SiO_2 , (b) oxygen vacancy in SiO_2 and an $\text{O}_3\text{Si-Si=O}_3$ molecule is formed. (c) Under the local electric field, the Si-Si bond breaks and the lower part of $\text{O}_3\text{Si-Si=O}_3$ molecule collapses to planar sp^2 configuration

Due to the large difference in electronegativity of the Si atom and O atom, the Si-O bond is polar, as depicted by the arrows in Fig. 9.26b. Since the Si-O bonds of the $\text{O}_3\equiv\text{Si}$ unit are not in the same plane after the $\text{O}_3\equiv\text{Si-Si}\equiv\text{O}_3$ molecule is formed, the $\text{O}_3\equiv\text{Si}$ unit has a permanent polarization and can be treated as a dipole. When an external electric field E_{ox} is applied to the oxide, the $\text{O}_3\equiv\text{Si-Si}\equiv\text{O}_3$ molecule experiences a local electric field E_{loc} as

$$\mathbf{E}_{\text{loc}} = \mathbf{E}_{\text{ox}} + L(\mathbf{P} / \epsilon_0) = (1 + L\chi)\mathbf{E}_{\text{ox}}, \quad (9.4.9)$$

where L is the Lorentz factor and $L = 1/3$ is assumed for amorphous SiO_2 . In deriving Eq. (9.4.9) the following relation has been used:

$$\mathbf{P} = \chi\epsilon_0\mathbf{E}_{\text{ox}} \quad (9.4.10)$$

where χ is the dielectric susceptibility and ϵ_0 is the permittivity of the vacuum.

Here we assume that the dipoles are either parallel or antiparallel to the applied electric field as shown in Fig. 9.26c, and that the magnitude of polarization of the dipoles is p_0 in the absence of an electric field. When the dipoles are under a local field E_{loc} , the energy of those parallel ones is given by

$$u_{\text{p}} = u_0 - p_0 E_{\text{loc}} - \frac{1}{2} \alpha E_{\text{loc}}^2, \quad (9.4.11)$$

and the energy of those antiparallel ones is

$$u_{\text{ap}} = u_0 + p_0 E_{\text{loc}} - \frac{1}{2} \alpha E_{\text{loc}}^2, \quad (9.4.12)$$

where u_0 is the reference energy of the dipoles without the local electric field, and α is the polarizability. Due to the different orientations, it is easy to see that the dipole of the upper $\text{O}_3\equiv\text{Si}$ is energetically favorable, as compared to that of the lower $\text{O}_3\equiv\text{Si}$.

If the Si-Si bond was to break, and the lower $\text{O}_3\equiv\text{Si}$ of the $\text{O}_3\equiv\text{Si-Si}\equiv\text{O}_3$ molecule was to collapse to a near planar sp^2 hybridization, as depicted in Fig. 9.27c, the dipolar energy contribution $p_0 \cdot E_{\text{loc}}$ is available to help lower the activation energy required for bond breakage, i.e.

$$(\Delta H) = (\Delta H)_0 - p_0 \cdot E_{\text{loc}}, \quad (9.4.13)$$

where $(\Delta H)_0$ is the activation energy in the absence of a local electric field. Then the reaction rate of bond breakage is given by

$$k = \nu_0 \exp \left[-\frac{(\Delta H)_0 - p \cdot E_{\text{loc}}}{k_B T} \right]. \quad (9.4.14)$$

Assuming that the time to breakdown of the dielectric is inversely proportional to the reaction rate, we can obtain t_{BD} as

$$t_{\text{BD}} \propto \frac{1}{k} \propto \exp \left[\frac{(\Delta H)_0 - p_0 \cdot E_{\text{loc}}}{k_B T} \right] = \exp \left[\frac{(\Delta H)_0 - p_0 \cdot (1 + L\chi)E_{\text{ox}}}{k_B T} \right]. \quad (9.4.15)$$

In deriving the last equation, Eq. (9.4.9) has been used.

From Eq. (9.4.15) we can see that t_{BD} has an exponential dependence on the electric field. This is why it is called the E model. McPherson and Mogul (1997) extended the above model to consider more than one molecule defect state with different bonding energy and activation energy of bond breakage. Some long term TDDDB experiments showed that the thermochemical model was more suitable to describe the low field regime of oxide breakdown compared to the AHI model (Boyko and Gerlach, 1989; Suehle et al., 1994; Chaparala et al., 1994; Kimura, 1997). Therefore it predicts a more accurate t_{BD} of the device when it is extrapolated to the operating conditions.

9.4.3 Anode Hydrogen Release Model

The anode hydrogen release model is similar to the AHI model except that it considers hydrogen species instead of holes as the cause of silicon oxide breakdown. Due to the complexity of hydrogen desorption, diffusion and defect generation, the AHR model is in fact a quasi-quantitative model. Nevertheless, there are a lot of experimental results supporting this model.

Electron devices are usually annealed in forming gas containing hydrogen above 400 °C to passivate the active defects at the interface. Therefore, hydrogen exists at the SiO₂/Si interface of most MOS devices. Experiments (Fig. 9.27a) showed that these hydrogen species could be released by hot electrons, and may play an important role in oxide breakdown (Gale et al., 1983; Feigl et al., 1984; Buchanan et al., 1993). There are two pieces of evidence for these experiments that support hydrogen participation in oxide degradation. Firstly, devices tested in a humid or hydrogen atmosphere show that the failure rate increases with increasing humidity or hydrogen pressure. Secondly, degradation of the oxide is observed at zones near both electrodes (DiMaria and Stasiak, 1989; Buchanan and DiMaria, 1990; DiMaria and Stathis, 1991). Since electrons gain energy in the oxide and become hot near the anode, it is natural to expect that degradation is concentrated near the anode. Therefore, the second piece of evidence indicates that some species probably travel back to the cathode and cause degradation there.

Hydrogen is a promising candidate.

Hydrogen release can generate defects with fast and slow interface states by three steps. (1) Hydrogen species are released by hot electrons at the SiO_2/Si interface. (2) Under the electric field these species diffuse across the oxide. (3) They react with the hydrogenated Si dangling bonds at the cathode as depicted in Fig. 9.27b. Such hydrogenated Si dangling bonds are usually called P_b centers. The reaction at the cathode can be expressed as follows:



Experiments showed the P_b center only contributed to a part of the interface states (Cartier and Stathis, 1995), indicating that hydrogen might also form defects in the bulk oxide. Furthermore, by using first-principles calculation, the charge-state levels of oxygen vacancies, hydrogen, complex and other promising defects in the oxide can be determined. There are two requirements for a defect to be a key candidate for causing SILC, as depicted in Fig. 9.28, i.e. (1) a charge-state level near the Fermi level of a cathode, and (2) a small relaxation in the charge-state level when the defect traps electrons. Only if its charge-state level shifts upon relaxation by less than the applied voltage, can the defect assist tunneling.

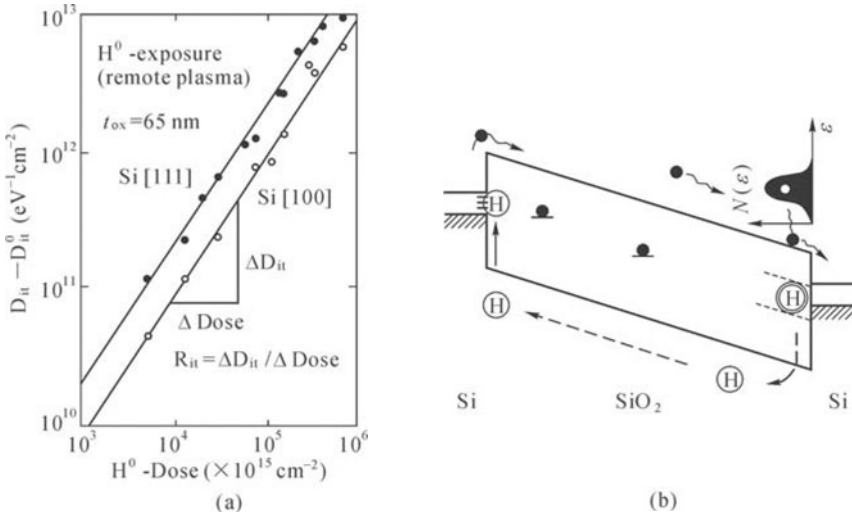


Fig. 9.27 (a) Interface-state buildup at mid-gap as a function of atomic hydrogen, H^0 , dose measured in a remote hydrogen plasma at room temperature. Reprinted with permission from Cartier et al. (1995). Copyright (1995) by Elsevier. (b) Scheme of hydrogen release process, where the right upper chart is the energy spectrum of electron at the anode. Reprinted with permission from DiMaria et al. (1992). Copyright (1992) by the American Institute of Physics

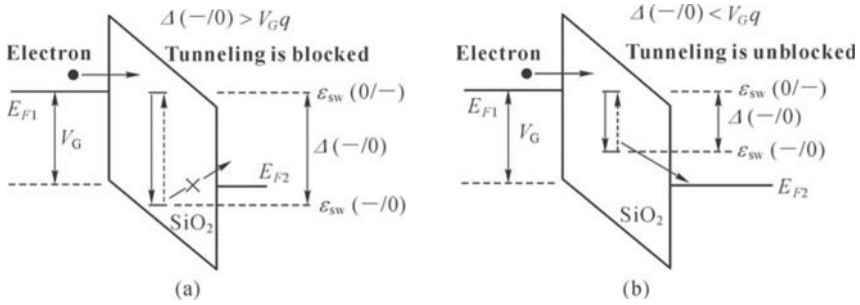


Fig. 9.28 Scheme of the requirements for trap-assisted tunneling. (a) Tunneling is blocked if relaxation of the charge-state level is too large. (b) Tunneling is unblocked as the relaxation charge-state level is appropriate

Comparison of the charge-state levels of various defects showed that the neutral hydrogen bridge, also called the E center in quartz, was probably responsible for stress-induced leakage current (Blöchl and Stathis, 1999). This provides further evidence of hydrogen participation in oxide degradation. Table 9.3 shows the computational results of charge-state levels of different defects which might be responsible for oxide degradation. Most defects in silica exhibit large charge-state level shift. In Table 9.3, ϵ_{th} is the thermodynamic charge-state level and ϵ_{av} is the mean value. Δ is the difference in the corresponding switch charge-state levels, e.g. $\Delta(0/-) = \epsilon_{sw}(0/-) - \epsilon_{sw}(-/0)$. $\epsilon_{av}(0/-)$ is defined as $[\epsilon_{sw}(0/-) + \epsilon_{sw}(-/0)]/2$.

Table 9.3 First-principles calculation results of various charge-state levels of several candidate defects that may be responsible for dielectric breakdown in SiO₂. Reproduced with the permission of Blöchl and Stathis (1999)

Defect	ϵ_{th}	$\epsilon_{av}(+/0)$	$\Delta(+/0)$	$\epsilon_{av}(0/-)$	$\Delta(0/-)$
(SiH) ₂	-2.74(+ / 0)	-2.37	2.60	—	—
O vacancy	-3.03(+ / 0)	-2.54	2.32	—	—
Si(3)+O(3)	-0.06(+ / 0)	-0.17	1.70	—	—
H	0.20(+ / -)	> -0.21	> 4.10	-0.58	4.23
SiH+Si(3)	0.30(+ / -)	> 0.25	> 2.92	-0.29	3.11
H bridge	0.33(+ / 0)	0.26	2.22	—	—
H bridge	0.74(0 / -)	—	—	0.72	1.71

Suñé and Wu (2001) proposed a quantitative hydrogen-based model, which assumed that those defects were generated by a two-step mechanism. Firstly, hot electrons release charged hydrogen atoms or protons at the interface of the anode, i.e. H^+ . Secondly, such protons react with oxygen vacancies in the oxide. The reaction rates of these two steps are denoted as ξ_1 and ξ_2 , respectively. They are expressed by

$$\xi_1(V_G, T) = \frac{K_1}{1 + \exp[(E_{\text{THI}} + q\phi_B - qV_G)/E_0]}, \quad (9.4.17a)$$

$$\xi_2(V_G, T) = K_2 V_G^2 \exp(-\frac{B}{V_G}), \quad (9.4.17b)$$

where K_1 and K_2 are constants related to the concentration of H^+ near the anode and to the bulk concentration of vacancies, respectively. E_{THI} is the height of the barrier for proton release. ϕ_B is the difference in electron affinity at the Si/SiO₂ interface. E_0 is the height of the saddle-point contact potential barrier (Suñé and Miranda, 2000). V_G is the gate voltage and B is related to the barrier height of the reaction of H^+ with the oxygen vacancy. Then the total defect generation rate is $\xi = \xi_1 \xi_2$. The kinetic equations of concentration of protons N_{H^+} and the concentration of defects are given by

$$\frac{dN_{\text{H}^+}}{dt} = \xi_1 \left\{ \frac{J(V)}{qt_{\text{ox}}} - \frac{N_{\text{H}^+}}{\tau_0} \right\} - \xi_2 \frac{N_{\text{H}^+}}{\tau_0} - \xi_3 \frac{N_{\text{H}^+}}{\tau_0}, \quad (9.4.18a)$$

$$\frac{dN}{dt} = \xi_2 \frac{N_{\text{H}^+}}{\tau_0}. \quad (9.4.18b)$$

9.4.4 Thermal Breakdown Model

Although dielectric breakdown is well known as a process with mechanical stress, the electric field and thermal field intercoupling with each other, relatively few papers have paid attention to the effect of the mechanical stress field and thermal field. Usually, the impact of the thermal field is incorporated into breakdown models by introducing a temperature accelerating factor obeying the Arrhenius law as the E model, i.e. $k \propto \exp(-E_a/k_B T)$. The activation energy E_a has to be determined empirically.

To reflect the effects of local heating, heat dissipation and temperature rise on dielectric breakdown, a more delicate treatment is needed. The coupling relation between the thermal field and electric field is usually described by the heat conduction equation (Wagner, 1922; O'Dwyer, 1973) as follows:

$$C_v \frac{dT}{dt} - \nabla \cdot K \nabla T = \sigma E^2, \quad (9.4.19)$$

where C_v is the specific heat capacity of the dielectric, K is the heat conductivity, and σ is the electric conductivity. The current density satisfies the continuity condition, i.e.

$$\nabla \cdot J = 0. \quad (9.4.20)$$

Eqs. (9.4.19) and (9.4.20) are usually solved by numerical methods. Both Whitehead (1953) and O'Dwyer (1973) gave impartial reviews of the thermal breakdown model.

Considering a device under long-term low stress, the temperature within the dielectric varies slowly. Then Eq. (9.4.19) reduces to

$$-\nabla \cdot K \nabla T = \sigma E^2. \quad (9.4.21)$$

Wintle (1981) solved Eq. (9.4.21) with different $K(T)$ and $\sigma(T)$. He pointed out that such an equation could not totally explain dielectric breakdown in polymers. Huang et al. (2004) used a similar time-independent heat equation to model the vacuum breakdown of a carbon nanotube.

When the stress is of the pulse type, the heat is not able to be dissipated immediately. Then we can neglect the heat conduction term in Eq. (9.4.19), i.e.

$$C_v \frac{dT}{dt} = \sigma E^2. \quad (9.4.22)$$

Vermeer (1956) used Eq. (9.4.22) to explain the breakdown of glass and Hikita (1985) also used it to explain his experimental results.

9.4.5 Mechanical-Stress-Induced Breakdown Model

Mechanical stress can be generated near the interface due to the volume expansion from Si to SiO₂ during the thermal oxidation. Such interfacial stress strongly depends on growth conditions, such as temperature, pressure and growth rate. It can largely affect the local bond state by changing the angle, length and force constant of the bond.

In the simplest case, we consider that mechanical stress mainly affects the bond angle (Yang and Saraswat, 2000). Similar to the thermochemical model, we consider bond breakage of Si-O-Si as the cause for dielectric breakdown and that the time to breakdown is related to the field-modified entropy change ΔH^* as

$$t_{BD} \propto \exp\left(\frac{\Delta H^*}{kT}\right). \quad (9.4.23)$$

The bonding energy of Si-O-Si is a function of the angle θ of Si-O-Si. Assuming that it reaches its minimum at θ_0 , we can express it by a Taylor expansion, i.e.

$$E(\theta) = E(\theta_0) + \frac{K_\theta}{2}(\theta - \theta_0)^2 + \frac{K'_\theta}{3!}(\theta - \theta_0)^3 + \frac{K''_\theta}{4!}(\theta - \theta_0)^4 \dots, \quad (9.4.24)$$

where K_θ , K'_θ , K''_θ ... are expansion coefficients and K_θ is also called the force constant. Neglecting the higher order terms, the bond energy is reduced to a harmonic oscillator

$$E(\theta) = E(\theta_0) + \frac{K_\theta}{2}(\theta - \theta_0)^2. \quad (9.4.25)$$

For two oxides under the same electric field but different mechanical stresses, the entropy change is different for the mechanical contribution; the time to breakdown of the oxides follows

$$\frac{t_{BD2}}{t_{BD1}} = \exp\left(\frac{\Delta H_2^* - \Delta H_1^*}{k_B T}\right) = \exp\left[\frac{\frac{K_\theta}{2}(\theta_1 - \theta_0)^2 - \frac{K_\theta}{2}(\theta_2 - \theta_0)^2}{k_B T}\right], \quad (9.4.26)$$

where θ_1 and θ_2 are the corresponding bond angles under different local mechanical stresses. Such angles can be detected by Fourier transform infrared spectroscopy (FTIR) experiments or determined by first-principles calculation.

9.4.6 Remarks

In summary, various degradation models of gate oxide have been presented in this section. As a key process in oxide breakdown, defect generation is still an intensively debated issue. Three popular models have been proposed to explain defect generation, i.e., the AHI model, the AHR model and the thermochemical model. Both the AHI model and AHR model can be related to the phenomenological $1/E$ model and the thermochemical model is related to the phenomenological E model. They can explain the specific range of experimental data. The $1/E$ model shows better agreement with experimental data in a high field while the E model is more accurate in a relatively low field.

We cannot use either model to predict gate oxide breakdown well over a wide stress range. Various mechanisms may coexist within a breakdown event. People now tend to establish a unified model which can resolve the controversy of the E

model and $1/E$ model (Hu and Lu, 1999; McPherson et al., 2000; Cheung, 2001). A unified model trying to resolve the contradiction of various possible breakdown mechanisms can be simply constructed by assuming that $1/T_{BD} = 1/T_{BD1} + 1/T_{BD2} + \dots + 1/T_{BDN}$, where T_{BDk} , $k=1, 2, \dots, N$, are the time to breakdown of different breakdown mechanisms. Such a unified model could explain a larger range of experimental results. Nevertheless, the coupling effects of electric, mechanical and thermal fields should be further well incorporated into the model. There is still a long way to go before settling this issue.

9.5 Statistical Models of Dielectric Breakdown

So far we have illustrated the two important processes of dielectric breakdown, i.e. tunneling phenomenon and defect generation. These processes are described by deterministic equations. Actually, those equations describe the evolution of average quantities, such as the tunneling current and the density of defects. They cannot account for the statistical characteristics of dielectric breakdown. Dielectric breakdown is likely to occur at the weakest spot and has the property of extreme value statistical distribution. In this section, statistical and percolation models that tackle the statistics of breakdown are presented.

9.5.1 A Basic Statistical Model

We shall first introduce the statistical picture of dielectric breakdown given by Suñé et al. (1990) as depicted in Fig. 9.29. Despite lacking predictive ability in relating the critical density of defects to oxide thickness or composition, such a model provides us with a physical interpretation of the breakdown statistics in terms of the critical density of defects. It also tells us the dependence of the distribution function on the oxide area.

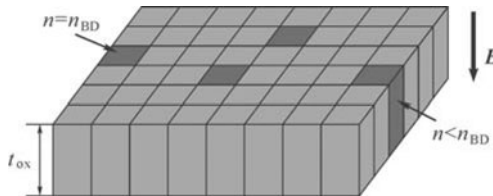


Fig. 9.29 Schematic of a device divided into smaller columnar cells. Cells of dark color are those that are defective

In this simple model, a device with area A is divided into N smaller columnar cells with area S_0 . Under electric stress, defects are assumed to be generated

randomly on the entire device at a constant rate. The cells are also considered independent of each other. Breakdown occurs when the defect density of one cell reaches a critical value N_{BD} .

We assume that the average density of defects is ρ , and that the probability to find a given number of defects n in a cell is described by the Poisson distribution if N is not too small, i.e.

$$P(n, \rho) = \frac{(\rho S_0 t_{ox})^n \exp(-\rho S_0 t_{ox})}{n!}. \quad (9.5.1)$$

When n reaches N_{BD} , the cell is considered to break down. Hence the broken probability of a cell is

$$P_{BD}(\rho) = 1 - \sum_{n=0}^{N_{BD}-1} \frac{(\rho S_0 t_{ox})^n \exp(-\rho S_0 t_{ox})}{n!}. \quad (9.5.2)$$

Then the probability of survival of a device of area $A = NS_0$ is

$$V(\rho) = [1 - P_{BD}(\rho)]^N. \quad (9.5.3)$$

Hence the cumulative failure rate of breakdown of a device with area $A = NS_0$ is

$$F(\rho) = 1 - [1 - P_{BD}(\rho)]^N = 1 - \left[\sum_{n=0}^{N_{BD}-1} \frac{(\rho S_0 t_{ox})^n \exp(-\rho S_0 t_{ox})}{n!} \right]^N. \quad (9.5.4)$$

Therefore, we derive the Weibull number as follows:

$$W(\rho) = \ln\{-\ln[1 - F(\rho)]\} = \ln A + \ln \left\{ \rho t_{ox} - \frac{1}{S_0} \ln \left[\sum_{n=0}^{N_{BD}-1} \frac{(S_0 t_{ox} \rho)^n}{n!} \right] \right\}. \quad (9.5.5)$$

From Eq. (9.5.5) we can see that the dependence of the Weibull number on the area is quite similar to the extreme value distribution.

Eq. (9.5.4) provides us with the relationship of the accumulated failed percentage of a given average defect density ρ , device area A and critical defect density N_{BD} . In this relationship, ρ is taken as the random variable. In order to apply this relationship to explain experimental results where the number of injection charges or stress time are usually used as the variables, we need to relate ρ to these variables. This is the issue of degradation models in the last section and will be discussed in the following.

9.5.2 A Three-Dimensional Statistical Model

As depicted in Fig. 9.30, we still consider a device with oxide thickness t_{ox} and area A . There are still N cells in a plane, i.e. $N = A/S_0$, but now each columnar cell is divided into m cubes with length a_0 . In this model, the oxide breaks down when all the cells in one column fail. The probability that a cell becomes broken is

$$P_{\text{cell}}(\rho) = 1 - \sum_{n=0}^{N_{\text{BD}}-1} \frac{(\rho a_0^3)^n \exp(-\rho a_0^3)}{n!}. \quad (9.5.6)$$

The probability of failure for one column of cells, the survival probability, the cumulative distribution of failure of the whole device and the Weibull number are given by

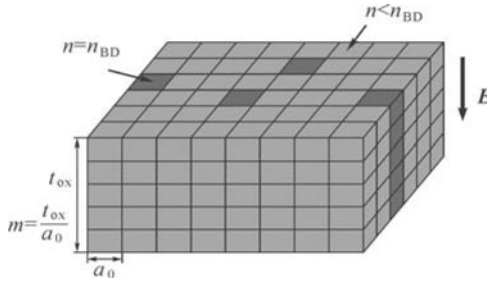


Fig. 9.30 Schematic of a device divided into smaller cube cells. Cells of dark color are those that are defective

$$P_{\text{BD}}(\rho) = [P_{\text{cell}}(\rho)]^m, \quad (9.5.7)$$

$$V(\rho) = [1 - P_{\text{BD}}(\rho)]^N = \left\{1 - [P_{\text{cell}}(\rho)]^m\right\}^N, \quad (9.5.8)$$

$$F(\rho) = 1 - \left\{1 - [P_{\text{cell}}(\rho)]^m\right\}^N, \quad (9.5.9)$$

$$W(\rho) = \ln \left\{ -\ln [1 - F(\rho)] \right\} = \ln \left\{ -N \ln [1 - P_{\text{cell}}^m(\rho)] \right\}. \quad (9.5.10)$$

If the probability that a cell becomes broken is quite small, i.e. $P_{\text{cell}}(\rho) \ll 1$, we can make the approximation

$$\ln [1 - P_{\text{cell}}^m(\rho)] \approx -P_{\text{cell}}^m(\rho). \quad (9.5.11)$$

Then the Weibull number is reduced to

$$W(\rho) = \ln \left(\frac{A}{a_0^2} \right) + m \ln [P_{\text{cell}}(\rho)]. \quad (9.5.12)$$

It is complicated to use Eq. (9.5.6). Often we simply assume that the probability of a cell being broken is proportional to the average density of a generated defect, i.e.

$$P_{\text{cell}}(\rho) = a_0^3 \rho. \quad (9.5.13)$$

Then Eq. (9.5.12) becomes

$$W(\rho) = \ln\left(\frac{A}{a_0^2}\right) + \frac{t_{\text{ox}}}{a_0} \ln(a_0^3 \rho). \quad (9.5.14)$$

By letting $W(\rho) = 0$, i.e. $F(\rho) = 0.632$, which is commonly used as the mean breakdown probability, it is easy to see that the mean critical density of a defect can be expressed as follows (Suñé and Wu, 2001):

$$\langle N_{\text{BD}} \rangle = \frac{1}{a_0^3} \exp\left[-\frac{a_0}{t_{\text{ox}}} \ln\left(\frac{A}{a_0^2}\right)\right]. \quad (9.5.15)$$

If we assume a power law relationship for the number of the injected charge and the density of the generated defect, i.e.

$$n = n_0 Q_{\text{inj}}^\alpha, \quad (9.5.16)$$

we can also use Eq. (9.5.14) or Eq. (9.5.15) to compute the mean charge to breakdown as

$$\langle Q_{\text{BD}} \rangle = \frac{1}{(a_0^3 n_0)^\alpha} \exp\left[-\frac{a_0}{\alpha t_{\text{ox}}} \ln\left(\frac{A}{a_0^2}\right)\right]. \quad (9.5.17)$$

Once we have known the relationship between defect generation and the stress time, we can similarly compute the mean time to breakdown.

9.5.3 Sphere and Cube Based Percolation Models

It is commonly accepted that percolation models are quite appropriate to describe the statistical properties of dielectric breakdown. Degraeve et al. (1995) first established such a kind of model to reproduce the distribution of breakdown. As is schematically shown in Fig. 9.31, a simulated sample with fixed dimensions is defined and inside this volume traps are randomly generated at random positions. These traps are described as spheres with fixed radius r .

If the spheres of two neighboring traps overlap, we say that conduction between these traps becomes possible. The algorithm continues generating traps until a conducting path is created from one interface to the other. This defines the

breakdown condition, and at this point we can define the critical density of traps N_{BD} and the charge to breakdown Q_{BD} . There are only two parameters in this model, i.e. the radius of traps r and the oxide thickness t_{ox} . More exactly, the results only depend on the radius thickness ratio r/t_{ox} .

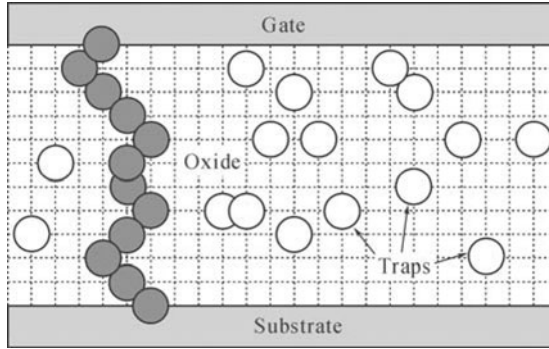


Fig. 9.31 Schematic illustration of sphere based percolation model

For different oxide thickness t_{ox} , the distributions of the critical density of traps can be directly computed using a Monte-Carlo type of simulation. Fig. 9.32a shows the simulated distributions of the critical density of traps N_{BD} . The Weibull characteristic is well represented. Often we measure the distributions of Q_{BD} rather than N_{BD} in experiments. Such distributions can be obtained from those of N_{BD} by assuming the power law relationship of the number of injected charges and the density of charges, as in Eq. (9.5.16). Fig. 9.32b shows the simulated distribution of Q_{BD} .

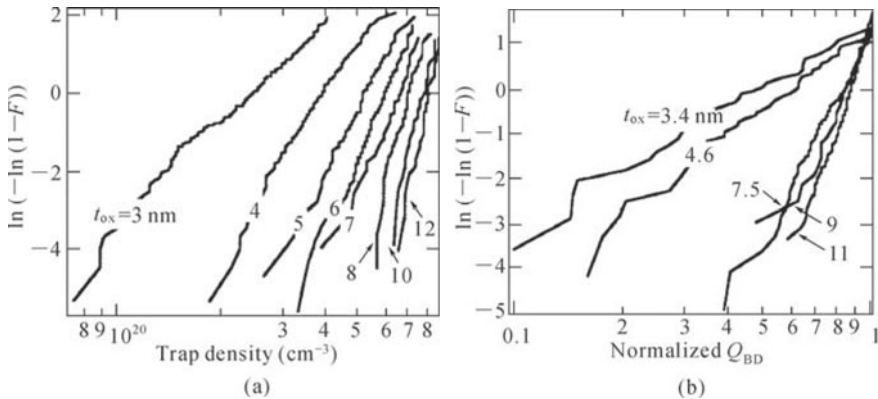


Fig. 9.32 Simulation results from a sphere based percolation model. (a) Distributions of critical density of traps and (b) distributions of charge to breakdown of various oxide thicknesses. Reprinted with permission from Degraeve et al. (1995). Copyright (1995) by IEEE

Stathis (1999) proposed a cube-based percolation model which was quite similar to the above sphere-based model except that it used cubes as the basic computational units instead of spheres. In this model, the size of defect could span more than one cube. To build a defect cell with radius r at r_0 , the neighboring cells within r are all marked broken. In the limiting case, the sphere percolation model can also be obtained. Compared with the above sphere percolation models, this model is simpler and thus can significantly reduce the computational time. Fig. 9.33 shows the result of simulated dependence of the critical density of defects at 63% accumulative failure rate on oxide thickness and the comparison with the experimental results.

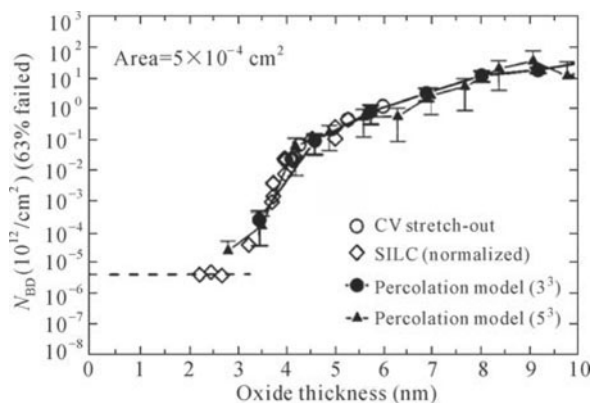


Fig. 9.33 Result of simulated dependence of critical density of defects at 63% accumulative failure rate on oxide thickness and comparison with the experimental results using a cube-based percolation model. Reprinted with permission from Stathis (1999). Copyright (1999) by the American Institute of Physics

9.5.4 Combination of Percolation Model and Degradation Model

In traditional percolation models, the probability of the cells being broken is often assumed to be uniform, i.e. with a constant probability P_0 . Alternatively, a non-uniform case (Stathis, 1999) can be considered by a given position dependent probability such as $P(x) = P_0 e^{-x/\alpha}$ which denotes an exponential dependence of broken probability on the position. Such a method cannot precisely describe defect generation and does not take account of surface roughness the local effects of the defective zone on the evolution. In order to accurately simulate the statistical characteristics of dielectric breakdown, it is necessary to combine tunneling, defect generation and percolation together.

The method is schematically illustrated in Fig. 9.34. At first, the simulated oxide is divided into columnar super cells. The thickness of each super cell is different, to simulate the roughness of the surface, which can be experimentally

obtained from a scanning tunneling microscope (STM) or atomic force microscope (AFM). Secondly, the potential profiles and distributions of carriers of biased super cells are calculated by solving Schrödinger's and Poisson's equations. Then the tunneling current including different components is computed using the method described in Section 9.3. After that, the number of injected holes (released hydrogen) generated by electrons is computed by the AHI (AHR) model. Holes (hydrogens) then travel back to the oxide. The defects are created according to the local holes or hydrogen flux. Using a percolation model to calculate the thickness dependent critical density of defect N_{BD} , the condition of breakdown can be determined. Alam et al. (1999) established such a comprehensive model based on an anode hole injection model.

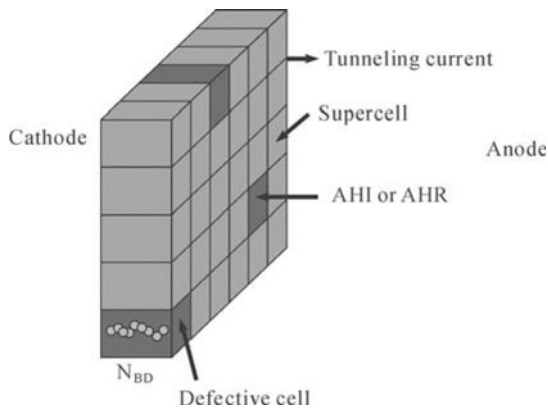


Fig. 9.34 Schematic illustration of a comprehensive percolation model

9.6 Damage of Dielectric Breakdown in MOSFET

A dielectric breakdown will induce serious damage in electronic devices, such as a MOSFET. The degree of damage is largely affected by the circuit environment of the device. When breakdown occurs in a MOSFET, various damage such as breakdown spots, dielectric-breakdown-induced silicon epitaxy (DBIE), dielectric-breakdown-induced metal migration (DBIM), severe substrate damage and the total burnout of the MOS structure could occur. Some damage, such as substrate damage and total burnout of the device, appears only when the stress is high enough or the compliant current is large enough, while some damage, such as DBIE, occur even when the stress is quite low. These damage processes are probably driven by the coupling of thermal heat, the electric field and mechanical stress field.

9.6.1 Lateral Propagation of Breakdown Spot

Experiments have shown that breakdown often starts at one spot, and then propagates laterally at the MOS interfaces as depicted in Fig. 9.35 (Lombardo et al., 2005). Such a process was found to be highly dependent on the oxide thickness. For a 10 nm-oxide layer stressed by constant voltage ~ 10 V at room temperature, the length and width of the propagating breakdown-spots line were about 200 μm and 300 nm, respectively. For an oxide layer of about 5 nm under the same conditions, the length and width of the spot line were lower by about 40 μm and 50 nm, respectively. As the oxide thickness was further decreased, no propagation of breakdown spots was found.

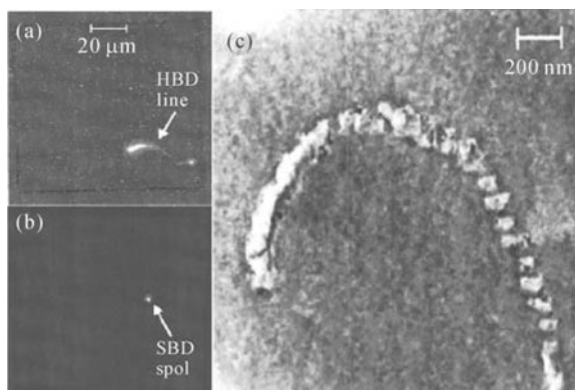


Fig. 9.35 Photon-emission micrographs of the lateral propagation of breakdown spots recorded after the BD in (a) 5.6 nm and (b) 3 nm oxides. (c) TEM micrograph of the region of breakdown spot line of the same sample of (a). Reprinted with permission from Lombardo et al. (2005). Copyright (2005) by the American Institute of Physics

The breakdown-spots line is believed to be caused by the subsequent breakdown events, although at first only one or a few percolative passages are formed and trigger breakdown. Meanwhile there exist a lot of other almost-percolative passages at that moment. As the first breakdown occurs, the breakdown spot acts as a strong sink of electrons and affects the local fields significantly. Such a large enhancement of fields may be responsible for triggering secondary breakdown events in the nearby almost-percolative passages.

9.6.2 Dielectric Breakdown-Induced Epitaxy

The silicon epitaxy hillock at the breakdown spot is often observed as dielectric breakdown occurs in a MOS structure as shown in Fig. 9.36. It is called dielectric-breakdown-induced epitaxy (DBIE). For a ploy-Si gate MOSFET, DBIE was

found to be polarity dependent; i.e., it always occurred at the cathode side. DBIE appears even at the earliest stage of breakdown and continues to grow as the process goes on. It can grow rapidly under high stress. DBIE is so easy to observe that people often use it to locate the breakdown zone.

It is believed that DBIE may also be responsible for boosting breakdown. It appears even at an early stage of breakdown and its continual growth during stressing can induce a high local mechanical field. Such a mechanical field would push aside the oxide, cause mechanical deformation in the oxide, weaken the dielectric property there and finally lead to breakdown.

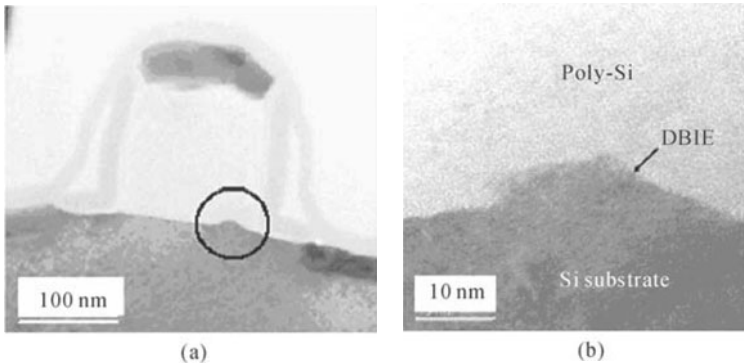


Fig. 9.36 TEM micrograph of DBIE in a failed transistor with HBD. (a) A DBIE, indicated by the circle, is seen near one of the poly-Si edges. (b) High-resolution TEM micrograph of the circled zone of (a) showing Si epitaxy formed on the Si substrate. Reprinted with permission from Tang et al. (2004). Copyright (2004) by IEEE

9.6.3 Dielectric Breakdown-Induced Migration

As the current compliance at breakdown is set large, one of the most severe types of damage in a MOSFET is the migration of silicide. This phenomenon is called dielectric-breakdown-induced migration (DBIM). Fig. 9.37 depicts the experimental observations of DBIM in a Ti-silicided MOSFET after breakdown. It can be seen that the migration of Ti species occurs not only from the silicided source and drain but also from the silicided gate. DBIM is considered to occur more easily in regions where a large thermal gradient exists, such as the zone near the DBIE.

9.6.4 Meltdown and Regrowth of Silicided Poly-Si Gate

When breakdown occurs, a large heat flux is generated in the MOS device. There is poorer thermal conduction in the surroundings of the poly-Si gate than in the

region of the substrate, including the drain and source, as this region is usually in contact with metal plugs. Due to asymmetry in heat conduction, severe meltdown in the poly-Si always takes place ahead of a similar meltdown in the substrate, even though the melting points of poly-Si and Si(100) substrate are similar. In most cases, this destructive process is severe enough to cause entire poly-Si structure melt, intermixed with the silicide, as shown in Fig. 9.38.

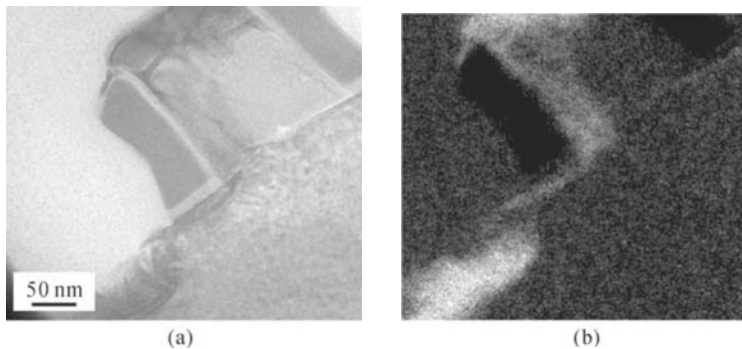


Fig. 9.37 (a) TEM micrograph of HBD damage in a 3.3 nm gate oxide of a Ti-silicided n-MOSFET. (b) Elemental map of Ti of the region shown in (a). Reprinted with permission from Tang et al. (2004). Copyright (2004) by IEEE

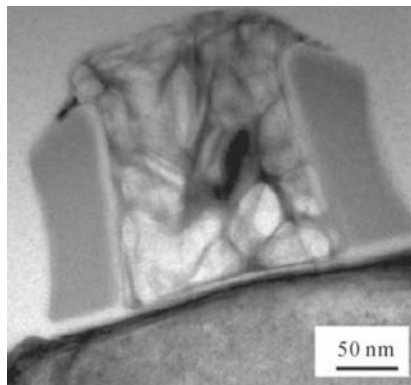


Fig. 9.38 Poly-Si gate meltdown in a 3.3 nm gate-oxide n-MOSFET stressed in inversion mode with a current compliance of 1 mA. In this case, the original TiSi on top of the poly-Si is totally absent. Reprinted with permission from Tang et al. (2004). Copyright (2004) by IEEE

9.6.5 *Damage in Substrate*

Much thermal energy generated at the breakdown spot can be readily conducted away via the high thermal conductive Si substrate and the W-plugs of the source/drain contacts. However, if the heat flux is quite large, microstructural

damage can possibly occur in the Si substrate subsequent to the meltdown and regrowth of silicide and poly-Si mixture in the gate electrode structures. Fig. 9.39 shows such damage in MOSFETs with ultra thin oxides.

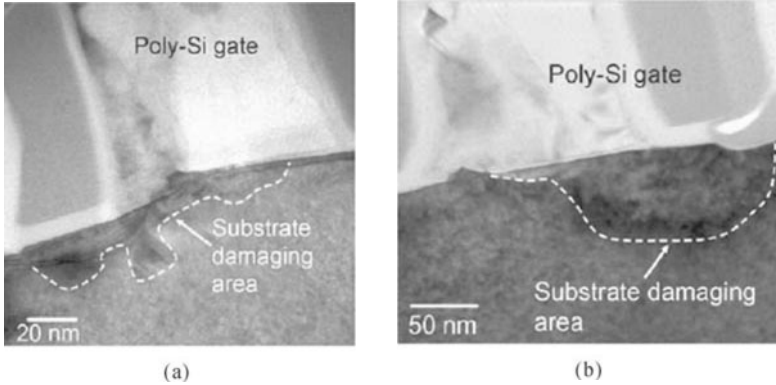


Fig. 9.39 TEM micrograph of (a) an n-MOSFET with 3.3 nm and (b) 2.5 nm gate oxide. Reprinted with permission from Tang et al. (2004). Copyright (2004) by IEEE

9.7 Remarks

Dielectric breakdown is a crucial issue in microelectronic and nanoelectronic devices as it largely affects the life span and operating range of the devices. Despite intensive investigations, dielectric breakdown is still a difficult subject. This is mainly due to the complexity of electric, thermal and mechanical stress fields intercoupling with each other, and because different mechanisms probably coexist within a breakdown event. Another reason comes from the difficulty in obtaining direct investigation of the microscopic process of dielectric breakdown using experimental techniques.

Generally speaking, there are two most imperative problems in this field. One is the microscopic mechanism of defect-participating processes, including the issue of how defects are generated, what structures the defects adopt and under what conditions the state of the defects are energetically favorable. So far, we are still not clear which defect is most responsible for dielectric breakdown, and this is one of the most important reasons why there are so many oxide breakdown models. The second imperative problem is to discover how various fields intercouple with each other. Nonlinear effects should be very important. A unified model that considers these intercoupling fields has not yet been established.

Nevertheless, there are two sides to everything. The sensitivity of dielectric breakdown in applied fields provides us with promising opportunities to control it. A lot of work could be done with respect to this issue. For example, we have shown in Section 9.3 that we can use substrate strain to modulate the band structures or the polarization states of the dielectrics, thus changing their electronic transport

properties. Furthermore, we can also control the damage caused by dielectric breakdown and make use of it. The resistance ratio before and after breakdown is so large that we can use this property to store logical states. Recently, the concept of one-time programmable non-volatile memory (OTP NVM) which makes use of hard dielectric breakdown has been proposed (Peng and Fong, 2004). Meanwhile, the recoverability of a soft breakdown has been found to have promising applications for non-volatile programmable memory. Actually, this is a very hot issue in a new extending research field, i.e., so-called resistive switching memories (Schindler et al., 2008; Waser et al., 2009; Jeong et al., 2009).

References

- [1] Agnello, S., Boscaino, R., Buscarino, G., Cannas, M., Gelardi, F.M.: Structural relaxation of E'_{γ} centers in amorphous silica. *Phys. Rev. B* **66**, 113201 (2002)
- [2] Alam, M., Weir, B., Bude, J., Silverman, P., Ghetti, A.: A computational model for oxide breakdown: theory and experiments. *Microelectron. Eng.* **59**, 137-147 (2001)
- [3] Alam, M.A., Bude, J., Ghetti, A.: Field acceleration for oxide breakdown—can an accurate anode hole injection model resolve the E vs. $1/E$ controversy? Reliability Physics Symposium, 2000, Proceedings 38th Annual 2000 IEEE International, pp. 21-26
- [4] Alam, M.A., Bude, J., Weir, B., Silverman, P., Ghetti, A., Monroe, D., Cheung, K.P., Moccio, S.: An anode hole injection percolation model for oxide breakdown—the “Doomsday” scenario revisited. *Electron Devices Meeting, 1999. IEDM Technical Digest, International*, pp. 715-718
- [5] Alam, M.A., Weir, B.E. Silverman, P.J.: A study of soft and hard breakdown-Part I: Analysis of statistical percolation conductance. *IEEE Trans. Electron Dev.* **49**, 232-238 (2002)
- [6] Ando, T.: Density-functional calculation of sub-band structure in accumulation and inversion layers. *Phys. Rev. B* **13**, 3468-3477 (1976)
- [7] Ando, T., Fowler, A.B., Stern, F.: Electronic properties of two-dimensional systems. *Rev. Mod. Phys.* **54**, 437-672 (1982)
- [8] Ando, Y., Itoh, T.: Calculation of transmission tunneling current across arbitrary potential barriers. *J. Appl. Phys.* **61**, 1497-1502 (1987)
- [9] Azizi, N., Yiannacouras, P.: Gate Oxide Breakdown. *Lecture Notes, Reliability of Integrated Circuits* (2003)
- [10] Baldereschi, A., Baroni, S., Resta, R.: Band offsets in lattice-matched heterojunctions: A model and first-principles calculations for GaAs/AlAs. *Phys. Rev. Lett.* **61**, 734-737 (1988)
- [11] Baraff, G.A., Appelbaum, J.A.: Effect of electric and magnetic fields on the self-consistent potential at the surface of a degenerate semiconductor. *Phys. Rev. B* **5**, 475-497 (1972)

- [12] Bardeen, J.: Tunnelling from a many-particle point of view. *Phys. Rev. Lett.* **6**, 57-59 (1961)
- [13] Blchl, P.E.: First-principles calculations of defects in oxygen-deficient silica exposed to hydrogen. *Phys. Rev. B* **62**, 6158-6179 (2000)
- [14] Blchl, P.E., Stathis, J.H.: Aspects of defects in silica related to dielectric breakdown of gate oxides in MOSFETs. *Phys. B Condens. Matter.* **273**, 1022-1026 (1999)
- [15] Blchl, P.E., Stathis, J.H.: Hydrogen electrochemistry and stress-induced leakage current in silica. *Phys. Rev. Lett.* **83**, 372-375 (1999)
- [16] Boyko, K.C., Gerlach, D.L.: Time Dependent Dielectric Breakdown at 210 Å oxides. *Proceedings of the IEEE Reliability Physics Symposium*, 1989, pp. 1-8
- [17] Buchanan, D.A., DiMaria, D.J.: Interface and bulk trap generation in metal-oxide-semiconductor capacitors. *J. Appl. Phys.* **67**, 7439-7452 (1990)
- [18] Buchanan, D.A., Marwick, A.D., DiMaria, D.J., Dori, L.: *The Physics and Chemistry of SiO₂ and the Si-SiO₂ Interface*. Plenum, New York (1993)
- [19] Buchanan, D.A., Stathis, J.H., Cartier, E., DiMaria, D.J.: On the relationship between stress induced leakage currents and catastrophic breakdown in ultra-thin SiO₂ based dielectrics. *Microelectron. Eng.* **36**, 329-332 (1997)
- [20] Cartier, E., Buchanan, D.A., Stathis, J.H., Maria, D.J.: Atomic hydrogen-induced degradation of thin SiO₂ gate oxides. *J. Non-cryst. Solids* **187**, 244-247 (1995)
- [21] Chang, C., Hu, C., Brodersen, R.W.: Quantum yield of electron impact ionization in silicon. *J. Appl. Phys.* **57**, 302-309 (1985)
- [22] Chaparala, P., Suehle, J.S., Messick, C., Roush, M.: Electric field dependent dielectric breakdown of intrinsic SiO₂ films under dynamic stress. *Reliability Physics Symposium*, 1996. 34th Annual Proceedings, IEEE International.
- [23] Chen, I.C., Holland, S., Young, K.K., Chang, C., Hu, C.: Substrate hole current and oxide breakdown. *Appl. Phys. Lett.* **49**, 669-671 (1986)
- [24] Chen, I.C., Holland, S.E., Hu, C.: Electrical breakdown in thin gate and tunneling oxides. *IEEE. T. Electron Dev.* **32**, 413-422 (1985)
- [25] Cheung, K.P.: A physics-based, unified gate-oxide breakdown model. *Electron Devices Meeting*, 1999. *IEDM Technical Digest International*, pp. 719-722
- [26] Cheung, K.P.: Unifying the thermal-chemical and anode-hole-injection gate-oxide breakdown models. *Microelectron. Reliab.* **41**, 193-199 (2001)
- [27] Chiabrera, A., Zitti, E.D., Costa, F., Bisio, G.M.: Physical limits of integration and information processing in molecular systems. *J. Phys. Appl. Phys.* **22**, 1571-1579 (1989)
- [28] Choi, Y.S., Park, H., Nishida, T., Thompson, S.E.: Reliability of HfSiON gate dielectric silicon MOS devices under [110] mechanical stress: Time dependent dielectric breakdown. *J. Appl. Phys.* **105**, 044503 (2009)
- [29] Degraeve, R., Groeseneken, G., Bellens, R., Depas, M., Maes, H.E. A consistent model for the thickness dependence of intrinsic breakdown in ultra-thin oxides. *IEEE* (1995).

- [30] Degraeve, R., Groeseneken, G., Bellens, R., Ogier, J.L., Depas, M., Roussel, P.J., Maes, H.E.: New insights in the relation between electron trap generation and the statistical properties of oxide breakdown. *IEEE. T. Electron Dev.* **45**, 904-911 (1998)
- [31] Degraeve, R., Kaczer, B., Groeseneken, G.: Reliability: a possible showstopper for oxide thickness scaling? *Semicond. Sci. Tech.* **15**, 436-444 (2000)
- [32] DiMaria, D.J.: *Insulating Films on Semiconductors*. Adam Hilger, Bristol (1991)
- [33] DiMaria, D.J.: Stress induced leakage currents in thin oxides. *Microelectron. Eng.* **28**, 63-66 (1995)
- [34] DiMaria, D.J., Arnold, D., Cartier, E.: Degradation and breakdown of silicon dioxide films on silicon. *Appl. Phys. Lett.* **61**, 2329-2331 (1992)
- [35] DiMaria, D.J., Cartier, E.: Mechanism for stress-induced leakage currents in thin silicon dioxide films. *J. Appl. Phys.* **78**, 3883-3894 (1995)
- [36] DiMaria, D.J., Cartier, E., Arnold, D.: Impact ionization, trap creation, degradation and breakdown in silicon dioxide films on silicon. *J. Appl. Phys.* **73**, 3367-3384 (1993)
- [37] DiMaria, D.J., Stasiak, J.W.: Trap creation in silicon dioxide produced by hot electrons. *J. Appl. Phys.* **65**, 2342-2356 (1989)
- [38] DiMaria, D.J., Stathis, J.H.: Trapping and trap creation studies on nitrated and reoxidized-nitrated silicon dioxide films on silicon. *J. Appl. Phys.* **70**, 1500-1509 (1991)
- [39] DiMaria, D.J., Stathis, J.H.: Anode hole injection, defect generation, and breakdown in ultrathin silicon dioxide films. *J. Appl. Phys.* **89**, 5015-5024 (2001)
- [40] DiStefano, T.H., Shatzkes, M.: Impact ionization model for dielectric instability and breakdown. *Appl. Phys. Lett.* **25**, 685-687 (1974)
- [41] DiStefano, T.H., Shatzkes, M.: Dielectric instability and breakdown in wide bandgap insulators. *J. Vac. Sci. Technol.* **12**, 37-46 (1975)
- [42] DiStefano, T.H., Shatzkes, M.: Dielectric instability and breakdown in SiO₂ thin films. *J. Vac. Sci. Technol.* **13**, 50-54 (1976)
- [43] Dumin, D.J.: Oxide wearout, breakdown, and reliability. *Int. J. High Speed Electron. Syst.* **11**, 617-718 (2001)
- [44] Dumin, D.J., Mopuri, S.K., Vanchinathan, S., Scott, R.S., Subramoniam, R., Lewis, T.G.: High field related thin oxide wearout and breakdown. *IEEE. T. Electron Dev.* **42**, 760-772 (1995)
- [45] Fang, F.F., Howard, W.E.: Negative field-effect mobility on (100) Si surfaces. *Phys. Rev. Lett.* **16**, 797-799 (1966)
- [46] Feigl, F.J., Gale, R., Chew, H., Magee, C.W., Young, D.R.: Current-induced hydrogen migration and interface trap generation in aluminum-silicon dioxide-silicon capacitors. *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms* **1**, 348-354 (1984)
- [47] Fischetti, M.V.: Model for the generation of positive charge at the Si-SiO₂

- interface based on hot-hole injection from the anode. *Phys. Rev. B* **31**, 2099-2113 (1985)
- [48] Fischetti, M.V., Weinberg, Z.: Effect of gate metal and SiO₂ thickness on the generation of donor states at the Si-SiO₂ interface. *J. Appl. Phys.* **57**, 418-425 (1985)
 - [49] Foster, A.S., Sulimov, V.B., Lopez Gejo, F., Shluger, A.L., Nieminen, R.M.: Structure and electrical levels of point defects in monoclinic zirconia. *Phys. Rev. B* **64**, 224108 (2001)
 - [50] Fowler, A.B., Fang, F.F., Howard, W.E., Stiles, P.J.: Magneto-oscillatory conductance in silicon surfaces. *Phys. Rev. Lett.* **16**, 901 (1966)
 - [51] Fowler, R.H., Nordheim, L.W.: Electron emission in intense electric fields. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* **119**, 173-181 (1928)
 - [52] Frank, D.J., Dennard, R.H., Nowak, E., Solomon, P.M., Taur, Y., Wong, H.S.P.: Device scaling limits of Si MOSFETs and their application dependencies. *Proc. IEEE* **89**, 259-288 (2001)
 - [53] Gale, R., Feigl, F.J., Magee, C.W., Young, D.R.: Hydrogen migration under avalanche injection of electrons in Si metal-oxide-semiconductor capacitors. *J. Appl. Phys.* **54**, 6938-6942 (1983)
 - [54] García, A., Northrup, J.E.: Compensation of p-Type doping in ZnSe: The role of impurity-native defect complexes. *Phys. Rev. Lett.* **74**, 1131-1134 (1995)
 - [55] Ghetti, A.: 6 gate oxide reliability: physical and computational models. *Predictive simulation of semiconductor processing: status and challenges* 201 (2004)
 - [56] Ghetti, A., Bude, J., Silverman, P., Hamad, A., Vaidya, H.: Modeling and simulation of tunneling current in MOS devices including quantum mechanical effects. *Ieice. T. Electron.* 1175-1182 (2000)
 - [57] Ghetti, A., Liu, C.T., Mastrapasqua, M., Sangiorgi, E.: Characterization of tunneling current in ultra-thin gate oxide. *Solid State Electron* **44**, 1523-1531 (2000)
 - [58] Green, M.L., Gusev, E.P., Degraeve, R., Garfunkel, E.L.: Ultrathin ([less-than] 4 nm) SiO₂ and Si-O-N gate dielectric layers for silicon microelectronics: Understanding the processing, structure, and physical and electrical limits. *J. Appl. Phys.* **90**, 2057-2121 (2001)
 - [59] Gusev, E.P., Buchanan, D.A., Cartier, E., Kumar, A., DiMaria, D., Guha, S., Callegari, A., Zafar, S., Jamison, P.C., Neumayer, D.A.: Ultrathin high-K gate stacks for advanced CMOS devices. *Electron Devices Meeting, 2001, IEDM Technical Digest, International*, pp.20.1.1-20.1.4
 - [60] Harari, E.: Dielectric breakdown in electrically stressed thin films of thermal SiO₂. *J. Appl. Phys.* **49**, 2478-2489 (1978)
 - [61] Harrison, W.A.: Tunneling from an independent-particle point of view. *Phys. Rev.* **123**, 85-89 (1961)
 - [62] Heh, D., Vogel, E.M., Bernstein, J.B.: Impact of substrate hot hole injection on ultrathin silicon dioxide breakdown. *Appl. Phys. Lett.* **82**, 3242-3244

- (2003)
- [63] Hikita, M., Tajima, S., Kanno, I., Ishino, I., Sawa, G., Ieda, M.: High-field conduction and electrical breakdown of polyethylene at high temperatures. *Jpn. J. Appl. Phys.* **24**, 988-996 (1985)
 - [64] Hu, C., Lu, Q.: A Unified Gate Oxide Reliability Model. Reliability Physics Symposium Proceedings, 1999. 37th Annual. 1999 IEEE International, pp. 47-51
 - [65] Huang, N.Y., She, J.C., Chen, J., Deng, S.Z., Xu, N.S., Bishop, H., Huq, S.E., Wang, L., Zhong, D.Y., Wang, E.G., Chen, D.M.: Mechanism responsible for initiating carbon nanotube vacuum breakdown. *Phys. Rev. Lett.* **93**, 075501 (2004)
 - [66] Imai, H., Arai, K., Imagawa, H., Hosono, H., Abe, Y.: Two types of oxygen-deficient centers in synthetic silica glass. *Phys. Rev. B* **38**, 12772-12775 (1988)
 - [67] Jeffery, S., Sofield, C.J., Pethica, J.B.: The influence of mechanical stress on the dielectric breakdown field strength of thin SiO₂ films. *Appl. Phys. Lett.* **73**, 172-174 (1998)
 - [68] Jeong, D.S., Schroeder, H., Breuer, U., Waser, R.: Characteristic electroforming behavior in Pt/TiO₂/Pt resistive switching cells depending on atmosphere. *J. Appl. Phys.* **104**, 123716-123718 (2008)
 - [69] Kadoshima, M., Hiratani, M., Shimamoto, Y., Torii, K., Miki, H., Kimura, S., Nabatame, T.: Rutile-type TiO₂ thin film for high-k gate insulator. *Thin Solid Films* **424**, 224-228 (2003)
 - [70] Kane, E.O.: Theory of tunneling. *J. Appl. Phys.* **32**, 83-91 (1961)
 - [71] Kimura, M.: Oxide breakdown mechanism and quantum physical chemistry for time-dependent dielectric breakdown. Reliability Physics Symposium, 1997. 35th Annual Proceedings, IEEE International, p. 190
 - [72] Klein, N.: A theory of localized electronic breakdown in insulating films. *Adv. Phys.* **21**, 605-645 (1972)
 - [73] Klein, N.: Electrical breakdown of insulators by one-carrier impact ionization. *J. Appl. Phys.* **53**, 5828-5839 (1982)
 - [74] Klein, N., Solomon, P.: Current runaway in insulators affected by impact ionization and recombination. *J. Appl. Phys.* **47**, 4364-4372 (1976)
 - [75] Laks, D.B., van de Walle, C.G., Neumark, G.F., Blöchl, P.E., Pantelides, S.T.: Native defects and self-compensation in ZnSe. *Phys. Rev. B* **45**, 10965-10978 (1992)
 - [76] Lee, J.C., Chen, I.C., Hu, C.: Modeling and characterization of gate oxide reliability. *IEEE. T. Electron Dev.* **35**, 2268-2278 (1988)
 - [77] Lenzlinger, M., Snow, E.H.: Fowler-Nordheim tunneling into thermally grown SiO₂. *J. Appl. Phys.* **40**, 278-283 (1969)
 - [78] Li, X., Tung, C.H., Pey, K.L.: The nature of dielectric breakdown. *Appl. Phys. Lett.* **93**, 072903 (2008)
 - [79] Lombardo, S., Stathis, J.H., Linder, B.P., Pey, K.L., Palumbo, F., Tung, C.H.: Dielectric breakdown mechanisms in gate oxides. *J. Appl. Phys.* **98**, 121301-121336 (2005)

- [80] Lui, W.W., Fukuma, M.: Exact solution of the Schrodinger equation across an arbitrary one-dimensional piecewise-linear potential barrier. *J. Appl. Phys.* **60**, 1555-1559 (1986)
- [81] Luo, X., Lin, S.P., Wang, B., Zheng, Y.: Impact of applied strain on the electron transport through ferroelectric tunnel junctions. *Appl. Phys. Lett.* **97**, 012905 (2010)
- [82] Luo, X., Wang, B., Zheng, Y.: First-principles study on energetics of intrinsic point defects in LaAlO_3 . *Phys. Rev. B* **80**, 104115 (2009)
- [83] Luo, X., Wang, B., Zheng, Y.: Microscopic mechanism of leakage currents in silica junctions. *J. Appl. Phys.* **106**, 073711 (2009)
- [84] Maserjian, J., Zamani, N.: Behavior of the Si/SiO_2 interface observed by Fowler-Nordheim tunneling. *J. Appl. Phys.* **53**, 559-567 (1982)
- [85] Mattila, T., Zunger, A.: Deep electronic gap levels induced by isovalent P and As impurities in GaN. *Phys. Rev. B* **58**, 1367-1373 (1998)
- [86] McPherson, J.W.: Stress dependent activation energy. Reliability Physics Symposium, 1986. 24th Annual, p. 12
- [87] McPherson, J.W., Baglee, D.A.: Acceleration factors for thin gate oxide stressing. Reliability Physics Symposium, 1985. 23rd Annual, p. 1
- [88] McPherson, J.W., Baglee, D.A.: Acceleration factors for thin oxide breakdown. *J. Electrochem. Soc.* **132**, 1903-1908 (1985)
- [89] McPherson, J.W., Khamankar, R.B., Shanware, A.: Complementary model for intrinsic time-dependent dielectric breakdown in SiO_2 dielectrics. *J. Appl. Phys.* **88**, 5351-5359 (2000)
- [90] McPherson, J.W., Mogul, H.C.: Impact of mixing of disturbed bonding states on time-dependent dielectric breakdown in SiO_2 thin films. *Appl. Phys. Lett.* **71**, 3721-3723 (1997)
- [91] McPherson, J.W., Mogul, H.C.: Underlying physics of the thermochemical E model in describing low-field time-dependent dielectric breakdown in SiO_2 thin films. *J. Appl. Phys.* **84**, 1513-1523 (1998)
- [92] McPherson, J.W., Reddy, V.K., Mogul, H.C.: Field-enhanced Si-Si bond-breakage mechanism for time-dependent dielectric breakdown in thin-film SiO_2 dielectrics. *Appl. Phys. Lett.* **71**, 1101-1103 (1997)
- [93] Merzbacher, E.: Quantum Mechanics, 2nd ed. John Wiley & Sons, Inc., New York (1970)
- [94] Meyer, W.K., Crook, D.L.: Model for oxide wearout due to charge trapping. Reliability Physics Symposium, 1983. 21st Annual, pp. 242-247
- [95] Moglestue, C.: Self-consistent calculation of electron and hole inversion charges at silicon-silicon dioxide interfaces. *J. Appl. Phys.* **59**, 3175-3183 (1986)
- [96] O'Dwyer, J.J.: Theory of high field conduction in a dielectric. *J. Appl. Phys.* **40**, 3887-3890 (1969)
- [97] O'Dwyer, J.J.: The theory of electrical conduction and breakdown in solid dielectrics. Clarendon Press, Oxford (1973)
- [98] Okada, K., Ota, H., Nabatame, T., Toriumi, A.: Dielectric breakdown in high-K gate dielectrics-Mechanism and lifetime assessment. Reliability

- Physics Symposium, 2007. Proceedings 45th Annual IEEE International, pp. 36-43
- [99] Okada, K., Yoneda, K.: A consistent model for time dependent dielectric breakdown in ultrathin silicon dioxides. Electron Devices Meeting, 1999. IEDM Technical Digest. International, pp. 445-448
 - [100] Pöykkö, S., Puska, M.J., Nieminen, R.M.: Ab initio study of fully relaxed divacancies in GaAs. *Phys. Rev. B* **53**, 3813-3819 (1996)
 - [101] Pals, J.A.: Quantization effects in semiconductor inversion and accumulation layers. Philips Research Reports Supplements 1972 (1972)
 - [102] Park, D.G., Cho, H.J., Lim, C., Yeo, I.S., Roh, J.S., Kim, C.T., Hwang, J.M.: Characteristics of Al₂O₃ gate dielectric prepared by atomic layer deposition for giga scale CMOS DRAM devices. VLSI Technology, 2000. Digest of Technical Papers. 2000 Symposium on, pp. 46-47
 - [103] Peng, J.Z., Fong, D.: High Density Semiconductor Memory Cell and Memory Array Using a Single Transistor. United States, Kilopass Technologies, Inc., Sunnyvale, CA (US), 6777757 B2 (2004)
 - [104] Peressi, M., Binggeli, N., Baldereschi, A.: Band engineering at interfaces: theory and numerical experiments. *J. Phys. D Appl. Phys.* **31**, 1273-1299 (1998)
 - [105] Pey, K.L., Tung, C.H., Radhakrishnan, M.K., Tang, L.J., Lin, W.H.: Dielectric breakdown induced epitaxy in ultrathin gate oxide—a reliability concern. Electron Devices Meeting, 2002. IEDM '02. Digest. International, pp.163-166
 - [106] Prange, R.E.: Tunneling from a many-particle point of view. *Phys. Rev.* **131**, 1083-1086 (1963)
 - [107] Radhakrishnan, M.K., Pey, K.L., Tung, C.H., Lin, W.H., Ong, S.H.: Physical analysis of reliability degradation in sub-micron devices. Electron Devices Meeting, 2001. IEDM Technical Digest. International, pp. 39.1.1-39.4.4
 - [108] Rana, F., Tiwari, S., Buchanan, D.A.: Self-consistent modeling of accumulation layers and tunneling currents through very thin oxides. *Appl. Phys. Lett.* **69**, 1104-1106 (1996)
 - [109] Register, L.F., Rosenbaum, E., Yang, K.: Analytic model for direct tunneling current in polycrystalline silicon-gate metal—oxide—semiconductor devices. *Appl. Phys. Lett.* **74**, 457-459 (1999)
 - [110] Schindler, C., Weides, M., Kozicki, M.N., Waser, R.: Low current resistive switching in Cu-SiO₂ cells. *Appl. Phys. Lett.* **92**, 122910-122913 (2008)
 - [111] Schrieffer, J.R.: Semiconductor Surface Physics. University of Penn. Press, Philadelphia, PA (1957)
 - [112] Shatzkes, M., Av-Ron, M.: Impact ionization and positive charge in thin SiO₂ films. *J. Appl. Phys.* **47**, 3192-3202 (1976)
 - [113] Solomon, P.: Breakdown in silicon oxide—a review. *J. Vac. Sci. Technol.* **14**, 1122-1130 (1977)
 - [114] Stathis, J.H.: Percolation models for gate oxide breakdown. *J. Appl. Phys.* **86**, 5757-5766 (1999)

- [115] Stathis, J.H.: Reliability limits for the gate insulator in CMOS technology. *IBM J. Res. Dev.* **46**, 265-286 (2002)
- [116] Stathis, J.H., DiMaria, D.J.: Reliability projection for ultra-thin oxides at low voltage. *Electron Devices Meeting, 1998. IEDM '98 Technical Digest, International*, pp. 167-170
- [117] Stathis, J.H., DiMaria, D.J.: Oxide scaling limit for future logic and memory technology. *Microelectron. Eng.* **48**, 395-401 (1999)
- [118] Stern, F.: Self-consistent results for n-type Si inversion layers. *Phys. Rev. B* **5**, 4891-4899 (1972)
- [119] Stern, F., Howard, W.E.: Properties of semiconductor surface inversion layers in the electric quantum limit. *Phys. Rev.* **163**, 816-835 (1967)
- [120] Suñé, J., Placencia, I., Barniol, N., Farres, E., Martin, F., Aymerich, X.: On the breakdown statistics of very thin SiO₂ films. *Thin Solid Films* **185**, 347-362 (1990)
- [121] Suñé, J., Wu, E.: Modeling the breakdown and breakdown statistics of ultra-thin SiO₂ gate oxides. *Microelectron. Eng.* **59**, 149-153 (2001)
- [122] Suñé, J., Miranda, E.: Post soft breakdown conduction in SiO₂ gate oxides. *Int. Electron. Devices Meeting (IEDM) 2000 Tech. Dig.*, pp. 553-556
- [123] Suñé, J., Mura, G., Miranda, E.: Are soft breakdown and hard breakdown of ultrathin gate oxides actually different failure mechanisms? *IEEE. Electr. Dev. Lett.* **21**, 167-169 (2000)
- [124] Suñé, J., Olivo, P., Ricco, B.: Self-consistent solution of the Poisson and Schrödinger equations in accumulated semiconductor-insulator interfaces. *J. Appl. Phys.* **70**, 337-345 (1991)
- [125] Suñé, J., Wu, E.Y.: Hydrogen-release mechanisms in the breakdown of thin SiO₂ films. *Phys. Rev. Lett.* **92**, 087601 (2004)
- [126] Suehle, J.S., Chaparala, P., Messick, C., Miller, W.M., Boyko, K.C.: Field and temperature acceleration of time-dependent dielectric breakdown in intrinsic thin SiO₂. *Reliability Physics Symposium, 1994. 32nd Annual Proceedings., IEEE International* pp. 120-125
- [127] Takagi, S., Yasuda, N., Toriumi, A.: A new I-V model for stress-induced leakage current including inelastic tunneling. *IEEE Trans. Electron Dev.* **46**, 348-354 (1999)
- [128] Tanaka, T., Matsunaga, K., Ikuhara, Y., Yamamoto, T.: First-principles study on structures and energetics of intrinsic vacancies in SrTiO₃. *Phys. Rev. B* **68**, 205213 (2003)
- [129] Tang, L.J., Pey, K.L., Tung, C.H., Radhakrishnan, M.K., Lin, W.H.: Gate dielectric-breakdown-induced microstructural damage in MOSFETs. *IEEE Trans. Dev. Mater. Reliab.* **4**, 38-45 (2004)
- [130] Timp, G., Bude, J., Bourdelle, K. K., Garno, J., Ghetti, A., Gossmann, H., Green, M.L., Forsyth, G., Kim, Y., Kleiman, R., Klemens, F., Kornblit, A., Loch stampfor, C., Mansfield, W., Moccio, S., Sorsch, T., Tennant, D. M., Timp, W., and Tung, R.: The Ballistic Nano-transistor. *Electron Devices Meeting, 1999. IEDM Technical Digest. International* pp. 55-58
- [131] Tung, C.H., Pey, K.L., Lin, W.H., Radhakrishnan, M.K.: Polarity-

- dependent dielectric breakdown-induced epitaxy (DBIE) in Si MOSFETs. *IEEE Electron Dev. Lett.* **23**, 526-528 (2002)
- [132] Umezawa, N., Shiraishi, K., Ohno, T., Watanabe, H., Chikyow, T., Torii, K., Yamabe, K., Yamada, K., Kitajima, H., Arikado, T.: First-principles studies of the intrinsic effect of nitrogen atoms on reduction in gate leakage current through Hf-based high-k dielectrics. *Appl. Phys. Lett.* **86**, 143507-143503 (2005)
 - [133] Vermeer, J.: On the relation between ionic conductivity and breakdown strength of glass. *Physica* **22**, 1269-1278 (1956)
 - [134] von Hippel, A.: Electric Breakdown of Solid and Liquid Insulators. *J. Appl. Phys.* **8**, 815-832 (1937)
 - [135] Wagner, K.W.: The physical nature of the electrical breakdown of solid dielectrics. *Trans. Am. Insti. Electr. Eng.* **41**, 288-299 (1922)
 - [136] Waser, R., Dittmann, R., Staikov, G., Szot, K.: Redox-based resistive switching memories-nanoionic mechanisms, prospects, and challenges. *Adv. Mater.* **21**, 2632-2663 (2009)
 - [137] Weibull, W.A.: A statistical distribution function of wide acceptability. *J. Appl. Mech.* **18**, 293-297 (1951)
 - [138] Weinberg, Z.A.: On tunneling in metal-oxide-silicon structures. *J. Appl. Phys.* **53**, 5052-5056 (1982)
 - [139] Whitehead, S.: *Dielectric Breakdown of Solids*. Clarendon Press, Oxford (1953)
 - [140] Wintle, H.J.: Electrothermal breakdown: Analytic solutions for currents and fields. *J. Appl. Phys.* **52**, 4181-4185 (1981)
 - [141] Wu, E., Vayshenker, A., Nowak, E., Sué, J., Vollertsen, R.P., Lai, W., Harmon, D.: Experimental evidence of TBD power-law for voltage acceleration factors for ultra-thin gate oxide. *IEEE Trans. Electron Dev.* **49**, 2244-2253 (2002)
 - [142] Wu, E.Y., Abadeer, W.W., Han, L.K., Lo, S.H., Hueckel, G.R.: Challenges for Accurate Reliability Projections in the Ultra-Thin Oxide Regime, *IEEE* (1999)
 - [143] Yang, T.C., Saraswat, K.C.: Effect of physical stress on the degradation of thin SiO₂ films under electrical stress. *IEEE Trans. Electron Devi.* **47**, 746-755 (2000)
 - [144] Zafar, S., Callegari, A., Gusev, E., Fischetti, M.V.: Charge trapping in high k gate dielectric stacks. *Electron Devices Meeting. IEDM '02. Digest, International*, pp. 517-520 (2002)
 - [145] Zhang, S.B., Northrup, J.E.: Chemical potential dependence of defect formation energies in GaAs: Application to Ga self-diffusion. *Phys. Rev. Lett.* **67**, 2339-2342 (1991)
 - [146] Zheng, Y., Woo, C.H.: Giant piezoelectric resistance in ferroelectric tunnel junctions. *Nanotechnology* **20**, 075401 (2009)
 - [147] Zhuravlev, M.Y., Sabirianov, R.F., Jaswal, S.S., Tsymbal, E.Y.: Giant electroresistance in ferroelectric tunnel junctions. *Phys. Rev. Lett.* **94**, 246802 (2005)

Index

A

ab initio 345
Accumulation layer 449, 450
AFM-induced electric field 152
Anisotropic elastic materials 60
Anode-hole-injection (AHI) 464
Anode-hydrogen-release (AHR) 511
Asymptotic solutions 203–205

B

Background dielectric constant 223, 233, 351
BaTiO₃ 353, 370
Bending modes 419
Bifurcation 212, 260
Bond breakage 446, 454
Boundary conditions 475, 479
Breakdown spot 511, 512

C

Capacitance versus voltage (CV) measurements 445
Charge ordering 392
Charge separation 464
Charge state level 469
Charge to breakdown (Q_{BD}) 509
Coercive field 163, 165
Compensation charges 197, 218
Compliant substrates 269, 273
Composite 378–380
Conformal transformation 29
Constitutive equations 8–11
Cooling critical thickness 201

Correlation length 348, 349
Crack tip fields 40
Cracks 40
Critical density of defects 456
Critical dimensions 171, 180, 194
Critical exponent 214
Critical thickness 223
Critical thickness for dislocation generation 273, 299
Cumulative failure rate 459, 460
Curie temperature 209, 223
Curie-Weiss law 211

D

Defect generation rate 456
Defects 456
Degradation 469, 471
Density functional theory 492
Depolarization energy 147, 161
Depolarizing field 212
Dielectric breakdown 443
Dielectric permittivity 118, 149
Dielectric substrates 197, 259
Dielectric susceptibility 343
Dielectric breakdown-induced epitaxy (DBIE) 447, 511
Dielectric breakdown-induced migration (DBIM) 513
Differential equations 6, 16
Differential model 112
Dilute model 103
Direct tunneling (DT) 452
Dislocations and line charges 52

Displacement boundaries 101

Domain growth 166

Domain shrinkage 188

E

1/*E* model 446

Effective Hamiltonian method 345

Effective material constants 103

Effective-mass approximation 472

Eigenstrains 53, 83

Elastic strain energy 113

Electric displacement 117

Electric field 5

Electromechanical resonance 381, 402

Electron affinity 470

Electrostriction 114, 138

Electrostrictive coefficient 211, 221

Ellipsoidal inclusions 118

Energy expressions 162

Energy formulation 138

Epitaxial stress 186, 194

Equivalent inclusion method 53

Eshelby's tensor 117

Evolution equations 166

Extrapolation lengths 182, 233

F

Ferroelectric composites 97

Ferroelectric domain 149

Ferroelectric films 215

Ferroelectric materials 235, 257

Ferroelectric random access memory
147

Ferroelectric superlattices 231

Ferroelectric thin film 234, 262

Ferromagnetic resonance 382, 418

Field accelerating factors 445, 454

First-principle 466, 489

Formation energy 312

Fourier integral 56

Fowler-Nordheim tunneling (FNT)
452

Free energy 497

Free-energy density 195

G

Gate oxide 493

Geometric distortion 386

Giant electro-resistance 180

Giant piezoelectric resistance (GPR)
effect 493

Ginzburg–Landau equation 186

Gradient coefficient 240

Green's function 6, 57

Griffith crack 30, 40

H

Hard breakdown (HBD) 454

Heterostructure 5

Hole affinity 470

I

Impact ionization 494

Impact-drift (ID) models 463

Impact-recombination (IR) models
463

Incompatibility 383

Inhomogeneities 35

Interface 43

Intrinsic free energy 113, 116

Inversion layer 446, 449

J

J-integral 30, 42

L

Landau free energy expansion 113,
187

Landau phase transition theory 273,
276

Landau-Ginzburg-Devonshire model
188, 244

Lattice dynamic theory 349

Leakage current 445

Lekhnitskii formalism 24

Liapunov criterion 149

M

Macroscopic properties 113

Magnetoelectric 401, 407, 418

Magnetostrictive 379, 400
 Manipulation 402, 415
 Matrix 473, 483
 Mechanical boundary conditions 183, 186
 Mechanical-stress-induced breakdown model 503
 Metal–oxide–semiconductor (MOS) structure 443
 Micromechanics models 97
 Minimum size of domain 170
 Misfit dislocations 186
 Misfit strain 209, 218
 Monte Carlo simulations 345
 Mori-Tanaka mean field model 108
 MOSFET 511
 Multiferroic materials 377
 Multilayers 187, 404
 Multi-step potential approximation 484

N

Nanograins 186
 Nanoisland 186, 257
 Nanotubes 184
 Nanowires 183
 Neutron scattering 334
 Newnham's connectivity theory 98
 Non-destructive read out 180
 Nonlocal theory 30
 Nonvolatility 147

O

Order of transition 208, 215
 Open-circuit boundary conditions 188

P

Paraelectric phase 116
 Paraelectric-ferroelectric phase transition 131
 PbTiO_3 139, 180
 Percolation models 447
 Phase diagrams 125
 Phase field model 188
 Phase transition 269
 Phase transition temperature 273

Phenomenological approach 322
 Phenomenological theory 323
 Piecewise-linear potential approximation 481
 Piezoelectric bending response 286
 Piezoelectric materials 5
 Polarization gradients 322
 Polarization switching 143
 Potential gradient 196
 Quantum yield of impact ionization 464

R

Radon transform 140
 Reliability 444
 Reliability projections 459, 460
 Renormalization theory 323
 Resonance frequency 419

S

Self-consistent model 107
 Self-consistent Schrödinger's and Poisson's equations 472
 Single phase 382
 Short-circuit boundary conditions 188, 197
 Size effects 387
 Soft breakdown (SBD) 454
 Soft phonon modes 350, 353
 Spiral spin order 381, 385
 Spontaneous polarizations 113
 Stability analysis 168
 Statistical mechanics 358, 363
 Steady-state 212, 414
 Strain engineering 418, 493
 Stress boundaries 100
 Stress induced leakage current (SILC) 448
 Stress intensity factors 80
 Stripe domains 180
 Strohm formalism 19
 Surface charge density 18, 196
 Surface effects 196
 Surface tension 240
 Susceptibility 248

Switching behavior 273, 286
Switching charge state level 470

T

Thermal breakdown model 502
Thermodynamic functions 324
Thermodynamic potentials 3, 10
Time dependent dielectric breakdown (TDDB) 454
Time to breakdown (T_{BD}) 445
Time-zero dielectric breakdown (TZDB) 454
TPOP approach 190
Transformation strain 194, 197
Transition temperature 198, 200

Transmission coefficient 480
Transverse Ising model 358
Transversely isotropic materials 40
Trap assisted tunneling (TAT) 452
Tunable 419, 429
Tunneling current 445
Tunneling junctions 181

V

Volume average 98

W

Weibull distribution 447
WKB approximation 480