Nanostructure Science and Technology
Series Editor: David J. Lockwood

Yoshitaka Umeno
Takahiro Shimada
Yusuke Kinoshita
Takayuki Kitamura

Multiphysics in Nanostructures

Springer

Nanostructure Science and Technology

Series editor

David J. Lockwood, FRSC
National Research Council of Canada
Ottawa, Ontario, Canada

More information about this series at http://www.springer.com/series/6331

Yoshitaka Umeno · Takahiro Shimada
Yusuke Kinoshita · Takayuki Kitamura

Multiphysics in Nanostructures

Springer

Yoshitaka Umeno
The University of Tokyo
Tokyo
Japan

Yusuke Kinoshita
Yazaki Corporation
Susono
Japan

Takahiro Shimada
Kyoto University
Kyoto
Japan

Takayuki Kitamura
Kyoto University
Kyoto
Japan

ISSN 1571-5744 ISSN 2197-7976 (electronic)
Nanostructure Science and Technology
ISBN 978-4-431-56571-0 ISBN 978-4-431-56573-4 (eBook)
DOI 10.1007/978-4-431-56573-4

Library of Congress Control Number: 2017946025

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer Japan KK
The registered company address is: Chiyoda First Bldg. East, 3-8-1 Nishi-Kanda, Chiyoda-ku, Tokyo 101-0065, Japan

Preface

Materials possess various properties such as magnetism, electric conductivity, ferroelectricity, photonics, thermal conductivity, mechanical properties, to name a few. In general, such properties do not alter independently of one another. Instead, strong couplings and interplay between these properties of materials can be noted. This is not surprising because various physical properties originate from the electronic structure of a material, i.e., a change in the electronic structure of a material will lead to changes in the physical properties of the material. Physical properties (usually representing functional aspects) can also be coupled with mechanical behaviors such as strain and deformation because a change in atomistic configuration brings about a change in electronic structure. The phenomena where multiple physical behaviors are intertwined is called 'multiphysics', which draws increasing interest due to its link to a wide range of possible applications such as strain engineering—which exploits the novel functions of devices by applying strain to them.

Multiphysics is significant and important in materials with nanostructures—nanometer-sized, low-dimensional materials such as ultrathin films, nanowires, nanoclusters, etc. Indeed, there has been a growing trend in fabricating nanodevices which utilize the interplay between the multiple properties of materials. The advent of first-principles calculation methods has been the driving force behind this trend because its methods enable highly accurate predictions of the properties and behaviors of nanostructures. The major drawback of first-principles calculations is surmountable due to the limited dimensions of the systems in question. There have been, therefore, a substantial number of theoretical studies, many of which have been based on first-principles calculation methods, to acquire an in-depth understanding and knowledge about 'multiphysics in nanostructures', this trend is expected to gain even more momentum in the following years and decades.

The purpose of this book is to systematically review recent advances in the theoretical investigations of the multiphysics phenomena in nanostructured materials by introducing a wide variety of studies including elastic strain engineering. This book mainly covers broad topics on the remarkable properties of multiphysics in low-dimensional nanoscale components, investigated by means of first-principles density functional theory. The book also includes studies based on semi-empirical

electronic structure calculations such as the tight-binding method. These methods help us understand the complicated nature of non-linear multiphysics couplings due to quantum mechanical effects. We believe that this book serves as motivation for readers to explore the rapidly expanding world of multiphysics in nanostructures, which itself paves the way for exploiting and designing novel functionalities at the nanoscale. We hope readers will be interested in this exciting multidisciplinary field and will be motivated to get involved in this promising research area.

We extend our sincere gratitude to Ms. Taeko Sato for her continuous and considerate support that has enabled the publication of this book.

Tokyo, Japan — Yoshitaka Umeno
Kyoto, Japan — Takahiro Shimada
Susono, Japan — Yusuke Kinoshita
Kyoto, Japan — Takayuki Kitamura

Contents

Chapter 1
Introduction

Abstract Multiphysics, which is the phenomenon where different physical properties are correlated with each other, is drawing much attention not only of scientific researchers but also of engineers. In the discussion of the physical properties in nanostructures, there are two targets, namely nanounderstructure and nanoshape. Classification of nanostructures is overviewed with an example of ferroelectric nanostructures. The scope of this book, mainly focusing on the investigation of nanoshaped components, is clarified.

Keywords Multiphysics · Nanostructure · Shape · Understructure

Insatiable demands have been imposed for the miniaturization and high performance for small technological devices such as electronic devices, micro-/nanosensors, and micro-/nanoelectromechanical systems (MEMS/NEMS). This trend demands high-density integration and multifunctions in their components in the nanometer scale. On the other hand, the rapid development in the manufacturing engineering has made production of components even atomic-scale components possible. Thus, atomic-scale design of functionality is required in near future.

As the dissimilar materials are composed in a device, the material property is influenced by the surrounding environment. For example, the ferroelectric property is sensitive to the mechanical strain, which is often brought in nanocomponents by the lattice mismatch between the neighboring materials. The interaction among physical properties, e.g., mechanical, electric, electronic, and magnetic ones, is called "multiphysics" and attracts many researchers and engineers in the field of materials science. Although there are many excellent textbooks available for individual behavior, explanatory books on the multiphysics are still not enough published yet. It is because many research works are going on and the scientific knowledge is expanding now. Here, it should be noted that the multiphysics properties possess sensitive correlation with each other through the boundary condition. As nanocomponents are under their characteristic boundary conditions, they might exhibit peculiar property in the multiphysics interaction. One of the unique aims in this book is that we spotlight the multiphysics in nanometer-scale structures.

Y. Umeno et al., *Multiphysics in Nanostructures*, Nanostructure Science and Technology, DOI 10.1007/978-4-431-56573-4_1

When we discuss the property in nanometer scale, there are two targets: (A) large material with nanounderstructure and (B) material in nanoshape. The former means multiscale investigation for understanding the hierarchical effect, while the latter focuses solely on the nanoscale properties. In the context of multiphysics in nanometer scale, these share common fundamentals; however, the effects of boundary conditions are remarkably different. Second unique point in this book is that the standpoint is mainly on the investigation and quest of (B), the nanoshaped components.

The structure creates various boundary conditions, which might bring about useful functions in terms of industrial application on the nanoshaped materials. In other words, geometric factor provides key on the nanometer-scale multiphysics. Here, material structure can be generally classified into two categories such as (1) "shape" (external geometry of material) and (2) "understructure" (internal geometry). The former is represented by low-dimensional components: films, bars/wires/tubes, and dots, where the surfaces/interfaces, edges, and vertexes play featured role in the characteristic material property. The latter in the atomic scale points out the effect of low-dimensional defects in crystallographic materials such as grain boundaries, dislocations, vacancies, and their networks. Both of (1) and (2) show critical influence on the multiphysics.

The physical quantities usually originated from the lattice-level or electron-level distortion have their directional behavior in a material under stable equilibrium condition. For example, the ferroelectricity in perovskite materials is caused by displacements of atoms inside a crystallographic unit cell and its spontaneous polarization is a vector at the lattice level in principle. Then, the existence of not only ferroelectric phase (FE), unidirectional distribution of moments, but also complex distortion, antiferrodistortive phase (AFD), and rotation of the oxygen octahedral is reported. Although the AFD distortion does not exist in a bulk, it emerges coupling with FE in nanostructures. This means that there exists lattice-level understructure, an inner geometric factor.

Moreover, the polarization usually forms polydomain structure. This points out that the polarization possesses topological spatial pattern. In particular, in a nanomaterial, it has been reported that exotic patterns come up. A vortex pattern, for example, is formed in a nanodot due to its surface effect. This provides another aspect of inner geometry, the understructure.

In short, the structural effects are categorized by the shape and understructure as summarized in Fig. 1.1. The latter is classified into the defect (atomic-level imperfection), the lattice-level distortion, and the topological pattern. The factors in Fig. 1.1 are nonlinearly coupled with each other. It should be noted that lattice-level distortion is switched by the external field and this brings temporal rearrangement of polarization pattern, temporal effect. The adventurous characteristic feature of this book is on the trial to approach the complexity of interaction.

When we consider the multiphysics, each quantity has such complexity explained above. The mechanical strain (stress) is a second-rank tensor so that the deformation pattern is more complex. Understanding of the multiphysics in structures means to clarify the interactions between the factors in each physical property

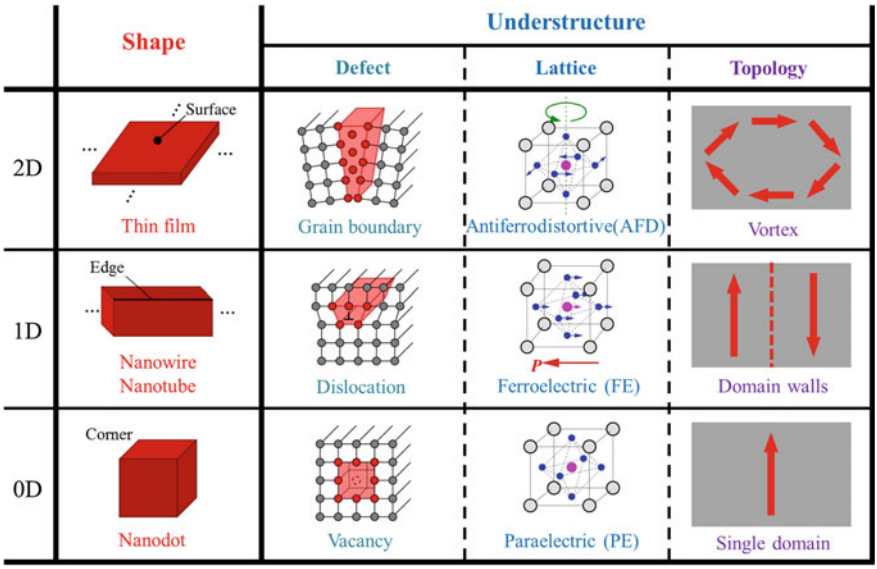

Fig. 1.1 Classification of ferroelectric nanostructures: (*Left side*) Shape of nanostructures including thin film (2D), nanowire and nanotube (1D), and nanodot (0D). (*Right side*) Understructure in ferroelectrics classified further into defect (grain boundary, dislocation, and vacancy), lattice-level distortion (ferroelectric, antiferrodistortive, and paraelectric), and polarization topology (single domain, domain walls, and vortex patterns)

shown in Fig. 1.1. Thus, the interaction extends diversely and the complexity provides us fertile field for the investigation.

Atomic simulations have made great progress in recent years. Ab initio (first-principles) calculation, which numerically solves the feature of electrons on the basis of quantum mechanics, has provided useful tool to analyze the material property at the atomic/electronic scale. Not only the enhancement of computer power but also the progress in numerical techniques has greatly contributed to the advancement in the quest of materials behavior. Considering the diversity and complexity of subjects, the approach inquired in this book is confined in the ab initio simulations and related techniques while excellent experimental works and other analytical efforts have been reported.

We try to describe the fundamentals of geometric effect in nanometer scale on the multiphysics characteristics in this book. As the nanoworld possesses enormous potential for more research on unexplored behavior and future industrial applications, it would be our great pleasure if some of readers get inspiration.

Chapter 2
Methodology of Quantum Mechanics/Atomic Simulations

Abstract Methodology for electronic structure calculations is overviewed. We place focus on first-principles calculation method based on the density-functional theory, which is most widely used for calculations of multiphysics in nanostructures. We explain the formulation for first-principles calculation method using pseudopotentials, which is a common practice in this field. Relatively advanced calculation methods including Beyond LDA approaches are also covered. Semi-empirical electronic structure calculation methods and empirical atomistic models are briefly introduced as well.

Keywords First-principles · Density-functional theory · Electronic structure calculation · Atomic model simulation

2.1 Method for Electronic Structure Calculation

Theoretical methods to obtain the electronic structure in solids and molecules have relatively a long history, which started in 1920s. In various fields, different methods suitable for each purpose have been proposed and improved. Needless to say, behind the advance of the increasingly sophisticated methods is the remarkable improvement of computational proficiency, which has enabled complicated and large-scale calculations.

Theories to obtain the electronic structure are basically based on the Schrödinger equation. When the method does not require empirical parameters and the calculation can be performed in a purely theoretical manner, it is called a first-principles calculation. "Ab initio" (Latin word meaning "from the beginning") is often used as a word equivalent to "first-principles" but there is a controversy whether these two words are exchangeable (see Appendix). It should be noted that first-principles (ab initio) calculations do *not* mean that no approximation is employed. Although first-principles calculations are basically regarded as rigorous and reliable and one can usually obtain the fundamental properties of materials in a good accuracy, one should well understand the introduced approximations and their limitations.

Y. Umeno et al., *Multiphysics in Nanostructures*, Nanostructure Science and Technology, DOI 10.1007/978-4-431-56573-4_2

There are a number of theoretical methods for electronic structure calculation. It is out of the scope of this book to cover all the methods proposed thus far. Classification of the methods may also be difficult because they are intermingled in the process of development. In this chapter, therefore, we explain about only some methods that are important and often used for the evaluation of the property of solids. Here we basically focus on the pseudopotential method with plane wave basis calculation scheme based on the density-functional theory (DFT) [1] which is mainly used for first-principles calculations for multiphysics of nanostructures introduced in the following chapters, while some of other methods are briefly covered.

2.1.1 Fundamental Approximations for Electronic Structure Calculation

In the electronic structure calculation, two fundamental approximations are usually employed: Born–Oppenheimer (adiabatic) and independent-electron approximations [2].

The adiabatic approximation assumes that electrons always take the ground state at any given configuration of atoms. In other words, the electron state is optimized fast enough to immediately follow varying atom configuration. This approximation is reasonable considering the fact that the mass of an electron is about three orders of magnitude lighter than that of an atom unless the change in atom configuration is extremely rapid. With this approximation, one can treat the behavior of electrons and that of atoms separately and independently, i.e., the electron state (and any physical quantity associated with it) can be obtained as the optimal solution to the given atom configuration, which is treated as a fixed environment.

An electron receives not only Coulomb interaction with its nucleus but also Coulomb and exchange–correlation interactions with other electrons. However, it is tremendously difficult to rigorously treat and solve the many-body problem of electrons. A way to circumvent this obstacle is to substitute the many-body interaction with a one-electron effective potential, i.e., it is regarded as a one-electron problem in a potential that represents contributions from other electrons. While this concept, which is called the independent-electron approximation, seems to be a rather drastic and bold approximation, it is often employed in most of the electron structure calculation methods, and it usually leads to successful results in various problems.

2.1.2 Hartree–Fock Method

The Hartree–Fock (HF) method [3] has been developed mainly for the electronic structure calculation of molecules or cluster systems. The molecular orbital

(MO) method calculation, which expresses wave functions with molecular orbitals (localized functions around atoms), is often performed based on the HF method. Here, the concept of the HF method is briefly explained.

Electrons are fermions, which are described by Fermi–Dirac statistics and follow the Pauli exclusion principle. Therefore, restrictions are imposed on the form of many-electron wave functions. We denote $\xi_i \equiv (\boldsymbol{r}_i, \sigma_i)$, where $\boldsymbol{r}_i$ is the coordinate of electron i and $\sigma_i (= \pm 1)$ its spin coordinate. The wave function of a N-electron system is written as $\Psi(\xi_1, \xi_2, \cdots, \xi_N)$. The wave function must change its sign when coordinates of two electrons are exchanged, i.e.,

$$\Psi\big(\cdots \xi_i \cdots \xi_j \cdots\big) = -\Psi\big(\cdots \xi_j \cdots \xi_i \cdots\big) \tag{2.1}$$

Now we express the wave function with one-electron orbitals, $\phi_{a_1}(\xi), \phi_{a_2}(\xi), \cdots$, as

$$\Psi_N(\xi_1, \cdots, \xi_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \phi_{a_1}(\xi_1) & \phi_{a_2}(\xi_1) & \cdots & \phi_{a_N}(\xi_1) \\ \phi_{a_1}(\xi_2) & \phi_{a_2}(\xi_2) & & \vdots \\ \vdots & & \ddots & \\ \phi_{a_1}(\xi_N) & & \cdots & \phi_{a_N}(\xi_N) \end{vmatrix} \tag{2.2}$$

This expression, called the Slater determinant, meets the Pauli exclusion principle. The exchange of two electrons alters the sign of wave function, meaning that electrons repel each other even without Coulomb interaction. While the many-electron state can in general be expressed as a linear combination of Slater determinants, the HF approximation uses one determinant. In the HF approximation, the one-electron Schrödinger equation (called the HF equation) becomes

$$-\frac{\hbar}{2m}\nabla^2 \Psi_i + V_H^i(\boldsymbol{r}_i)\Psi_i = \varepsilon_i \Psi_i \tag{2.3}$$

$$\begin{aligned} V_H^i(\boldsymbol{r}_i) = &-e^2 \sum_{\alpha} \frac{Z_\alpha}{|\boldsymbol{r}_i - \boldsymbol{R}_\alpha|} + e^2 \sum_{j \neq i} \int \frac{\left|\Psi_j\left(\boldsymbol{r}_j'\right)\right|^2}{\left|\boldsymbol{r}_i - \boldsymbol{r}_j'\right|} d\boldsymbol{r}_j' \\ &- e^2 \sum_{j \neq i} \left[\int \frac{\Psi_j^*\left(\boldsymbol{r}_j'\right)\Psi_i(\boldsymbol{r}_i)}{\left|\boldsymbol{r}_i - \boldsymbol{r}_j'\right|} d\boldsymbol{r}_j' \right] \Psi_j(\boldsymbol{r}_j)\delta_{\sigma_i\sigma_j} \end{aligned} \tag{2.4}$$

The second and third terms of the effective potential (V_H) are electron–electron Coulomb interaction and exchange interaction, respectively.

2.1.3 Density-functional Theory

Another approach to electronic structure calculation that started being used actively in 1970s is based on the density-functional theory (DFT) [1], which originates in the work by Hohenberg and Kohn (1964) that showed the total energy of a system can be expressed as a functional of charge density (ρ)

$$E_{\text{tot}} = \int V_{\text{ext}}(r)\rho(r)dr + T[\rho] + E_{ee}[\rho] \tag{2.5}$$

where V_{ext} is the external potential (potential of the nucleus ion), T the kinetic energy functional, and E_{ee} the electron–electron interaction functional. A practical calculation method was presented by Kohn and Sham (1965). According to the Kohn–Sham (KS) theory [4], one is to solve one-electron eigenvalue problem (Kohn–Sham equation)

$$[T + V_{\text{ext}} + V_H + V_{\text{XC}}]\psi_i(r) = \varepsilon_i\psi_i(r) \tag{2.6}$$

where V_H is the Hartree (Coulomb) potential and V_{XC} the exchange–correlation potential (details will be described in the next section). The energy functional in the Kohn–Sham theory is described as

$$\begin{aligned} E_{\text{tot}} = {} & T[\rho(r)] + \int V_{\text{ext}}(r)\rho(r)dr + \frac{1}{2}\int\int \frac{\rho(r)\rho(r')}{|r-r'|}dr'dr \\ & + E_{\text{XC}}[\rho(r)] \end{aligned} \tag{2.7}$$

The ground state can be therefore obtained by optimizing wave functions so that corresponding charge density minimizes the total energy. In other words, DFT is a charge density-based theory, while the HF method is a wave-function-based one.

Usually, the KS equation is written in the matrix form and solution (eigenvalues and eigenvectors) is found by an iterative method. As the Hamiltonian terms contain the contribution of charge density of electrons, solving the KS equation cannot be done only once but a so-called self-consistent loop (SCL) calculation is needed, i.e., the Hamiltonian is constructed with a provisional charge density distribution (ρ_{in}), and the KS equation is solved to obtain ψ, from which charge density is renewed (ρ_{out}). Then the created charge density constructs the new Hamiltonian, and this process is repeated until the equation becomes self-consistent ($\rho_{\text{in}} = \rho_{\text{out}}$).

Electronic structure calculation for crystals (periodic structures) is sometimes called "electronic band structure calculation." Various theoretical methods for band calculation have been proposed and developed. All-electron calculation methods, which obtain the states of both valence and core electrons, will be only briefly introduced in Sect. 2.2.8. Electronic band structure calculation based on DFT is also conducted with a plane wave basis set, where the pseudopotential method is usually employed to (explicitly) treat valence electrons only. One of the advantages

of such approach is its suitability for the calculation of forces exerted on atoms due to the Hellmann–Feynman theorem [5], which enables efficient structural optimization and obtains mechanical properties such as stress tensor and elastic constants. This approach is widely used owing to sophisticated simulation packages including VASP [6], CASTEP [7], ABINIT [8], and PWSCF [9]. The detail of the calculation method of DFT with pseudopotentials will be described in the next section.

2.2 First-Principles Calculation with Plane Wave Basis Set

In this section, we describe the DFT calculation method using a plane wave basis set. Hartree atomic units (Table 2.1) are used in the following description of the formalism. For the sake of brevity, we assume the nonmagnetic case without spin polarization unless otherwise stated.

The plane wave basis is not suitable for expressing wave functions with many nodes. The plane wave basis DFT calculations are therefore usually performed with employing the pseudopotential method, which makes the effective potential remarkably softer to reduce the basis set required to represent wave functions. For pseudopotential methods, see later section.

2.2.1 Kohn–Sham Equation

In DFT, the Kohn–Sham equation

$$[T + V_{\text{ext}} + V_{\text{H}} + V_{\text{XC}}]\psi_i(\boldsymbol{r}) = \varepsilon_i \psi_i(\boldsymbol{r}) \tag{2.8}$$

is to be solved to find the wave vector (eigenvector), $\psi_i(\boldsymbol{r})$, and the corresponding energy level (eigenvalue), ε_i, of one-electron state (i). T is the operator of the kinetic energy of the electron

$$T[\rho(\boldsymbol{r})] = \sum_{i}^{\text{occ}} \left\langle \psi_i \middle| -\frac{1}{2}\nabla^2 \middle| \psi_i \right\rangle \tag{2.9}$$

Table 2.1 Hartree and Rydberg atomic units and SI unit

	Hartree (a.u)	Rydberg (a.u)	SI
Mass	1	½	9.10965×10^{-31} kg
Time	1	½	2.41889×10^{-17} s
Energy	1	2	4.35982×10^{-18} J
Length	1	1	5.29177×10^{-11} m
Force	1	2	8.23892×10^{-8} N

and the charge density ρ is obtained as

$$\rho(\boldsymbol{r}) = 2\sum_{i}^{\text{occ}} |\psi_i(\boldsymbol{r})|^2 \tag{2.10}$$

Here "occ" means that the summation ranges over the occupied electron states. The other three terms ($V_{\text{eff}} \equiv V_{\text{ext}} + V_{\text{H}} + V_{\text{HC}}$) are the effective one-electron potential. V_{ext} represents the interaction between the electron and the "core ion," which consists of the nucleus and the surrounding core electrons in the framework of the pseudopotential method (see Sect. 2.2.4). V_{H} represents the electron–electron (Coulomb) interaction

$$V_{\text{H}} = \int \frac{\rho(\boldsymbol{r}')}{|\boldsymbol{r} - \boldsymbol{r}'|} d\boldsymbol{r}' \tag{2.11}$$

which is also called Hartree term. V_{XC} is the exchange–correlation potential

$$V_{\text{XC}} = \frac{\delta E_{\text{XC}}[\rho(\boldsymbol{r})]}{\delta \rho(\boldsymbol{r})} \tag{2.12}$$

where all quantum effect is included. The exact exchange–correlation functional ($E_{\text{XC}} = E_{\text{C}} + E_{\text{X}}$; E_{C} and E_{X} are the correlation and the exchange energies, respectively) is not known, and an approximation must be introduced to write this term (see, e.g., Sect. 2.2.2).

2.2.2 Local Density Approximation

To solve the Kohn–Sham equation, it is required to determine the exchange–correlation functional form. This is, however, a many-electron term and it is hardly possible to come up with the exact solution for general cases. The local density approximation (LDA) [4] was introduced to give a major breakthrough to this problem. LDA assumes that the gradient of the charge density with respect to space is modest and writes

$$E_{\text{XC}}[\rho(r)] = \int \varepsilon_{\text{XC}}(\boldsymbol{r})\rho(\boldsymbol{r}) d\boldsymbol{r} \tag{2.13}$$

where $\varepsilon_{\text{XC}}(\boldsymbol{r})$ is the exchange–correlation energy density, which is the sum of the exchange (ε_{X}) and correlation (ε_{C}) energy densities, i.e.,

$$\varepsilon_{\text{XC}}(\boldsymbol{r}) = \varepsilon_{\text{X}}(\boldsymbol{r}) + \varepsilon_{\text{C}}(\boldsymbol{r}) \tag{2.14}$$

The exchange–correlation potential is then

$$V_{\mathrm{XC}} = \frac{\delta E_{\mathrm{XC}}[\rho(\boldsymbol{r})]}{\delta\rho} = \varepsilon_{\mathrm{XC}}(\boldsymbol{r}) + \rho(\boldsymbol{r})\frac{\delta\varepsilon_{\mathrm{XC}}(\boldsymbol{r})}{\delta\rho(\boldsymbol{r})} \tag{2.15}$$

Although LDA is a rather bold approximation, it has been empirically known that the approximation works well for various cases. The reason for this was intensively discussed by Gunnarsson [10]. To perform analysis of magnetic materials, LDA was extended to the spin-polarized case, where majority and minority spins do not take an identical state. This is called the local spin density approximation (LSDA). Denoting the opposite spins by + and −, the Kohn–Sham equation becomes as follows:

$$\left[-\frac{1}{2}\nabla^2 + V_{\mathrm{eff}}^{\pm}(\boldsymbol{r})\right]\psi_i^{\pm}(\boldsymbol{r}) = \varepsilon_i^{\pm}\psi_i^{\pm}(\boldsymbol{r}) \tag{2.16}$$

$$V_{\mathrm{eff}}^{\pm}(\boldsymbol{r}) = V_{\mathrm{ext}}(\boldsymbol{r}) + V_{\mathrm{H}}(\boldsymbol{r}) + V_{\mathrm{XC}}^{\pm}(\boldsymbol{r}) \tag{2.17}$$

$$V_{\mathrm{XC}}^{\pm}(\boldsymbol{r}) = \frac{\delta E_{\mathrm{XC}}}{\delta\rho^{\pm}(\boldsymbol{r})} \tag{2.18}$$

Here,

$$\rho(\boldsymbol{r}) = \rho^{+}(\boldsymbol{r}) + \rho^{-}(\boldsymbol{r}) \tag{2.19}$$

$$\rho^{\pm}(\boldsymbol{r}) = \sum_{i}^{\mathrm{occ}}\left|\psi_i^{\pm}(\boldsymbol{r})\right|^2 \tag{2.20}$$

The exchange–correlation energy reads

$$E_{\mathrm{XC}}[\rho^{+}, \rho^{-}] = \int\rho(\boldsymbol{r})\varepsilon_{\mathrm{XC}}\{\rho^{+}(\boldsymbol{r}), \rho^{-}(\boldsymbol{r})\}d\boldsymbol{r} \tag{2.21}$$

Various expressions for the exchange and correlation energies have been suggested thus far [11]. Here we introduce an example, which is one of the most commonly used functions. In the following expressions, superscripts P and F denote paramagnetic and ferroelectric states, respectively. Kohn, Sham, and Gaspar gave an approximation for the exchange energy density as [4, 12]

$$\varepsilon_{\mathrm{X}}^{\mathrm{P}} = -\frac{3}{4\pi}\left(\frac{9\pi}{4}\right)^{1/3}\frac{1}{r_s} \tag{2.22}$$

and

$$\varepsilon_{\mathrm{X}}^{\mathrm{F}} = 2^{1/3}\varepsilon_{\mathrm{X}}^{\mathrm{P}} \tag{2.23}$$

where

$$r_s = \left(\frac{3}{4\pi\rho(\boldsymbol{r})}\right)^{1/3} \tag{2.24}$$

A functional form for the correlation energy density was proposed by Perdew and Zunger [13] based on the results of Monte Carlo calculations of electron gas by Ceperley and Alder [14] as

$$\varepsilon_{\mathrm{C}}^{i}(\boldsymbol{r}) = \begin{cases} -\dfrac{\gamma^i}{1+\beta_1^i\sqrt{r_s}+\beta_2^i r_s} & (r_s > 1) \\ A^i \ln r_s + B^i + C^i r_s \ln r_s + D^i r_s & (r_s \le 1) \end{cases} \tag{2.25}$$

where i = P or F. The parameters are given as follows: $\gamma^{\mathrm{P}} = -0.07115$, $\beta_1^{\mathrm{P}} = 1.0529$, $\beta_2^{\mathrm{P}} = 0.3334$, $A^{\mathrm{P}} = 0.01555$, $B^{\mathrm{P}} = -0.024$, $C^{\mathrm{P}} = 0.001$, $D^{\mathrm{P}} = -0.0116$, $\gamma^{\mathrm{F}} = -0.04215$, $\beta_1^{\mathrm{F}} = 1.3981$, $\beta_2^{\mathrm{F}} = 0.2611$, $A^{\mathrm{F}} = 0.007775$, $B^{\mathrm{F}} = -0.01345$, $C^{\mathrm{F}} = 0.00035$, and $D^{\mathrm{F}} = -0.0024$. For the spin polarization $\zeta(\equiv(\rho^+ - \rho^-)/\rho)$ between 0 and 1, the energy densities can be obtained using the spin interpolation formula by von Barth and Hedin [15], i.e.,

$$\varepsilon_{\mathrm{XC}}(\zeta) = \varepsilon_{\mathrm{XC}} + f(\zeta)\left(\varepsilon_{\mathrm{XC}}^{\mathrm{F}} - \varepsilon_{\mathrm{XC}}^{\mathrm{P}}\right) \tag{2.26}$$

$$f(\zeta) = \frac{(1+\zeta)^{4/3} + (1-\zeta)^{4/3} - 2}{2^{4/3} - 2} \tag{2.27}$$

LDA (LSDA) has been and still is widely used due to its good performance in spite of the simplicity. Nevertheless, it is known that the approximation does not work well for some cases. Some of the nontrivial, well-known problems are listed below [16].

(1) LDA tends to underestimate the lattice constant. (In general, however, this trait is not considered to be fatal because the deviation from the experimental value can be restricted within 1% in most cases.)
(2) Structure and magnetism of 3D transition metals cannot be reproduced well. For instance, the hcp (paramagnetic) structure of Fe is evaluated as the most stable state, as opposed to the bcc structure with the ferromagnetic state in reality.
(3) The band gap energy of semiconductors and insulators tends to be underestimated.

2.2.3 Generalized Gradient Approximation

Generalized gradient approximation (GGA) [17, 18], which is to take into account the spatial gradient of charge density, is a natural and reasonable extension of LDA. The exchange–correlation energy in GGA is expressed as

$$E_{\mathrm{XC}}[\rho^+, \rho^-] = \int f\{\rho^+, \rho^-, \nabla\rho^+, \nabla\rho^-\} d\boldsymbol{r} \tag{2.28}$$

GGA tends to improve various energy evaluations including total energies and structural energy differences. Evaluation of magnetism can be improved in some cases. In GGA, interatomic bonds are evaluated to be softer than in LDA, which leads to (over)correction to the underestimation of lattice constants by LDA.

Among GGA, functionals being used most frequently are PW91 proposed by Perdew and Wang [19], and PBE by Perdew et al. [18]. Omitting the details of its derivation, explicit forms of the PBE functional are given below. The correlation energy is given by

$$E_C^{\mathrm{PBE}}[\rho^+, \rho^-] = \int \rho \left[\varepsilon_c^{\mathrm{LDA}}(r_s, \zeta) + H(r_s, \zeta, t)\right] d\boldsymbol{r} \tag{2.29}$$

where $r_s = (3/4\pi\rho)^{1/3}$ is the Wigner–Seitz radius, $\zeta = (\rho^+ - \rho^-)/\rho$ is the relative spin polarization, and $t = |\nabla\rho|/2\phi k_s\rho$ is a dimensionless density gradient ($\phi(\zeta) = \left[(1+\zeta)^{2/3} + (1-\zeta)^{2/3}\right]/2$, $k_s = \sqrt{4k_F/\pi a_0}$). The gradient contribution is given by

$$H = \left(\frac{e^2}{a_0}\right)\gamma\phi^3 \ln\left\{1 + \frac{\beta}{\gamma}t^2\left[\frac{1 + At^2}{1 + At^2 + A^2t^4}\right]\right\} \tag{2.30}$$

$$A = \frac{\beta}{\gamma}\left[\exp\left\{\frac{-\varepsilon_c^{\mathrm{LDA}} a_0}{\gamma\phi^3 e^2}\right\} - 1\right]^{-1} \tag{2.31}$$

$$(e =, a_0 =, \beta = 0.066725, \gamma = (1 - \ln 2)/\pi^2)$$

The exchange energy is written as

$$E_X^{\mathrm{PBE}} = \int \rho \left[\varepsilon_X^{\mathrm{LDA}}(\rho) + F_X(s)\right] d\boldsymbol{r} \tag{2.32}$$

$$F_X(s) = 1 + \kappa - \kappa/\left(1 + \mu s^2/\kappa\right) \tag{2.33}$$

$$(\kappa = 0.804)$$

2.2.4 Pseudopotential Method and Norm-Conserving Pseudopotential

Potentials that electrons feel basically have a deep valley around the nucleus. Wave functions have therefore many nodes, which makes it impractical to represent such wave functions with plane wave basis sets. The pseudopotential method is to circumvent this problem by dealing with only valence electrons and constructing "effective" (pseudo-)potentials created by the nucleus and the core electrons (see Fig. 2.1). Core electrons, chemically inert, are considered to be fixed (frozen-core approximation), and only the Schrödinger equation for valence electrons is solved.

Pseudopotentials are constructed from the solution of the all-electron Schrödinger equation for an isolated atom and are put in use for condensed matters. Thus, it is essential for the pseudopotentials to possess good transferability (i.e., reliability of the potential when being put in different environment). In addition, the pseudopotentials are desired to be sufficiently "soft" so that only small basis sets are required.

The norm-conserving pseudopotential (NCPP) method was proposed by Hamann et al. [20], and various potentials were constructed with modifications later on [21–25]. NCPPs are constructed to meet the following conditions (ψ_l^{AE} and ε_l^{AE} are the wave function and the energy (eigenvalue) of the all-electron calculation, respectively. ψ_l^{PS} and ε_l^{PS} are the pseudo wave function and the corresponding (pseudo) energy, respectively):

(1) Pseudo wave functions (ψ_l^{PS}, where l is the angular momentum) are nodeless.
(2) Pseudo wave functions are identical with all-electron wave functions (ψ_l^{AE}) outside the core radius, r_{cl}.
(3) Energies (eigenvalues) of the pseudo wave functions, ε_l^{PS}, are identical with all-electron energies (ε_l^{AE}).
(4) The integral of pseudo charge density (norm of pseudo wave functions) inside the cutoff sphere is identical with that of true charge density, namely

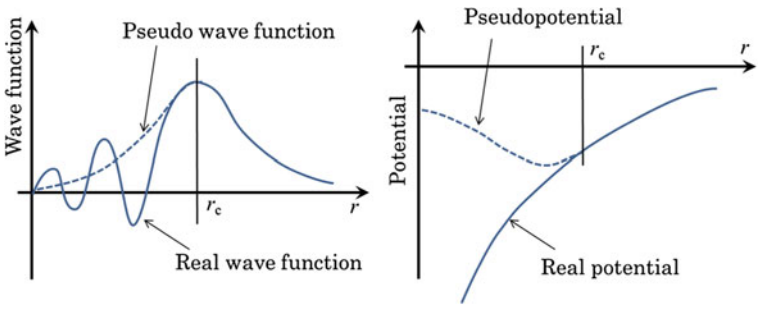

Fig. 2.1 Schematics explaining pseudopotential method

$$\int_0^{r_{cl}} \left| r\psi_l^{PS}(r) \right|^2 dr = \int_0^{r_{cl}} \left| r\psi_l^{AE}(r) \right|^2 dr \tag{2.34}$$

The following identity is obtained.

$$\left[-\frac{1}{2} r^2 \psi(r)^2 \frac{d}{d\varepsilon} \frac{d}{dr} \ln \psi(r) \right]_{r_{cl}} = \int_0^{r_{cl}} r^2 \psi(r)^2 dr \tag{2.35}$$

The left-hand side is unchanged between ψ_l^{AE} and ψ_l^{PS} if the above condition (4) is satisfied, i.e., the logarithmic derivative of the radial wave function, which describes the scattering property of electrons by the ion, is correct. This ensures the transferability of NCPPs.

2.2.5 Hamiltonian in NCPP

Here we show the Hamiltonian operator in the matrix form within the framework of the NCPP method [26]. Consider an object (crystal) cell, whose primitive translation vectors are $\boldsymbol{a}_1, \boldsymbol{a}_2,$ and $\boldsymbol{a}_3$. The reciprocal lattice is given as

$$\boldsymbol{G} = m_1 \boldsymbol{b}_1 + m_2 \boldsymbol{b}_2 + m_3 \boldsymbol{b}_3 \tag{2.36}$$

with

$$\boldsymbol{b}_1 = 2\pi \frac{\boldsymbol{a}_2 \times \boldsymbol{a}_3}{\boldsymbol{a}_1 \cdot (\boldsymbol{a}_2 \times \boldsymbol{a}_3)} \tag{2.37}$$

$$\boldsymbol{b}_2 = 2\pi \frac{\boldsymbol{a}_3 \times \boldsymbol{a}_1}{\boldsymbol{a}_2 \cdot (\boldsymbol{a}_3 \times \boldsymbol{a}_1)} \tag{2.38}$$

$$\boldsymbol{b}_3 = 2\pi \frac{\boldsymbol{a}_1 \times \boldsymbol{a}_2}{\boldsymbol{a}_3 \cdot (\boldsymbol{a}_1 \times \boldsymbol{a}_2)} \tag{2.39}$$

and $m_1, m_2,$ and m_3 are integers. We use the notation of a plane wave

$$|\boldsymbol{k} + \boldsymbol{G}\rangle = \frac{1}{\sqrt{\Omega}} \exp[i(\boldsymbol{k} + \boldsymbol{G}) \cdot \boldsymbol{r}] \tag{2.40}$$

where $\boldsymbol{k}$ is a sampling point in the Brillouin zone. Plane waves are orthonormal as

$$\langle \boldsymbol{k} + \boldsymbol{G} | \boldsymbol{k} + \boldsymbol{G}' \rangle = \frac{1}{\sqrt{\Omega}} \int_\Omega \exp[i(\boldsymbol{G} - \boldsymbol{G}') \cdot \boldsymbol{r}] d\boldsymbol{r} = \delta_{GG'} \tag{2.41}$$

(Ω is the volume of the entire crystal.) The wave function that has the nth eigenvalue for $\boldsymbol{k}$, $\psi_{kn}(\boldsymbol{r})$, is expanded into the plane wave basis as

$$\psi_{kn}(\boldsymbol{r}) = \sum_{G} C^{n}_{k+G} |\boldsymbol{k}+\boldsymbol{G}\rangle \tag{2.42}$$

Practically, the summation ranges over all $\boldsymbol{G}$'s with $\frac{1}{2}|\boldsymbol{k}+\boldsymbol{G}|^2$ smaller than a certain value (the cutoff value for plane waves). The charge density is given by

$$\rho(\boldsymbol{r}) = \sum_{G}\sum_{G'}\sum_{n}^{\text{occ}}\sum_{k} f_n f_k \frac{1}{\Omega} C^{n*}_{k+G'} C^{n}_{k+G} \exp[i(\boldsymbol{G}-\boldsymbol{G}')\cdot\boldsymbol{r}] \tag{2.43}$$

where f_n and f_k indicate the weight of each $\boldsymbol{k}$-point and the occupation number of the energy level n, respectively.

In the plane wave basis formalism, the Kohn–Sham equation becomes the eigenvalue problem whose eigenvectors are the expansion coefficients written as

$$\sum_{G'} H_{k+G,k+G'} C^{n}_{k+G'} = \varepsilon_{kn} C^{n}_{k+G} \tag{2.44}$$

where ε_{kn} is the eigenvalue for the nth state of the sampling point $\boldsymbol{k}$.

Here we show the elements of the Hamiltonian matrix

$$H_{k+G,k+G'} = \langle \boldsymbol{k}+\boldsymbol{G} \left| -\frac{1}{2}\nabla^2 + v_{\text{eff}}(\boldsymbol{r}) \right| \boldsymbol{k}+\boldsymbol{G}' \rangle \tag{2.45}$$

The kinetic energy term is

$$\langle \boldsymbol{k}+\boldsymbol{G} \left| -\frac{1}{2}\nabla^2 \right| \boldsymbol{k}+\boldsymbol{G}' \rangle = \frac{1}{2}|\boldsymbol{k}+\boldsymbol{G}|^2 \delta_{GG'} \tag{2.46}$$

The term for the Coulomb interaction with the ion core is split into the local term (V_{loc}), which is independent of the angular momentum, and the angular-dependent nonlocal term (V_{nloc}). The local term

$$\begin{aligned} &\langle \boldsymbol{k}+\boldsymbol{G} | V^{\text{pp}}_{loc}(\boldsymbol{r}) | \boldsymbol{k}+\boldsymbol{G}' \rangle \\ &= \frac{1}{\Omega}\int_{\Omega} V^{\text{pp}}_{loc}(r) \exp[-i(\boldsymbol{k}+\boldsymbol{G})\cdot\boldsymbol{r})] \exp[i(\boldsymbol{k}+\boldsymbol{G}')\cdot\boldsymbol{r})] d\boldsymbol{r} \\ &= V^{pp}_{loc}(\boldsymbol{G}-\boldsymbol{G}') \end{aligned} \tag{2.47}$$

is given by the coefficients of the Fourier transformation of $V_{\mathrm{loc}}^{\mathrm{pp}}(\boldsymbol{r})$, i.e.

$$V_{\mathrm{loc}}^{\mathrm{pp}}(\boldsymbol{G}) = \frac{1}{\Omega_{\mathrm{a}}} \sum_{a} \exp(-i\boldsymbol{G}\cdot\boldsymbol{r}_a) V_a^{\mathrm{pp,loc}}(\boldsymbol{G}) \tag{2.48}$$

$$V_{\mathrm{a}}^{\mathrm{pp,loc}}(\boldsymbol{G}) = \int_{\Omega} V_a^{\mathrm{pp,loc}}(r) \exp(-\boldsymbol{G}\cdot\boldsymbol{r}) d\boldsymbol{r} \tag{2.49}$$

Here, Ω_a is the volume of the simulation cell, and $V_a^{\mathrm{pp,loc}}(r)$ and $\boldsymbol{r}_a$ are the local pseudopotential and the position of atom a, respectively. The nonlocal part is

$$\begin{aligned} \left\langle \boldsymbol{k}+\boldsymbol{G} \middle| V_{\mathrm{nloc}}^{\mathrm{pp}}(\boldsymbol{r}) \middle| \boldsymbol{k}+\boldsymbol{G}' \right\rangle &= \frac{1}{\Omega_a} \sum_{a} \exp[-i(\boldsymbol{G}-\boldsymbol{G}')\cdot\boldsymbol{r}_a] V_a^{\mathrm{pp,nloc}}(\boldsymbol{k}+\boldsymbol{G}, \boldsymbol{k}+\boldsymbol{G}') \\ &= V_{\mathrm{nloc}}^{\mathrm{pp}}(\boldsymbol{k}+\boldsymbol{G}, \boldsymbol{k}+\boldsymbol{G}') \end{aligned} \tag{2.50}$$

Here,

$$\begin{aligned} &V_{\mathrm{nloc}}^{\mathrm{pp}}(\boldsymbol{k}+\boldsymbol{G}, \boldsymbol{k}+\boldsymbol{G}') \\ &= 4\pi \sum_{l} (2l+1) P_l(\cos\omega) \\ &\times \int V_{a,l}^{\mathrm{pp,nloc}}(r) j_l(|\boldsymbol{k}+\boldsymbol{G}|r) j_l(|\boldsymbol{k}+\boldsymbol{G}'|r) r^2 dr \end{aligned} \tag{2.51}$$

$V_{a,l}^{\mathrm{pp,nloc}}(r)$ is the nonlocal pseudopotential of atom a for angular momentum l, P_l the Legendre polynomial, and j_l the spherical Bessel function. ω is the angle between $\boldsymbol{k}+\boldsymbol{G}$ and $\boldsymbol{k}+\boldsymbol{G}'$.

The Fourier transformation of the charge density reads

$$\rho(\boldsymbol{r}) = \sum_{G} \rho(\boldsymbol{G}) \exp(i\boldsymbol{G}\cdot\boldsymbol{r}) \tag{2.52}$$

$$\rho(\boldsymbol{G}) = \frac{1}{\Omega} \int \rho(\boldsymbol{r}) \exp(-i\boldsymbol{G}\cdot\boldsymbol{r}) dr \tag{2.53}$$

The Coulomb term meets the Poisson equation,

$$\nabla^2 V_{\mathrm{coul}}(\boldsymbol{r}) = -4\pi\rho(\boldsymbol{r}) \tag{2.54}$$

therefore

$$\nabla^2 V_{\mathrm{coul}}(\boldsymbol{r}) = -4\pi \sum_{G} \rho(\boldsymbol{G}) \exp(i\boldsymbol{G}\cdot\boldsymbol{r}) \tag{2.55}$$

Thus, we get

$$V_{\mathrm{coul}}(\boldsymbol{r}) = 4\pi \sum_{G} \frac{\rho(\boldsymbol{G})}{|\boldsymbol{G}|^2} \exp(i\boldsymbol{G}\cdot\boldsymbol{r}) \tag{2.56}$$

and its Fourier transformation becomes

$$\begin{aligned} V_{\mathrm{coul}}(\boldsymbol{G}) &= \frac{1}{\Omega}\int_{\Omega} 4\pi \sum_{G'} \frac{\rho(\boldsymbol{G}')}{|\boldsymbol{G}'|^2} \exp(i\boldsymbol{G}'\cdot\boldsymbol{r}) \exp(-i\boldsymbol{G}\cdot\boldsymbol{r}) d\boldsymbol{r} \\ &= 4\pi \frac{\rho(\boldsymbol{G})}{|\boldsymbol{G}|^2} \end{aligned} \tag{2.57}$$

The corresponding element of the Hamiltonian matrix therefore becomes

$$\begin{aligned} &\langle \boldsymbol{k}+\boldsymbol{G}|V_{\mathrm{coul}}(\boldsymbol{r})|\boldsymbol{k}+\boldsymbol{G}'\rangle \\ &= \frac{1}{\Omega}\int_{\Omega} V_{\mathrm{coul}}(\boldsymbol{r}) \exp(-i\boldsymbol{G}\cdot\boldsymbol{r}) \exp(i\boldsymbol{G}'\cdot\boldsymbol{r}) d\boldsymbol{r} \\ &= V_{\mathrm{coul}}(\boldsymbol{G}-\boldsymbol{G}') \end{aligned} \tag{2.58}$$

In the same manner, the exchange–correlation term can be obtained as

$$\langle \boldsymbol{k}+\boldsymbol{G}|\mu_{\mathrm{XC}}(\boldsymbol{r})|\boldsymbol{k}+\boldsymbol{G}'\rangle = \mu_{\mathrm{XC}}(\boldsymbol{G}-\boldsymbol{G}') \tag{2.59}$$

Summarizing the above equations, we get

$$\begin{aligned} H_{k+G,k+G'} = &\frac{1}{2}|\boldsymbol{k}+\boldsymbol{G}|^2 \delta_{GG'} + V_{loc}^{pp}(\boldsymbol{G}-\boldsymbol{G}') \\ &+ V_{nloc}^{\mathrm{pp}}(\boldsymbol{k}+\boldsymbol{G},\boldsymbol{k}+\boldsymbol{G}') + V_{\mathrm{coul}}(\boldsymbol{G}-\boldsymbol{G}') \\ &+ \mu_{\mathrm{XC}}(\boldsymbol{G}-\boldsymbol{G}') \end{aligned} \tag{2.60}$$

2.2.6 Ultrasoft Pseudopotential Method

Softer pseudopotentials require smaller basis sets, leading to the improvement in calculation efficiency. The ultrasoft pseudopotential (USPP) scheme proposed by Vanderbilt [27] succeeded in making substantially soft potentials without sacrificing calculation accuracy by removing the norm-conserving condition and introducing augmentation charges. A scheme to include more than one reference energies in the construction of pseudopotential for better transferability was also suggested. This "multireference energy" scheme can be applied to the NCPP method in principle, but such an approach was employed in few studies.

Based on all-electron calculations, pseudo wave functions ($\tilde{\psi}_i$) are constructed so that the scattering properties are correct for reference energy levels. For a given angular momentum l, more than one energy levels, ε_i, are chosen. Pseudo wave functions are constructed under the generalized norm-conserving condition as follows.

$$Q_{ij} = \langle \psi_i | \psi_j \rangle_R - \left\langle \tilde{\psi}_i | \tilde{\psi}_j \right\rangle_R = 0 \tag{2.61}$$

Here $\langle \psi_i | \psi_j \rangle_R$ denotes the integral of $\psi_i^*(r)\psi_j(r)$ inside the sphere with a radius of R. Now, projector functions $|\beta_i\rangle = \sum_j (B^{-1})_{ji} \langle \chi_j |$ are defined, which are dual to the pseudofunctions $\left| \tilde{\psi}_i \right\rangle$ (i.e., $\left\langle \beta_i | \tilde{\psi}_j \right\rangle = \delta_{ij}$). Then, the nonlocal pseudopotential operator can be chosen as

$$V_{NL} = \sum_{i,j} B_{ij} |\beta_i\rangle \langle \beta_j| \tag{2.62}$$

The generalized norm-conserving condition $Q_{ij} = 0$ is not necessary if we accept the generalized eigenvalue problem where the overlapping operator

$$S = 1 + \sum_{i,j} Q_{ij} |\beta_i\rangle \langle \beta_j| \tag{2.63}$$

Then the nonlocal pseudopotential operator is redefined with $D_{ij} \equiv B_{ij} + \varepsilon_j Q_{ij}$ as

$$V_{NL} = \sum_{i,j} D_{ij} |\beta_i\rangle \langle \beta_j| \tag{2.64}$$

Here we get the following relation

$$\left\langle \tilde{\psi}_i | S | \tilde{\psi}_j \right\rangle_R = \langle \psi_i | \psi_j \rangle_R \tag{2.65}$$

Thus, the pseudofunctions are the solution of the generalized eigenvalue problem ($H = T + V_{\mathrm{loc}} + V_{\mathrm{NL}}$)

$$(H - \varepsilon_i S) \left| \tilde{\psi}_i \right\rangle = 0 \tag{2.66}$$

Though B is no longer Hermitian due to the removal of the generalized norm-conserving condition, the hermiticity of the pseudo-Hamiltonian is restored because D and Q are Hermitian operators. Valence charge density is written as

$$\rho(\boldsymbol{r}) = \sum_{i}^{\mathrm{occ}} \tilde{\psi}_i^*(r) \tilde{\psi}_i(\boldsymbol{r}) + \sum_{i,j} \rho_{ij} Q_{ji}(\boldsymbol{r}) \tag{2.67}$$

where

$$\rho_{ij} \equiv \sum_{k}^{\text{occ}} \langle \beta_i | \psi_k \rangle \langle \psi_k | \beta_j \rangle \tag{2.68}$$

$$Q_{ij}(r) \equiv \psi_i^*(r)\psi_i(r) - \tilde{\psi}_i^*(r)\tilde{\psi}_i(r) \tag{2.69}$$

The advantage of the removal of the generalized norm-conserving condition is that the pseudofunction construction should only meet the requirement that the pseudofunctions should only be smoothly connected to all-electron functions at a radius of R; therefore, the core radius can be much larger than that of NCPPs without sacrificing accuracy (accuracy can be retained by introducing the auxiliary function Q and the overlapping operator S) [28]. In fact, the logarithmic derivative of the wave function is conserved as

$$-\frac{1}{2}(r\tilde{\psi}_i(r))^2 \frac{d}{d\varepsilon_i}\frac{d}{dr}\ln \tilde{\psi}_i(r)\bigg|_R = \left\langle \tilde{\psi}_i | \tilde{\psi}_i \right\rangle_R + Q_{ii} = \langle \psi_i | \psi_i \rangle_R \tag{2.70}$$

$$-\frac{1}{2}(r\psi_i(r))^2 \frac{d}{d\varepsilon_i}\frac{d}{dr}\ln \psi_i(r)\bigg|_R = \langle \psi_i | \psi_i \rangle_R \tag{2.71}$$

2.2.7 *Projector-augmented Wave Method*

The projector-augmented wave (PAW) method proposed by Blöchl [29] is another solution to the problem that wave functions are sharp in the core region requiring a substantial number of plane waves. In PAW, a linear transformation operator, $\widehat{T}$, is introduced to efficiently describe wave function features that are largely different between core and interstitial regions.

The operator $\widehat{T}$ transforms an smooth auxiliary function, $|\psi_i\rangle$, to the true all-electron Kohn–Sham single-particle wave function, $|\psi_i\rangle$

$$|\psi_i\rangle = \widehat{T}\left|\tilde{\psi}_i\right\rangle \tag{2.72}$$

Then we get a transformed Kohn–Sham equation,

$$\widehat{T}^{\dagger} H \widehat{T} = \varepsilon_i \widehat{T}^{\dagger}\widehat{T}\left|\hat{\psi}_i\right\rangle \tag{2.73}$$

which is to be solved instead of the ordinary Kohn–Sham equation. The transformation operator is determined so that the auxiliary function as the solution of the above equation becomes smooth. $\widehat{T}$ has only to affect the core region, so we define

$$\widehat{T} = 1 + \sum_a \widehat{T}^a \tag{2.74}$$

where a is an atom index. $\widehat{T}^a$ is an atom-centered transformation, which has no effect outside augmentation spheres with a radius of r_c^a, i.e., $|\boldsymbol{r} - \boldsymbol{R}^a| > r_c^a$. Augmentation spheres do not overlap with each other. In the augmentation sphere, the true wave function is expanded to partial waves, ϕ_j^a. For each partial wave, a smooth auxiliary wave, $\tilde{\phi}_j^a$, is defined, and the following condition is required.

$$\left|\phi_j^a\right\rangle = \left(1 + \widehat{T}^a\right)\left|\tilde{\phi}_j^a\right\rangle \tag{2.75}$$

ϕ_j^a and $\tilde{\phi}_j^a$ coincide with each other outside the augmentation sphere. Now, the smooth wave function $\left|\tilde{\psi}_i\right\rangle$ is expanded with the smooth partial waves as

$$\left|\tilde{\psi}_i\right\rangle = \sum_j P_{ij}^a \left|\tilde{\phi}_j^a\right\rangle;\ |\boldsymbol{r} - \boldsymbol{R}^a| < r_c^a \tag{2.76}$$

Recalling $\left|\phi_j^a\right\rangle = \widehat{T}\left|\tilde{\phi}_j^a\right\rangle$, we can make the following expansion with the same coefficients,

$$|\psi_i\rangle = \widehat{T}\left|\tilde{\psi}_i\right\rangle = \sum_j P_{ij}^a \left|\phi_j^a\right\rangle;\ |\boldsymbol{r} - \boldsymbol{R}^a| < r_c^a \tag{2.77}$$

The linearity of the transformation operator gives

$$P_{ij}^a = \left\langle \tilde{p}_j^a | \tilde{\psi}_i \right\rangle \tag{2.78}$$

$\widetilde{P}_j^a$ is called a smooth projector function.

The projector function must satisfy the following completeness relation,

$$\sum_j \left|\tilde{\phi}_j^a\right\rangle \left\langle \tilde{p}_j^a \right| = 1;\ |\boldsymbol{r} - \boldsymbol{R}^a| < r_c^a \tag{2.79}$$

which implies that the projector function must be orthonormal to the smooth partial waves within the augmentation sphere ($\left\langle p_j^a | \tilde{\phi}_k^a \right\rangle = \delta_{jk}; |\boldsymbol{r} - \boldsymbol{R}^a| < r_c^a$). Using the completeness relation, we get

$$\widehat{T}^a = \sum_j \widehat{T}^a \left|\tilde{\phi}_j^a\right\rangle \left\langle \tilde{p}_j^a\right| = \sum_j \left(\left|\phi_j^a\right\rangle - \left|\tilde{\phi}_j^a\right\rangle\right) \left\langle \tilde{p}_j^a\right| \tag{2.80}$$

and therefore

$$\hat{T} = 1 + \sum_a \sum_j \left(\left|\phi_j^a\right\rangle - \left\langle \tilde{\phi}_j^a\right|\right) \left\langle \tilde{p}_j^a\right| \tag{2.81}$$

The all-electron Kohn–Sham wave function can be obtained as

$$\psi_i(\boldsymbol{r}) = \tilde{\psi}_i(\boldsymbol{r}) + \sum_a \sum_j \left(\phi_j^a(\boldsymbol{r}) - \tilde{\phi}_j^a(\boldsymbol{r})\right) \left\langle \tilde{p}_j^a | \hat{\psi}_i\right\rangle \tag{2.82}$$

By this decomposition, the original wave function is divided to the auxiliary wave function (smooth everywhere) and the contribution including fast oscillation (affecting limited region in space). It is the advantage of the introduction of the transformation operator to be able to treat the two functions independently.

2.2.8 All-Electron Method

Besides the pseudopotential approach, there exist calculation methods to obtain the states of all electrons (core and valence electrons). One of the methods that have been widely used for all-electron (including both valence and core electrons) calculation has its origin in the APW (augmented plane wave) method proposed by Slater (1937) [30]. In APW, wave functions are represented by plane waves for interstitial region that are connected to wave functions of (atomic) core region. The potential of the core region is approximated with muffin-tin-type functions. LAPW (linearized APW) developed by Anderson (1975) [31] enabled efficient calculation by linearizing radial wave functions with respect to energy, with which the calculation can be solved as a generalized eigenvalue problem. To lift the restriction of spherical wave functions due to the muffin-tin approximation, it was further developed to FLAPW (full-potential LAPW) [32]. WIEN2k [33] is a well-known simulation package for calculations based on FLAPW.

There is also a lineage of the KKR method [34], which was originally proposed by Koringa (1947), Kohn, and Rostoker (1954). This is also called the Green function method, as the method defines the one-particle Green function of the Kohn–Sham equation. Calculation of wave functions and eigenvalues is circumvented, and charge density or local density of states can be obtained directly through the Green function. It has been confirmed by many studies that results obtained by KKR and APW exhibit good agreement with each other.

2.2.9 Beyond LDA and GGA

While DFT calculations with LDA and GGA have enjoyed much success, it is also known that there are a number of problems that cannot be addressed with these approximations. Attempts to eliminate these problems, which are often called "Beyond LDA and GGA," include the $+U$ method (often called DFT + U, LDA + U or GGA + U as well) [35], the GW approximation [36], and DFT–HF hybrid functionals [16]. In general, these advanced methods require much increased computational resources or introduction of additional, adjustable parameters.

It has been known that the "standard" DFT calculation with LDA or GGA can fail dramatically in materials containing electrons with strong correlation, whose ground state is characterized by pronounced localization of electrons. A typical and serious problem is the substantial underestimation of the band gap energy. This is because the approximate exchange–correlation functionals do not cancel out the electronic self-interaction, which makes charge density portions associated with one atom repel each other, in the Hartree (Coulomb) term, resulting in over-delocalized valence electrons. The $+U$ approach is based on the Hubbard model, which considers the interaction between electrons within the same atom (on-site Coulomb interaction), and corrects the self-interaction error in the standard DFT that tends to overly delocalize the metal d and f states. In the $+U$ approach, the strength of the on-site interactions is described by parameters U and J, which represent on-site Coulomb and on-site exchange contributions, respectively. These parameters can be evaluated by ab initio calculations, but usually determined empirically (i.e., treated as adjustable parameters) to reproduce experimental results. One of the main advantages of this method is that computational cost is not much different from the standard DFT calculation.

If we denote the mth energy level in the Kohn–Sham equation of a N-electron system as $\epsilon_M^{(N)}$, the actual (experimental) energy gap is $E = \epsilon_{N+1}^{(N+1)} - \epsilon_N^{(N)}$. In a standard DFT calculation, however, the gap is obtained as $E_{\mathrm{DFT}} = \epsilon_{N+1}^{(N)} - \epsilon_N^{(N)}$, resulting in the underestimation by $\Delta = \epsilon_{N+1}^{(N+1)} - \epsilon_{N+1}^{(N)}$. The GW approximation is a way to correct this discrepancy based on the quantum theory of many-body system. In the approximation, Δ is written as the product between the one-particle Green's function (G) and the screened Coulomb interaction (W), $\Delta = iGW$, where G and W are obtained using the HF approximation. It has been shown that GW can remarkably improve the evaluation of the band gap energy for the cases where LDA works to some extent (e.g., the band gap is underestimated but not completely closed). A major disadvantage of GW is the requirement of tremendous increase in computational cost.

As explained in Sect. 2.1.2, the exchange interaction is correctly implemented in the HF method. One way to circumvent the problem of inaccurate representation of the exchange energy in the standard DFT is therefore to incorporate a portion of exact exchange from the HF method. The exact exchange functional in HF is written as

$$E_X^{\mathrm{HF}} = -\frac{1}{2}\sum_{i,j}\int\int \psi_i^*(\boldsymbol{r}_1)\psi_j^*(\boldsymbol{r}_1)\frac{1}{r_{12}}\psi_i(\boldsymbol{r}_2)\psi_j(\boldsymbol{r}_2)d\boldsymbol{r}_1 d\boldsymbol{r}_2 \tag{2.83}$$

where $\psi(\boldsymbol{r})$ is one-electron Bloch states of the system. The most popular B3LYP functional (standing for "Becke, three-parameter, Lee-Yang-Parr") writes the hybrid exchange–correlation functional as $E_{\mathrm{XC}}^{\mathrm{B3LYP}} = E_X^{\mathrm{LDA}} + a_0\left(E_X^{\mathrm{HF}} - E_X^{\mathrm{LDA}}\right) + a_X\left(E_X^{\mathrm{GGA}} - E_X^{\mathrm{LDA}}\right) + E_C^{\mathrm{LDA}} + a_C\left(E_C^{\mathrm{GGA}} - E_C^{\mathrm{LDA}}\right)$, where a_0, a_X and a_C are parameters. The DFT–HF hybrid functional method also requires substantial increase in computational effort.

2.2.10 Evaluation of Physical Quantities

The total energy of the system, E_{tot}, is expressed as

$$\begin{aligned} E_{\mathrm{tot}} = &\sum_k \sum_n^{\mathrm{occ}} \varepsilon_{kn} - \frac{1}{2}\int V_{\mathrm{coul}}(r)\rho(r)dr \\ &+ \int[\varepsilon_{\mathrm{xc}}(r) - \mu_{\mathrm{xc}}(r)]\rho(r)dr + E_{\mathrm{Ewald}} \end{aligned} \tag{2.84}$$

where E_{Ewald} is the Ewald summation representing interaction between nuclei (ions). In the NCPP formalism, we get

$$\begin{aligned} E_{\mathrm{tot}} = &\frac{1}{2}\sum_k f_k \sum_n^{\mathrm{occ}} f_n \sum_G |k+G|^2 \left|C_{k+G}^n\right|^2 + \Omega_a \sum_G V_{\mathrm{loc}}^{\mathrm{pp}}(G)\rho(-G) \\ &+ \sum_k f_k \sum_n^{\mathrm{occ}} f_n \sum_G \sum_{G'} C_{k+G}^{n*} C_{k+G'}^n V_{\mathrm{nloc}}^{\mathrm{pp}}(k+G, k+G') \\ &+ \frac{1}{2}\Omega_a \sum_G V_{\mathrm{coul}}(G)\rho(-G) \\ &+ \Omega_a \sum_G \varepsilon_{\mathrm{xc}}(G)\rho(-G) + E_{\mathrm{Ewald}} \end{aligned} \tag{2.85}$$

Considering the fact that the diverging terms $V_{\mathrm{loc}}^{\mathrm{pp}}(G)\big|_{G=0}$ and $V_{\mathrm{coul}}(G)\big|_{G=0}$ offset with the diverging term in E_{Ewald}, we obtain

$$\begin{aligned} E_{\mathrm{tot}} = & \frac{1}{2}\sum_k f_k \sum_n^{\mathrm{occ}} f_n \sum_G |k+G|^2 \left|C_{k+G}^n\right|^2 + \Omega_a \sum_{G\neq 0} V_{loc}^{pp}(G)\rho(-G) \\ & + \sum_k f_k \sum_n^{\mathrm{occ}} f_n \sum_G \sum_{G'} C_{k+G}^{n*} C_{k+G'}^n V_{nloc}^{pp}(k+G, k+G') \\ & + \frac{1}{2}\Omega_a \sum_{G\neq 0} V_{\mathrm{coul}}(G)\rho(-G) \\ & + \Omega_a \sum_G \varepsilon_{\mathrm{xc}}(G)\rho(-G) + E'_{\mathrm{Ewald}} + \sum_a \frac{\alpha_a Z_a}{\Omega_a} \end{aligned} \tag{2.86}$$

Here, Z_a is the number of valence electrons of each atom and α_0 is given by

$$\alpha_a = \int_{\Omega_a} \left(V_a^{\mathrm{pp,loc}}(r) - \left(-\frac{Z_a}{r}\right)\right) d\boldsymbol{r} = 4\pi \int_0^\infty r^2 \left(V_a^{\mathrm{pp,loc}}(r) + \frac{Z_a}{r}\right) dr \tag{2.87}$$

E'_{Ewald} is the Ewald sum subtracted by the diverging term, given as

$$\begin{aligned} E'_{\mathrm{Ewald}} = & \sum_a \sum_{a'} Z_a Z_{a'} \sum_{G\neq 0} \frac{2\pi}{\Omega_a |\boldsymbol{G}|^2} \exp[-i\boldsymbol{G}\cdot(\boldsymbol{r}_a - \boldsymbol{r}_{a'})] \exp\left(-\frac{|\boldsymbol{G}|^2}{4\gamma^2}\right) \\ & + \frac{1}{2}\sum_a \sum_{a'} Z_a Z_{a'} \sum_R \frac{\mathrm{erfc}(|\boldsymbol{r}+\boldsymbol{r}_{a'}-\boldsymbol{r}_a|\gamma)}{|\boldsymbol{r}+\boldsymbol{r}_{a'}-\boldsymbol{r}_a|} - \sum_a \frac{Z_a^2\gamma}{\sqrt{\pi}} - \frac{Z^2\pi}{2\Omega_a\gamma^2} \end{aligned} \tag{2.88}$$

where γ is a parameter such that the series expansion converges fast and $Z = \sum_a Z_a$.

The force exerted on atom a, $\boldsymbol{F}_a$, is given as the derivative of the total energy with respect to $\boldsymbol{r}_a$,

$$\begin{aligned} \boldsymbol{F}_a = -\frac{\partial E_{\mathrm{tot}}}{\partial \boldsymbol{r}_a} = & -\frac{1}{\Omega_a}\sum_k f_k \sum_n^{\mathrm{occ}} f_n \sum_G \sum_{G'} C_{k+G}^{n^*} C_{k+G'}^n i(\boldsymbol{G}'-\boldsymbol{G}) \exp[-i(\boldsymbol{G}-\boldsymbol{G}') \\ & \cdot \boldsymbol{r}_a] \times \left[V_a^{\mathrm{pp,loc}}(\boldsymbol{G}-\boldsymbol{G}') + V_a^{\mathrm{pp,nloc}}(\boldsymbol{k}+\boldsymbol{G}, \boldsymbol{k}+\boldsymbol{G}')\right] - \frac{\partial E'_{\mathrm{Ewald}}}{\partial \boldsymbol{r}_a} \end{aligned} \tag{2.89}$$

The last term can be written as

$$\begin{aligned}\frac{\partial E'_{\text{Ewald}}}{\partial \boldsymbol{r}_a} &= -\sum_{a'} Z_a Z_{a'} \frac{4\pi}{\Omega_a} \sum\nolimits_{G\neq 0} \frac{\boldsymbol{G}}{|\boldsymbol{G}|^2} \sin\{\boldsymbol{G}\cdot(\boldsymbol{r}_a - \boldsymbol{r}_{a'})\} \exp\left(-\frac{|\boldsymbol{G}|^2}{4\gamma^2}\right) \\ &+ \sum_{a'} Z_a Z_{a'} \sum\nolimits_{\boldsymbol{R}} \frac{\boldsymbol{R}+\boldsymbol{r}_{a'}-\boldsymbol{r}_a}{|\boldsymbol{R}+\boldsymbol{r}_{a'}-\boldsymbol{r}_a|^3} \\ &\times \left\{ \operatorname{erfc}(|\boldsymbol{R}+\boldsymbol{r}_{a'}-\boldsymbol{r}_a|\gamma) \right. \\ &\left. -|\boldsymbol{R}+\boldsymbol{r}_{a'}-\boldsymbol{r}_a|\gamma \frac{\partial \operatorname{erfc}(|\boldsymbol{R}+\boldsymbol{r}_{a'}-\boldsymbol{r}_a|\gamma)}{\partial(|\boldsymbol{R}+\boldsymbol{r}_{a'}-\boldsymbol{r}_a|\gamma)} \right\}\end{aligned} \tag{2.90}$$

The local term can be rewritten as below for faster calculation,

$$\begin{aligned}&-\frac{1}{\Omega_{at}} \sum_k f_k \sum_n^{\text{occ}} f_n \sum_G \sum_{G'} C^{n^*}_{k+G} C^n_{k+G'} i(\boldsymbol{G}'-\boldsymbol{G}) \exp[-i(\boldsymbol{G}-\boldsymbol{G}')\cdot\boldsymbol{r}_a] V_a^{\text{pp,loc}}(\boldsymbol{G}-\boldsymbol{G}') \\ &= -\frac{1}{\Omega_{at}} \sum_k f_k \sum_n^{\text{occ}} f_n \sum_G \sum_{G'} C^{n^*}_{k+G} C^n_{k+G'} i(-\boldsymbol{G}') \exp[-i\boldsymbol{G}' \\ &\cdot\boldsymbol{r}_a] V_a^{\text{pp,loc}}(\boldsymbol{G}') = \sum_{\boldsymbol{G}} \rho(-\boldsymbol{G}) i\boldsymbol{G} \exp(-i\boldsymbol{G}\cdot\boldsymbol{r}_a) V_a^{\text{pp,loc}}(\boldsymbol{G})\end{aligned} \tag{2.91}$$

The global stress exerted on the simulation box (supercell), $\sigma_{\alpha\beta}(\alpha,\beta = x,y,z)$, is calculated as the derivative of the total energy with respect to strain. Using $S_a(G) = \exp(-iG\cdot r_a)$, the global stress is expressed as

$$\begin{aligned}\sigma_{\alpha\beta} &= \frac{1}{\Omega_a}\frac{\partial E_{\text{tot}}}{\partial \varepsilon_{\alpha\beta}} = -\frac{1}{\Omega_a}\sum_k f_k \sum_n^{\text{occ}} f_n \sum_G \left|C^n_{k+G}\right|^2 (k+G)_\alpha (k+G)_\beta \\ &-\frac{1}{\Omega_a}\sum_{G\neq 0}\sum_a S_a(G)\left[\frac{\partial V_a^{\text{pp,loc}}(G)}{\partial(G^2)} 2G_\alpha G_\beta \right. \\ &\left. + V_a^{\text{pp,loc}}(G)\delta_{\alpha\beta}\right]\rho(-G) \\ &+ \sum_k f_k \sum_n^{\text{occ}} f_n \sum_G \sum_{G'} \sum_l \sum_a S_a(G \\ &-G') C^{n^*}_{k+G} C^n_{k+G'} \frac{\partial}{\partial \varepsilon_{\alpha\beta}} \left[\frac{1}{\Omega_a} V_{a,l}^{\text{pp,nloc}}(k+G,k+G')\right] \\ &+ \frac{1}{2}\sum_{G\neq 0} V_{\text{coul}}(G)\rho(-G)\left(\frac{2G_\alpha G_\beta}{G^2} - \delta_{\alpha\beta}\right) + \delta_{\alpha\beta}\sum_G [\varepsilon_{\text{xc}}(G) - \mu_{\text{xc}}(G)]\rho(-G) \\ &+ \frac{1}{\Omega_a}\frac{\partial E_{\text{Ewald}}}{\partial \varepsilon_{\alpha\beta}} - \delta_{\alpha\beta}\frac{Z}{\Omega_a}\sum_a \alpha_a\end{aligned} \tag{2.92}$$

The Ewald term becomes

$$\begin{aligned}\frac{\partial E_{\mathrm{Ewald}}}{\partial \varepsilon_{\alpha\beta}} = {} & \frac{2\pi}{\Omega_a}\sum_{G\neq 0}\frac{1}{G^2}\exp\left(-\frac{G^2}{4\gamma^2}\right)\left|\sum_a Z_a \exp(iG\cdot r_a)\right|^2 \\ & \times\left[\frac{2G_\alpha G_\beta}{G^2}\left(\frac{G^2}{4\gamma^2}+1\right)-\delta_{\alpha\beta}\right] \\ & +\frac{1}{2}\gamma\sum_a\sum_{a'}\sum_R Z_a Z_{a'} H'(D\gamma)\frac{D_\alpha D_\beta}{D^2}\bigg|_{D=R+r_{a'}-r_a} + \frac{Z^2\pi}{2\Omega_a\gamma^2}\delta_{\alpha\beta}\end{aligned} \tag{2.93}$$

where $H'(x) = \frac{\partial \mathrm{erfc}(x)}{\partial x} - \frac{\mathrm{erfc}(x)}{x}$.

A practical scheme to evaluate local energy and local stress was established by Shiihara et al. [37] within the framework of the stress density developed by Filippetti and Fiorentini [38]. With the method, one can evaluate the distribution of energy and stress in a system containing nonuniform structure, such as surfaces and grain boundaries. Energy density ($e_{\mathrm{tot}}(\boldsymbol{r})$) and stress density ($\tau_{\alpha\beta}(\boldsymbol{r})$) are defined as integrands of macroscopic total energy and stress tensor, respectively, as

$$E_{\mathrm{tot}} = \int_V e_{\mathrm{tot}}(\boldsymbol{r})d\boldsymbol{r} \tag{2.94}$$

$$\sigma_{\alpha\beta} = \frac{1}{V}\frac{\partial E_{\mathrm{tot}}}{\partial \varepsilon_{\alpha\beta}} = \frac{1}{V}\int_V \tau_{\alpha\beta}(\boldsymbol{r})d\boldsymbol{r} \tag{2.95}$$

where V is the total volume of the supercell and $\varepsilon_{\alpha\beta}$ is strain tensor. Therefore, local energy and local stress for partial region indicated by i can be defined, respectively, as

$$E_{\mathrm{tot}}(i) = \int_{V_i} e_{\mathrm{tot}}(\boldsymbol{r})d\boldsymbol{r} \tag{2.96}$$

$$\sigma_{\alpha\beta}(i) = \frac{1}{V_i}\int_{V_i} \tau_{\alpha\beta}(\boldsymbol{r})d\boldsymbol{r} \tag{2.97}$$

where V_i indicates partial volume. The local values are not well defined because the expressions can contain functions that are gauge-dependent, which integrates to zero over V but does not over V_i. The gauge dependency stems from the fact that local energy density can be defined in symmetric and asymmetric expressions. The former is

$$e_{\mathrm{kin},S}(\boldsymbol{r}) = \frac{1}{2}\sum_i f_i \nabla\psi_i^*(\boldsymbol{r})\cdot\nabla\psi_i(\boldsymbol{r}) \tag{2.98}$$

and the latter is

$$e_{\mathrm{kin},AS}(\boldsymbol{r}) = -\frac{1}{2}\sum_i f_i \psi_i^*(\boldsymbol{r})\nabla^2\psi_i(\boldsymbol{r}) \tag{2.99}$$

where ψ_i is a valence wave function and f_i is an occupation number. If we take partial region i such that the symmetric and asymmetric expressions coincide, local energy can be described in a well-defined form. It was shown that the differences between the symmetric and asymmetric expressions for local energy and local stress are proportional to $\nabla^2\rho$ and $\nabla_\alpha\nabla_\beta\rho$, respectively. Thus, the conditions of gauge dependence for e_{tot} and $\tau_{\alpha\beta}$ are

$$\int_{V_i} \nabla^2\rho(\boldsymbol{r})d\boldsymbol{r} = 0 \tag{2.100}$$

and

$$\int_{V_i} \nabla_\alpha\nabla_\beta\rho(\boldsymbol{r})d\boldsymbol{r} = 0 \tag{2.101}$$

respectively. As practical ways to divide a supercell to meet the above conditions, Shiihara et al. established a layer-by-layer and Bader integral methods [37].

2.3 Semi-empirical and Empirical Theories for Nanostructure Properties

2.3.1 Semi-empirical Calculation of Electronic State

One of the major disadvantages in the first-principles electronic state calculation is that the calculation requires tremendous computer resources, which severely limits the size of simulation objects. It is already challenging to handle a system consisting of thousands of atoms with a laboratory-class cluster server. When a relatively large simulation cell is required to investigate, e.g., properties of materials with defects, it can be a reasonable way to choose a method of semi-empirical electronic structure calculation, where the Schrödinger equation is solved but with empirically constructed functions and parameters.

The tight-binding (TB) method is a most widely used MO method for semi-empirical electronic structure calculation [39, 40]. The method is based on the assumption that electrons are strongly bound to atoms which they belong to so that they cannot move to other orbitals. In that sense, the method is opposite to free-electron models. Hopping of electron states between different orbitals is, however, allowed to some extent because orbitals are slightly overlapped. It should

be noted here that different definitions of the TB method seem to exist. Nevertheless, the TB method is in most cases considered to be equivalent to the extended Hückel method, where electron–electron interaction is neglected but overlapping (hopping) integral between different orbitals is considered. In usual TB calculations, the self-consistent loop calculation is not performed.

We write the atomic orbital α of the ath atom positioned at $\boldsymbol{r}_a$ as $\phi_{a\alpha}(\boldsymbol{r}-\boldsymbol{r}_a)$. The wave function Ψ^i of electron state i is written as

$$\Psi^i = \sum_{a,\alpha} C^i_{a\alpha}\phi_{a\alpha}(\boldsymbol{r}-\boldsymbol{r}_a) \tag{2.102}$$

The Hamiltonian is written with the potential from atom a, V_a, as

$$H = -\frac{1}{2}\nabla^2 + \sum_a V_a(\boldsymbol{r}-\boldsymbol{r}_a) \tag{2.103}$$

The Hamiltonian matrix element therefore becomes

$$\begin{aligned} H_{a\alpha b\beta} &= \int \phi_{a\alpha}(\boldsymbol{r}-\boldsymbol{r}_a) H \phi_{b\beta}(\boldsymbol{r}-\boldsymbol{r}_b)d\boldsymbol{r} \\ &= \int \phi_{a\alpha}(\boldsymbol{r}-\boldsymbol{r}_a)\left\{-\frac{1}{2}\nabla^2 + \sum_k V_k(\boldsymbol{r}-\boldsymbol{r}_k)\right\}\phi_{b\beta}(\boldsymbol{r}-\boldsymbol{r}_b)d\boldsymbol{r} \\ &= \int \phi_{a\alpha}(\boldsymbol{r}-\boldsymbol{r}_a)\left\{-\frac{1}{2}\nabla^2 + V_a(\boldsymbol{r}-\boldsymbol{r}_a) + V_b(\boldsymbol{r}-\boldsymbol{r}_b)\right\}\phi_{b\beta}(\boldsymbol{r}-\boldsymbol{r}_b)d\boldsymbol{r} \\ &\quad + \int \phi_{a\alpha}(\boldsymbol{r}-\boldsymbol{r}_a)\left\{\sum_{k\neq a,b} V_k(\boldsymbol{r}-\boldsymbol{r}_k)\right\}\phi_{b\beta}(\boldsymbol{r}-\boldsymbol{r}_b)d\boldsymbol{r} \end{aligned} \tag{2.104}$$

The last term (the three-center integral) is neglected in conventional TB calculations, which is called the two-center approximation. The Hamiltonian matrix elements (also called the resonance integral) under the two-center approximation are calculated by the direction cosine of atom pairs and parameters, which is shown by Slater and Koster (Slater–Koster table, see Table 2.2) [40]. The overlap integral is defined as

$$S_{a\alpha b\beta} = \int \phi_{a\alpha}(\boldsymbol{r}-\boldsymbol{r}_a)\phi_{b\beta}(\boldsymbol{r}-\boldsymbol{r}_b)d\boldsymbol{r} \tag{2.105}$$

and then the generalized eigenvalue problem

Table 2.2 Slater–Koster table (only part). l, m, n are direction cosine from atom a to atom b

$H_{as,bs}$	$= V_{s,s}$	$= V_{ss\sigma}$
H_{as,bp_x}	$= V_{s,x}$	$= lV_{sp\sigma}$
H_{ap_x,bp_x}	$= V_{x,x}$	$= l^2V_{pp\sigma} + (1-l^2)V_{pp\pi}$
H_{ap_x,bp_y}	$= V_{x,y}$	$= lm(V_{pp\sigma} - V_{pp\pi})$
H_{ap_x,bp_z}	$= V_{x,z}$	$= ln(V_{pp\sigma} - V_{pp\pi})$

$$\boldsymbol{HC} = \boldsymbol{ESC} \tag{2.106}$$

is solved to obtain eigenvalues (energy levels) and eigenfunctions (wave functions).

In a system with periodic boundaries, basis functions are constructed by taking the Bloch sum for a $\boldsymbol{k}$-point as follows.

$$\psi_{a\alpha,\boldsymbol{k}}(\boldsymbol{r}) = \frac{1}{\sqrt{N}} \sum_{l} \exp[i\boldsymbol{k} \cdot (\boldsymbol{r}_a + \boldsymbol{R}_l)] \psi_{a\alpha}(\boldsymbol{r} - \boldsymbol{r}_a - \boldsymbol{R}_l) \tag{2.107}$$

Here, $\boldsymbol{r}_a$ is the position vector of atom a in the fundamental cell, l the index for cells, and $\boldsymbol{R}_l$ is the translation vector for cell l. N is the number of periodically arranged cells (N is infinity but will be canceled out later). An individual Hamiltonian for each $\boldsymbol{k}$-point is constructed, and its eigenvalue problem is to be solved. Diagonal and nondiagonal terms are written as

$$\begin{aligned} H^{\boldsymbol{k}}_{a\alpha,a\alpha} &= \int \psi^*_{a\alpha,\boldsymbol{k}}(\boldsymbol{r}) H \psi_{a\alpha,\boldsymbol{k}}(\boldsymbol{r}) d\boldsymbol{r} \\ &= \sum_{l} \exp[i\boldsymbol{k} \cdot \boldsymbol{R}_l] \int \psi^*_{a\alpha,\boldsymbol{k}}(\boldsymbol{r} - \boldsymbol{r}_i) H \psi_{a\alpha,\boldsymbol{k}}(\boldsymbol{r} - \boldsymbol{r}_i - \boldsymbol{R}_l) d\boldsymbol{r} \end{aligned} \tag{2.108}$$

and

$$\begin{aligned} H^{\boldsymbol{k}}_{a\alpha,b\beta} &= \int \psi^*_{a\alpha,\boldsymbol{k}}(\boldsymbol{r}) H \psi_{b\beta,\boldsymbol{k}}(\boldsymbol{r}) d\boldsymbol{r} \\ &= \sum_{l} \exp\left[\left(i\boldsymbol{k} \cdot (\boldsymbol{r}_j + \boldsymbol{R}_l - \boldsymbol{r}_i)\right)\right] \\ &\quad \int \psi^*_{a\alpha,\boldsymbol{k}}(\boldsymbol{r} - \boldsymbol{r}_i) H \psi_{b\beta,\boldsymbol{k}}\left(\boldsymbol{r} - \boldsymbol{r}_j - \boldsymbol{R}_l\right) d\boldsymbol{r} \end{aligned} \tag{2.109}$$

respectively.

The total energy of a system, E_{tot}, is given as the sum of the band energy, E_{TB}, and the repulsive energy, E_{rep},

$$E_{\text{tot}} = E_{TB} + E_{\text{rep}} \tag{2.110}$$

Here, the band energy is the sum of the energy eigenvalues of occupied states,

$$E_{TB} = 2 \sum_{i}^{\text{occ}} E^i \tag{2.111}$$

The repulsive energy is often given as the form of simple pairwise functions.

As an example, a set of TB parameters (function forms for $\boldsymbol{H}$ and $\boldsymbol{S}$) for silicon atoms by Kohyama [41] is presented below.

$$E_{\text{rep}} = \frac{1}{2}\sum_{j\neq i}\sum_{i}\varphi\left(r_{ij}\right) \tag{2.112}$$

$$\varphi\left(r_{ij}\right) = A_{ij}S\left(r_{ij}\right)r_{ij}^{-v} \tag{2.113}$$

$$V_{ll'm} = \eta_{ll'm}S\left(r_{ij}\right)r_{ij}^{-v} \tag{2.114}$$

$$S\left(r_{ij}\right) = \left\{1 + \exp\left[\mu\left(r_{ij} - R_c\right)\right]\right\}^{-1} \tag{2.115}$$

where r_{ij} is the separation between atoms i and j, and $\eta, v,$ and μ and parameters. A_{ij} is defined as

$$A_{ij} = b_0 - b_1\left(Z_i + Z_j\right) \tag{2.116}$$

Z_i is the effective coordination number of atom i

$$Z_i = \sum_{j\neq i}\exp\left[-\lambda_1\left(r_{ij} - R_i\right)^2\right] \tag{2.117}$$

$$R_i = \sum_{j\neq i} r_{ij}\exp\left(-\lambda_2 r_{ij}\right)\left[\sum_{j\neq i}\exp\left(-\lambda_2 r_{ij}\right)\right]^{-1} \tag{2.118}$$

$R_c, b_0, b_1, \lambda_1,$ and λ_2 are parameters. In addition, the diagonal elements of the Hamiltonian (on-site terms) are constants given as parameters.

2.3.2 *Atomistic Modeling Using Empirical Interatomic Potential*

To obtain electronic structure and evaluate related physical properties, it is basically needed to solve the Schrödinger equation. However, some relatively sophisticated interatomic models include terms that represent the structure of electrons, e.g., charge density, so that the models can evaluate physical properties determined by the electron structure, such as magnetism and ferroelectricity. These approaches may be a solution of problems that require a very large number of atoms and therefore cannot be addressed by electron structure calculations.

The shell model [42] is a crude model to mimic charge polarization around atoms by pairs of cation and anion particles (a cation–anion pair corresponds to an atom and surrounding charge). In this way, one can simulate changes in atom positions and charge polarization by optimizing the structure of the particles

according to the environment. It has been demonstrated that the model works for perovskites to reproduce ferroelectricity.

The dipole potential model by Tangney and Scandolo [43] is originated in the same idea to represent charge polarization around atoms but in a slightly different way. The TS model incorporates electrostatic dipole vectors assigned to atoms. Because of the similarity in the fundamental concepts, there seems to be no significant difference in the two models except technical matters in computation.

An interatomic model was developed by Dudarev and Derlet [44] to represent the effect of magnetism in Fe. The model, which describes potential energy of atoms within the framework of the embedded atom method (EAM) [45], was constructed so that the effect of paramagnetic–ferromagnetic transition on potential energy difference is represented by employing an additional term in the embedding energy function.

Though abovementioned empirical atomistic models have enjoyed successful representation of the objective properties they are designed for, such approaches are available only for a limited variety of properties. For example, electrostatic calculations are necessary for the evaluation of the band gap energy, even for qualitative analysis.

2.4 Conclusion

This chapter gave an overview of methods for computational analysis of solid materials. To evaluate physical properties originated in electron states, it is necessary to perform electron structure calculations, which usually require solving the Schrödinger equation. Currently, ab initio DFT seems to be the method of choice for solid materials, especially for problems of multiphysics because with the approach one can evaluate various properties, both mechanical and physical, with an excellent accuracy. In fact, a growing number of researches are being conducted using the approach not only due to the reliability of the theory par se but also due to software packages available on the market or for free, which still keep being developed with incorporating new methods that improve the prediction accuracy and the computational efficiency. Owing to seemingly everlasting advance of computational power, it will presumably keep getting easier to deal with larger models (simulation cell containing a large number of atoms) that are necessary to address problems of complex structures.

It is also indispensable, however, to employ other methods that are not ab initio-based but computationally efficient when necessary. In a theoretical approach with numerical simulations, a typical pitfall is producing artifacts due to the limitation of cell size, i.e., if the property in question has a strong size effect, setting up a sufficiently large simulation cell should be prioritized than conducting rigorous electronic structure calculations. One should always be aware of the theoretical

background of the computational method and its drawbacks as there is no versatile method. It should also be noted that computational methods for electronic structure calculations are making a rapid progress. Present challenges referred to in this chapter may be addressed in the near future.

Appendix: First-Principles and Ab Initio

The term "ab initio" is often used as a synonym of "first-principles," as is found in many scientific papers. Rigorously speaking, however, they have different meanings. Ab initio, meaning "from the beginning" in Latin, is calculation of electron states using no empirical parameters where the usage of only fundamental physical constants, such as Planck constant, electron mass, elementary charge, are allowed. The term "ab initio" should be used when the calculation is wave-function-based theories as opposed to DFT, which is density-based. Thus, it is strange to say "ab initio DFT" although it is not rare to find such combination of the terms in scientific reports. To be exact, when a DFT calculation is done without empirical parameters, it should be called "first-principles DFT." However, in this book, we do not stick to the slight difference of the meaning between "ab initio" and "first-principles," and we accept the use of "ab initio" for DFT calculations.

References

1. P. Hohenberg, W. Kohn, Phys. Rev. **136**, B864 (1964)
2. M. Born, J.R. Oppenheimer, Ann. Physik **84**, 457 (1927)
3. V. Fock, Z. Phys. **61**, 126 (1930)
4. W. Kohn, L.J. Sham, Phys. Rev. **140**, A1133 (1965)
5. R.P. Feynman, Phys. Rev. **56**, 340 (1939)
6. G. Kresse and J. Furthmüller, Phys. Rev. B 54, 11169 (1996). (www.vasp.at)
7. www.castep.org
8. www.abinit.org
9. www.quantum-espresso.org
10. O. Gunnarsson, M. Jonson, B.I. Lundqvist, Phys. Rev. B **20**, 3136 (1979)
11. J.M. MacLaren, D.P. Clougherty, M.E. McHenry, M.M. Donovan, Comput. Phys. Comm. **66**, 383 (1991)
12. R. Gaspar, Acta Phys. Hungaria **3**, 263 (1954)
13. J.P. Perdew, A. Zunger, Phys. Rev. B **23**, 5048 (1981)
14. D.M. Ceperley, B.J. Alder, Phys. Rev. Lett. **45**, 566 (1980)
15. U. von Barth, L. Hedin, J. Phys. C: Solid State Phys. **5**, 1629 (1972)
16. A.J. Cohen, P. Mori-Sánchez, W. Yang, Chem. Rev. **112**, 289 (2012)
17. J.P. Perdew, K. Burke, Int. J. Quant. Chem. **57**, 309 (1996)
18. J.P. Perdew, K. Burke, M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996)
19. J.P. Perdew, Y. Wang, Phys. Rev. B **45**, 13244 (1992)
20. D.R. Hamann, M. Schlüter, C. Chiang, Phys. Rev. Lett. **43**, 1494 (1979)
21. G.B. Bachelet, D.R. Hamann, M. Schlüter, Phys. Rev. B **26**, 4199 (1982)

22. D.R. Hamann, Phys. Rev. B **40**, 2980 (1989)
23. G.P. Kerker, J. Phys. C: Solid St. Phys. **13**, L189 (1980)
24. N. Troullier, J.L. Martins, Phys. Rev. B **43**, 1993 (1991)
25. M. Rappe, K.M. Rabe, E. Kaxiras, J.D. Joannopoulos, Phys. Rev. B **41**, 1227 (1990)
26. S. Ogata, Ph.D Thesis, Osaka University (1998)
27. D. Vanderbilt, Phys. Rev. B **41**, 7892 (1990)
28. C. Lee, J. Korean Phys. Soc. **31**, S278 (1997)
29. P.E. Blöechl, Phys. Rev. B **50**, 17953 (1994)
30. J.C. Slater, Phys. Rev. **51**, 846 (1937)
31. O.K. Andersen, Phys. Rev. B **12**, 3060 (1975)
32. P. Blaha, K. Schwarz, P. Sorantin, Comput. Phys. Commun. **59**, 399 (1990)
33. www.wien2k.at
34. W. Kohn, N. Rostoker, Phys. Rev. **94**, 1111 (1954)
35. B. Himmetoglu, A. Floris, S. de Gironcoli, M. Cococcioni, Int. J. Quantum Chem. **114**, 14 (2014)
36. M. Marsili, S. Botti, M. Palummo, E. Degoli, O. Pulci, H.-C. Weissker, M.A.L. Marques, S. Ossicini, R. Del Sole, J. Phys. Chem. C **117**(27), 14229 (2013)
37. Y. Shiihara, M. Kohyama, S. Ishibashi, Phys. Rev. B **81**, 075441 (2010)
38. A. Filippetti, V. Fiorentini, Phys. Rev. B **61**, 8433 (2000)
39. D.A. Papaconstantopoulos, M.J. Mehl, J. Phys Condens. Matter **15**, R413 (2003)
40. J.C. Slater, G.F. Koster, Phys. Rev. **94**, 1498 (1954)
41. M. Kohyama, R. Yamamoto, Phys. Rev. B **49**, 17102 (1994)
42. B.J. Dick, A.W. Overhauser, Phys. Rev. **112**, 90 (1958)
43. P. Tangney, S. Scandolo, J. Chem. Phys. **117**, 8898 (2002)
44. S.L. Dudarev, P.M. Derlet, J. Phys, Condens. Matter **17**, 7097 (2005)
45. M.S. Daw, M.I. Baskes, Phys. Rev. B **29**, 6443 (1984)

Chapter 3
Ideal Strength in Low-Dimensional Nanostructures

Abstract The ideal strength is the theoretical strength of materials setting the limit of stress attained and is therefore an important fundamental quantity describing mechanical behavior of materials. While a substantial number of studies have been dedicated to the ideal strength of various types of crystals, the ideal strength of nanostructures can be peculiar due to strong influence of low-dimensional structures on the mechanical properties. We introduce recent studies of the ideal strength of nanostructures, mainly discussing that of two-dimensional (e.g., films) and one-dimensional (e.g., nanowires and nanotubes) structures.

Keywords Ideal strength · Low-dimensional nanostructure · Mechanical property

3.1 Mechanical Properties of Nanostructures

Mechanical properties of materials are represented by various quantities, among which ideal strength (also referred to as theoretical strength) and elastic constants (describing elastic response at equilibrium) are most fundamental indicators of mechanical traits. As both the properties are influenced by structures, it is important to understand the ideal strength and elastic constants of nanostructures. In this chapter, we mainly focus on the former, while the latter is partially covered with some examples.

3.1.1 Ideal Strength

Ideal strength is defined as the fracture strength of a perfect (pristine) crystal under homogeneous stress or strain. Ideal strength is equivalent to the stress at which unstable deformation (structural instability) is initiated in the crystal, i.e., deformation is sustained under a constant or decreasing load. Analysis of ideal strength is, therefore, usually to evaluate the mechanical stability of a perfect crystal at 0 K. This

Y. Umeno et al., *Multiphysics in Nanostructures*, Nanostructure Science and Technology, DOI 10.1007/978-4-431-56573-4_3

is achieved by observing the sign of the eigenvalues of the Hessian matrix (H), whose elements are the second derivatives of the potential energy of the system with respect to strain. In the case where the crystal has a large degrees of freedom in terms of deformation, namely internal strains, the Hessian matrix is written in such a way that the degrees of freedom are included (i.e., the dimension of the Hessian matrix becomes large because of the large number of degrees of freedom, with respect to which the potential energy is differentiated). The evaluation of crystal stability at finite temperatures can be done by replacing potential energy with free energy.

The well-known Born criterion [1], which is often referred to as elastic stability condition, corresponds to the case where six lattice parameters of a unit crystal cell, three side lengths, and three angles between them, are considered. In this case, the Hessian matrix becomes the elastic coefficient matrix C with the dimension of 6×6. From the condition for the matrix C to be positive definite (i.e., all eigenvalues are positive), the condition for crystal stability is derived. For example, the condition for the stability of a cubic crystal is $C_{11} + 2C_{12} > 0$, $C_{44} > 0$, and $C_{11} - C_{12} > 0$, meaning that the system becomes unstable unless bulk modulus ($B = (C_{11} + 2C_{12})/3$), shear modulus ($G = C_{44}$), and tetragonal shear modulus ($\mathrm{G}' = (C_{11} - C_{12})/2$) are all positive. These three conditions are called spinodal instability, shear instability, and Born instability, respectively, which correspond to different modes of unstable deformation.

Another well-known condition for crystal stability $\omega^2(\boldsymbol{q}, s) > 0$, which is often referred to as dynamical stability condition, is derived by considering $3N$ atom displacements as mechanical degrees of freedom. In this case, the Hessian matrix becomes the dynamical matrix $D(\boldsymbol{q})$ with the dimension of $3N \times 3N$. Here, ω is phonon frequency, $\boldsymbol{q}$ is wave vector, s labels whether the wave is longitudinal or transverse, and N is the number of atoms. The eigenvalues of the dynamical matrix are $\omega^2(\boldsymbol{q}, s)$, and the eigenvectors represent vibration modes. The condition for the dynamical matrix to be positive definite is therefore written as $\omega^2(\boldsymbol{q}, s) > 0$.

As seen above, both elastic and dynamical stability conditions are derived from the positive definiteness of the Hessian matrix. The former is the expression of the stability criterion in terms of C_{ij} given that the elastic coefficients are the second derivatives of free (or potential) energy with respect to strain tensor. This approach can be expanded to the case of bulk containing defects or complex structures if the elastic coefficients of the system are properly defined. On the other hand, the latter requires that the system has complete periodicity. The two stability conditions, which are explained in more detail in a review paper [2], therefore correspond to macroscopic and microscopic stabilities, respectively.

In the last couple of decades, ideal strength analyses by means of first-principles calculations have been intensively performed [2, 3]. In general, a unit cell is used as simulation cells in first-principles calculations, where the periodic boundary conditions are applied in the x, y and z directions to realize a perfect crystal. The initial configuration (equilibrium state under vanishing mechanical load) is obtained by structural relaxation while adjusting cell dimensions so that all stress components, σ_{ij}, become zero. Then, increasing stress or strain is applied while structural

relaxation is performed. Ideal strength is obtained by evaluating the mechanical stability of the system during the process of increasing stress or strain. To analyze elastic and dynamical stabilities, elastic coefficients C_{ijkl} and force constants $K_{\alpha i \beta j}$ are calculated at each stress or strain and the elastic coefficient matrix C and dynamical matrix $D(\boldsymbol{q})$ are evaluated, where i, j, k and l are labels of the Cartesian axes, and α and β are those of atoms. The evaluation of the elastic coefficients with first-principles calculations is in general done by the finite difference method (i.e., $C_{ijkl} \approx \Delta\sigma_{ij}/\Delta\varepsilon_{kl}$ where $\Delta\sigma_{ij}$ is change of stress as a response to infinitesimal strain $\Delta\varepsilon_{kl}$), or by the density functional perturbation theory (DFPT) for analytical evaluation of the derivatives. Similarly, the force constants are evaluated by $K_{\alpha i \beta j} \approx \Delta f_{\alpha i}/\Delta r_{\beta j}$ where $\Delta f_{\alpha i}$ is change of force as a response to infinitesimal atom displacement $\Delta r_{\beta j}$, or by DFPT. A detailed explanation about DFPT can be found in, for example, Ref. [4].

In the case of cleavage fracture, rupture of interatomic bonds by tension occurs at the fracture, which is a typical form of fracture in brittle materials. In metals, on the other hand, bifurcation of deformation paths may arise during tension. In such a case, the stress at the bifurcation is the ideal strength. As an example, we introduce here a study of the Cu perfect crystal undergoing uniaxial tension along the [100] direction [5]. Note that the x, y and z directions are set along [100], [010], and [001], respectively. Figure 3.1a shows the relationship between tensile stress (σ_{xx}) and tensile strain (ε_{xx}). The tensile stress increases with increasing strain until the strain shows its maximum value of 9.0 GPa at $\varepsilon_{xx} = 0.098$, followed by decrease. The decrease in the stress occurs due to a deformation bifurcation corresponding to the Born instability ($C_{22} - C_{23} < 0$). As the tensile strain grows, the value of $C_{22} - C_{23}$ decreases and becomes negative when ε_{xx} exceeds 0.098 (Fig. 3.1b), which is associated with the structure change from fcc to bcc (Fig. 3.2). That is, $\varepsilon_{xx} = 0.098$ is the point when the structural instability occurs and the corresponding stress, $\sigma_{xx} = 9.0$ GPa, is the ideal strength of Cu under the uniaxial tension along [100].

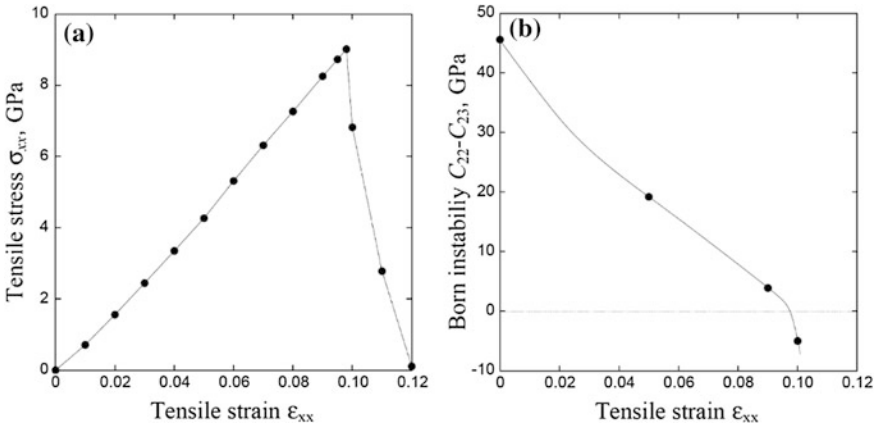

Fig. 3.1 Cu single crystal in [100] uniaxial tension [5]. **a** Tensile stress σ_{xx} and **b** Born instability C_{22}–C_{23}, as a function of tensile strain ε_{xx}

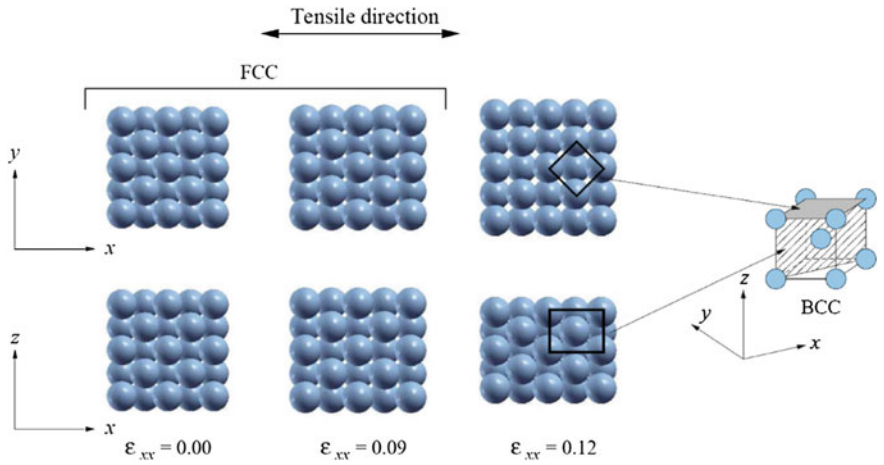

Fig. 3.2 Change in crystal structure of Cu single crystal in [100] uniaxial tension [5]

Ideal strength and structural instability modes of a material alter depending on loading conditions—crystal orientations and the combination of six stress components. For example, the ideal strengths of Al under uniaxial tension along ⟨100⟩, ⟨110⟩, and ⟨111⟩ evaluated by first-principles calculations are 9.20, 4.89, and 8.95 GPa, respectively [6]. In the analysis, the dynamical instability occurs prior to the elastic instability during tension in any direction. A first-principles calculation study of Ni under uniaxial ⟨100⟩ and hydrostatic tension/compression [7] demonstrates the case where the combination of stress components alters instability modes. While Ni under uniaxial tension along ⟨100⟩ exhibits a deformation bifurcation by the Born instability from isotropic to anisotropic Poisson's contractions, the crystal causes a phase transition by the spinodal instability under uniaxial ⟨100⟩ compression and hydrostatic tension. As relatively simple combinations of stress components, first-principles calculations have been done for normal stress along i (σ_i) combined with isotropic transverse stress (σ_b), e.g., $\{\sigma\} = \{\sigma_b, \sigma_b, \sigma_z, 0, 0, 0\}$. These calculations have reported that the ideal strength obtained as the maximum value of σ_i depends on the transverse stress. Besides, for the case of shear stress on plane i along j (τ_{ij}) combined with normal stress along i and isotropic transverse stress, e.g., $\{\sigma\} = \{\sigma_b, \sigma_b, \sigma_z, 0, \tau_{zx}, 0\}$, the ideal shear strength obtained as the maximum value of τ_{ij} depends on σ_i and σ_b. The ideal strength of various crystals for different crystal orientations and combinations of stress components is reviewed, for example, in Ref. [3].

In the case of fcc and bcc perfect crystals, the unit cell contains only one atom and the crystals therefore exhibit uniform deformation having no internal strain. Thus, it is valid to evaluate the mechanical stability of the system in terms of the elastic or dynamical stability conditions. On the other hand, the mechanical stability under inhomogeneous deformation must be evaluated in the case of crystals with complex structures. This is also the case of materials with structures that will be

discussed from the next section. An extended method to analyze the mechanical stability that works for inhomogeneous deformation was proposed by Kitamura et al. [8, 9]. The method considers the Hessian matrix ($\boldsymbol{H}$) constructed taking into account all mechanical degrees of freedom (i.e., atom displacements) that influences the potential energy of the system in question and evaluates the mechanical stability by the positive definiteness of the Hessian matrix. Although this method has been applied thus far only to calculations using empirical interatomic potentials, it can be expanded in principle to first-principles calculations. Such an extended method is needed for a rigorous analysis of ideal strength beyond the restriction of homogeneous deformation.

In Sects. 3.2 and 3.3, we will discuss the strength of materials having characteristic (ideal) understructures (Sect. 3.2) and shapes (Sect. 3.3). The strength of understructures may be referred to as "fundamental structural strength" and that of ideal shapes as "ideal structural strength." The ideal strength introduced in this section represents the essential strength of materials and serves as the bedrock for the discussion of the fundamental and ideal structural strengths.

3.1.2 Elastic Constants

Elastic constants (or elastic moduli) are factors that represent the slope of linear stress–strain relationship in the region of small strain. It is known that the elastic response can be altered by structures; e.g., the elastic property is influenced by surfaces. Thus, the elastic response of nanostructures differs from that of bulk.

While the stress–strain relationship (or its slope) of a nanostructured material depends on the size (e.g., surface–volume ratio in a thin film or nanowire), the effect of structures on elasticity can be evaluated by defining elastic constants of structures such as surfaces. As elastic constants are derivative of stress components with respect to strain, our problem here is to consider the definition of stress of structures. For example, the definition of surface and interface stress is explained in Ref. [10]. In the Lagrangian coordinate system denoted by subscript L, surface area and surface free energy are represented as follows:

$$A_L = A/(1+\varepsilon_{ii}), \tag{3.1}$$

$$\gamma_L = \gamma(1+\varepsilon_{ii}), \tag{3.2}$$

respectively, where A and γ are physical values. ε_{ii} is the trace of the elastic strain, ε_{ij}. Surface stress is then given as

$$f_{ij} = \frac{\partial \gamma_L}{\partial \varepsilon_{ij}}. \tag{3.3}$$

Note that the dimension of the surface stress is the same as that of surface energy, i.e., different from that of bulk stress. Similar to the definition of bulk elastic constants, surface elastic constants can be defined as the derivative of the surface stress with respect to strain. Therefore, elastic constants of structures such as surfaces and interfaces are well-defined values. Elastic constants of lower dimensional structures can be defined exactly in the same manner.

The material strength can be remarkably influenced by the strength of understructures or shapes because the strength of a material is governed by the weakest part of the material. In contrast, elastic constants of understructures or shapes do not significantly matter because elastic constants are the averaged quantity of the entire structure. The issue of elastic constants of nanostructures will be briefly discussed later in this chapter (Sect. 3.3.1).

3.2 Ideal Understructure

We define the ideal understructure as the structure where one characteristic inhomogeneity is included in a perfect crystal and discuss the strength of materials having ideal understructures, i.e., the fundamental structural strength. In such cases, the fundamental structural strength is the stress at which a material with an ideal understructure becomes mechanically unstable under stress or strain that is homogeneous sufficiently away from the location of its inhomogeneous structure. By comparing the fundamental structural strength with the ideal strength, the effect of the internal inhomogeneity of the structure on strength and instability modes can be clarified. In this section, examples of characteristic internal inhomogeneity of zero-, one-, and two-dimensional structures are presented to discuss their fundamental structural strengths.

3.2.1 Zero-Dimensional Understructure

The vacancy is a typical internal inhomogeneous structure of zero-dimension. A structure containing a vacancy in a perfect crystal is therefore regarded as a zero-dimensional ideal understructure. Although a simulation cell with an infinite size is necessary to rigorously realize such a structure in first-principles calculations, a sufficiently large simulation cell is used in practice. The cell dimension must be over a critical value (L_c) where the energy per atom is sufficiently converged; that is, the energy change should be negligibly small with respect to change in the cell size. Here, a physical quantity in question must be used for judging the convergence. For the analysis of the mechanical instability, the convergence should be examined in terms of the Hessian matrix (or the second derivatives of potential energy in practice). It may be worth noting here that, in general, the first derivatives of energy (e.g., force and stress) and the second derivatives (e.g., elastic coefficients

and phonon dispersion) are more sensitive than the energy to the change in the cell size, cutoff energy, k-point mesh, etc. It should also be noted that, in first-principles calculations for materials with a vacancy, what part of the material should be focused for local stress or strain is unclear. Because there arises inhomogeneous distribution of stress and strain around a vacancy, it is desirable to evaluate local stress and strain for the determination of critical stress and strain for the onset of unstable deformation, i.e., fundamental structural strength. A method for first-principles calculations of local energy and stress can be found in Ref. [11].

As a first-principles analysis of the fundamental structural strength of a material with vacancies, W crystal containing a vacancy undergoing uniaxial tension along $\langle 100 \rangle$ was studied [12]. The ideal strength of W was evaluated to be 27.3 GPa and the fundamental structural strength of W with a vacancy to be 26.8 GPa, where the tensile strength was evaluated by the inflexion point of the energy–strain curve. Here, a caveat should be noted for the quantitative comparison between the values. The calculations used a simulation model of a $3 \times 3 \times 3$ stuck of bcc cubic unit cells, where it was not examined whether the model is sufficiently large. Besides, the tensile stress was evaluated in terms of the global stress of the model, not local stress around the vacancy, and therefore, the fundamental structural strength was defined as the maximum stress of the stress–strain curve. Nevertheless, a rigorous evaluation of the onset of unstable deformation should be made by analyzing the mechanical stability of the system under inhomogeneous deformation, which may find an instability before the peak of the stress–strain curve. To sum up, one should beware that the fundamental structural strength obtained in the study has room of discussion in terms of the simulation cell size, the definition of stress, and the mechanical stability. Lack of mechanical instability analysis left instability modes unclarified. Having said that observation of the change in atom arrangements during tension reveals that interatomic bond lengths are shorter in the vicinity of the vacancy than those at a distance and that the bonds away from the vacancy break simultaneously at the peak of the stress–strain curve. This result indicates that a cleavage fracture in the region away from the vacancy, where stress distribution is homogeneous, occurs as the unstable deformation mode. Although this is the only study of the fundamental structural strength of a material with a vacancy reported thus far, the ideal strength and instability modes of a material with a vacancy are naturally expected to depend on the atom species (electronic structure), crystal structure, crystal orientation, and combination of stress components, as is the case for perfect crystals.

3.2.2 One-Dimensional Understructure

The dislocation, which is a line defect defined as the boundary between a region where a crystal slip has occurred and that where no slip exists, is a one-dimensional internal inhomogeneous structure. Dislocations are categorized into two types: edge and screw. The edge (screw) dislocation is the configuration where the dislocation

line and the Burgers vector, or the displacement of the slip deformation, are normal (parallel) to each other. Dislocations found in real materials are often mixed and consist of edge and screw components. Thus, a configuration of an edge or screw dislocation embedded in a perfect crystal can be regarded as one of the most fundamental forms of one-dimensional ideal understructures. To evaluate its fundamental structural strength with first-principles calculations, it is needed to properly set the simulation cell size and clarify the definition of stress (or strain), similarly to the case of vacancy models. According to the linear elasticity theory, the stress field around a dislocation decays with r^{-1} regardless of the type of dislocation. Such a long-range stress field requires a substantially large simulation cell compared to the case dealing with force fields that decay with r^{-2} or faster. More details about the dislocation theory can be found, for example, in Ref. [13].

There has been no study of the fundamental structural strength of materials with a dislocation by first-principles calculations. A conceivable unstable deformation mode is the motion of a dislocation. The motion of a dislocation occurs when the resolved shear stress along the Burgers vector, τ_{rss}, reaches the critical stress for dislocation motion, τ_p (Peierls stress). Peierls stress has never been directly obtained by first-principles calculations. Instead, Peierls stress can be estimated based on the Peierls–Nabbaro model, which is a continuum theory for dislocation motion, in the combination with general stacking-fault energy of perfect crystals obtained by first-principles calculations [14]. In the case of uniaxial tension with a stress of σ, the resolved shear stress becomes $\tau_{\mathrm{rss}} = \sigma \cos\phi \cos\lambda$, and the tensile stress that meets $\tau_{\mathrm{rss}} = \tau_p$ is the fundamental structural strength. Here, ϕ is the angle between the tensile axis and the normal direction of the slip plane, and λ is the angle between the tensile axis and the slip direction. As the resolved shear stress varies with the crystal orientation and the combination of stress components, the fundamental structural strength of a system with a dislocation depends on them. A system becomes unstable with instability modes that is not for dislocation motion when the resolved shear stress is vanishing (e.g., $\phi = 90^\circ$ or $\lambda = 90^\circ$ in uniaxial tension) or the dislocation is sessile (e.g., the Lomer-Cottrell lock). One of such modes is for fracture from the dislocation and another is for fracture in the homogeneous region far away from the dislocation. Fracture should occur from the dislocation if interatomic bonds around the dislocation are weaker than those in the homogeneous region, and vice versa.

The grain boundary, which is a boundary between crystals with different orientations, is a two-dimensional internal inhomogeneous structure. A configuration of a grain boundary embedded in a perfect crystal can be regarded as one of two-dimensional ideal understructures. There exist numerous grain boundary structures depending on the combination of crystal orientations. Here, we overview the classification and geometric models for grain boundaries, as more details can be found, for example, in Ref. [4]. Consider that the crystal orientation on one side of the grain boundary is expressed by rotation of that on the other side. The grain boundary is called a tilt grain boundary when the rotation axis lies on the grain boundary; a twist grain boundary, when the rotation axis is perpendicular to the

grain boundary. Grain boundaries are divided into small- and large-angle grain boundaries according to the rotation angle; symmetric and asymmetric grain boundaries, according to the symmetry between the adjoining crystal configurations. If the crystal lattice points of one side expanded to the other side find coinciding lattice points arranged periodically, it is called a coincidence grain boundary and denoted by the Σ value and the misorientation angle Θ. In general, grain boundaries possess asymmetric and complex structures with tilt and twist boundaries mixed. Attention must be paid again for setting a proper simulation model size and clarifying the definition of stress or strain when analyzing the fundamental structural strength of a grain boundary using first-principles calculations.

3.2.3 Two-Dimensional Understructure

First-principles calculations of the fundamental structural strength of materials containing grain boundaries have been performed only for coincidence grain boundaries. Since coincidence grain boundaries possess periodic and ordered structures, calculations can be done with relatively small simulation cells with the periodic boundary conditions. In the following, we introduce a study of coincidence tilt and twist grain boundaries in Al [15]. The study calculated the fundamental structural strength of the $\Sigma 3(111)[111]60^\circ$ twist, $\Sigma 3(11\bar{2})[110]109^\circ$ tilt, $\Sigma 11(1\bar{1}3)[110]129^\circ$ tilt, and $\Sigma 9(1\bar{1}4)[110]141^\circ$ tilt grain boundaries under tension normal to the brain boundary, where (hkl) and $[uvw]$ indicate the grain boundary plane and rotation axis, respectively. While the ideal strengths of Al under uniaxial tensions along $\langle 111\rangle$, $\langle 112\rangle$, $\langle 113\rangle$, and $\langle 114\rangle$ are 12.6, 12.8, 12.7, and 13.2 GPa, respectively, the fundamental structural strengths of the above-mentioned grain boundaries are 11.5, 6.8, 12.0, and 12.1 GPa, respectively. The results, obtained as the global stress of the entire simulation cell, indicate the reduction in strength by the presence of a grain boundary. The tensile simulation was done by rigidly separating the crystals on the opposite sides of the grain boundary, considering the occurrence of brittle fracture at the grain boundary. On the other hand, another analysis was done considering grain boundary sliding as the instability mode [16]. The study evaluated the fundamental structural strength of the above four and $\Sigma 9(2\bar{2}\bar{1})[110]39^\circ$ brain boundaries under shear along the grain boundary plane. Depending on the grain boundary structure and the combination of stress components, other instability modes such as grain boundary migration and dislocation emission from the grain boundary may occur.

When more than one inhomogeneous structures exist in a material, characteristic organized microstructures such as dislocation assemblies may emerge. A microstructure can also be regarded as an ideal understructure because the microstructure is one type of internal inhomogeneous structure. It is well known that two- or three-dimensional dislocation assemblies, such as dislocation veins,

walls, and cells, are formed due to interactions between dislocations, dislocation dipoles, and quadrupoles. As the size of each structure of dislocation assembly is of the order of μm, it is currently extremely difficult to simulate such structures by atomistic modeling simulations using empirical potentials, let alone first-principles calculations. On the other hand, a simulation scheme based on the discrete dislocation dynamics method, which is a well-known numerical simulation method for dislocations, has been proposed to analyze the mechanical stability of dislocation assemblies [17]. The basic concept of this scheme, which evaluates the mechanical stability by the positive definiteness of a Hessian matrix that is constructed with respect to the degrees of freedom of all dislocations, is the same as the approach by Kitamura et al. [8, 9]. The critical shear stress and corresponding instability modes of ideal vein and wall structures with periodic arrangement of dislocation quadrupoles were evaluated, demonstrating the capability of the scheme for examining the condition and deformation modes of instabilities of dislocation microstructures.

The instability modes of dislocation microstructures depend on the type of dislocation assembly and the combination of stress components. While all dislocations in the microstructure may start moving simultaneously, only part of dislocations may move. Instead of dislocation motion, instability modes of fracture from dislocations or in homogeneous region away from dislocations may emerge. The evaluation of such instability modes cannot be done by schemes based on discrete dislocation dynamics but requires atomistic modeling. Besides the dislocation microstructure, the grain boundary network is a characteristic microstructure. While no first-principles calculation is found for the study of a grain boundary network structure, analyses using empirical interatomic potentials have been performed [18], revealing that triple junctions of grain boundaries contribute in a complicated fashion to dislocation nucleation, dislocation emission, grain boundary sliding, grain boundary migration, crack initiation, and crack propagation. If a periodic grain boundary network with high symmetry is formed (e.g., two-dimensional squares or three-dimensional cubes arranged periodically), an instability mode that corresponds to macroscopic deformation transforming the shape of the grain boundary network may emerge.

3.2.4 Understructure of Two or More Elements

While the previous subsections were confined to unary materials, we here discuss characteristic inhomogeneity of understructures in systems consisting of more than one atom species. For example, the substitutional or interstitial impurity and the interface between dissimilar materials are zero- and two-dimensional internal inhomogeneous structures, respectively. As was discussed already, attention should be paid to proper setting of simulation cell sizes and appropriate definition of stress or strain when evaluating the fundamental structural strength of materials containing such structures by first-principles calculations. Note that, in the following examples of first-principles calculations for systems containing impurities or

interfaces, global stress and strain for the entire simulation cell are used, meaning that the fundamental structural strength is evaluated as the maximum value of the global stress.

The hydrogen atom is a prototypical interstitial impurity that causes hydrogen embrittlement. A first-principles study was done for W with interstitial hydrogens undergoing ⟨100⟩ uniaxial tension, finding that the fundamental structural strength of the system is 7% lower than that of the perfect W crystal [19]. The reduction in the strength is due to the weakening of W–W bonds caused by charge transfer from W to H, suggesting the occurrence of fracture from the weak W–W bonds. First-principles calculations for ⟨100⟩ uniaxial tension of W containing substitutional impurities of Re, Ta, and V were also performed [20], showing that their fundamental structural strengths are only 1–2% smaller than the strength of the perfect crystal. These results indicate that, similarly to the case of the perfect crystal, fracture occurs from W–W bonds, not from W–Re, W–Ta, or W–V bonds.

As for systems with an interface between dissimilar materials, first-principles calculations were performed for a Ag–Al multilayer system containing a coherent Ag(111)/Al(111) interface under [111] tension and $[11\bar{2}]$ shear [21]. Here, atomic layers near the interface are denoted as $Ag_{-3}Ag_{-2}Ag_{-1}/Al_1Al_2Al_3\cdots$, where the subscript *i* indicates the *i*th layer from the interface. While Al_1–Al_2 bonds break in the [111] tension, Ag_{-2}–Ag_{-1} bonds break in the $[11\bar{2}]$ shear, demonstrating that an interface can exhibit fractures from different parts according to the loading condition. Note that the fracture does not occur from the Ag/Al interface layer but from the adjoining layers, which is due to the formation of strong Ag_{-1}–Al_1 bonds by charge transfer resulting in a strong interface strength. A study of the coherent Al (111)/TiN(111) interface under [111] tension [22] serves as an example of the cases where interface strength is highly influenced by the combination of atom species that form the interface. The study showed that the strength of the interface differs much depending on the termination of the TiN layer at the interface; that is, the strength of the Al/N interface is three times as large as that of the Al/T interface.

It is naturally conjectured that, besides the loading condition and the atom species, the form of the interface, such as coherent, semi-coherent, and noncoherent, should influence the fundamental structural strength and instability modes of the interface. Thus, further extensive studies are required to fully understand the fundamental structural strength and instability modes of systems containing characteristic internal inhomogeneity of structure.

3.3 Nanostructures with Ideal Shape

We define the ideal shape as the structure having no defects but surfaces. The ideal shape is thus constructed by removing periodicity in certain direction(s) from a perfect crystal. For example, a thin film is an ideal shape with a two-dimensional periodicity, while a wire or a tube is one with a one-dimensional periodicity. Similar to the case of ideal understructures discussed in Sect. 3.3, the strength of a

material with a shape can be regarded as the ideal structural strength. The ideal structural strength in this case is the stress at which the structure with a low-dimensional periodicity exhibits unstable deformation under uniform stress or strain. The smaller the size of the low-dimensional periodic structure is, the larger the ratio of its surface area to the volume is. The effect of surfaces on ideal structural strength, therefore, becomes significant when the size of the structure goes down to the nanometer scale. In this section, we discuss the ideal structural strength of low-dimensional periodic nanostructures presenting examples of nanofilms, nanowires, and nanotubes.

3.3.1 Two-Dimensional Nanostructures

Characteristic features of the two-dimensional periodic nanostructure are the presence of periodic surface(s) and a large surface-to-volume ratio. It is therefore primarily important to understand the surface structure so as to discuss the ideal structural strength of two-dimensional periodic nanostructures. It is well known that, depending on the surface orientation and atom species, a surface is reconstructed to exhibit an atomic arrangement different from that in bulk. For example, Si surfaces are reconstructed by dangling bonds forming new bonds to reduce the system energy [23, 24]. Figure 3.3 shows examples of Si(001) surface reconstruction (Fig. 3.3a for the unreconstructed surface and Fig. 3.3b–e for four patterns of reconstruction). White and black circles in the figures denote atoms on the first (surface) layer and the second layer, respectively, and the larger circles represent the atoms displaced further in the normal direction to the surface than the smaller circles. The unit cell of the surface structure is marked in blue. The surface structure in Fig. 3.3a is called $p(1 \times 1)$, where each atom in the first surface layer has two bonds with other atoms and two dangling bonds. The surface structures in Fig. 3.3b–e emerge by the formation of dimers from two adjoining dangling bonds. Figure 3.3b shows the surface structure called $p(2 \times 1)$ symmetric, where two atoms forming a dimer have the same height. When the heights of the two dimer atoms are different, various surface structures emerge depending on the dimer arrangement as depicted Fig. 3.3c–e, which illustrate the $p(2 \times 1)$ asymmetric, $p(2 \times 2)$, and $c(4 \times 2)$ structures, respectively. Energy analysis by first-principles calculations has shown that the $c(4 \times 2)$ structure is most favored, followed by $p(2 \times 2)$, $p(2 \times 1)$ asymmetric, $p(2 \times 1)$ symmetric, and $p(1 \times 1)$ in this order [23]. According to the analysis, the energy differences between $p(1 \times 1)$ and $p(2 \times 1)$ symmetric, $p(2 \times 1)$ symmetric and $p(2 \times 1)$ asymmetric, and $p(2 \times 1)$ asymmetric and $c(4 \times 2)$ are 1.8 ± 0.1, 0.12 ± 0.01, 0.048 ± 0.018, and 0.003 ± 0.013 eV/dimer, respectively, showing that the $p(2 \times 1)$ asymmetric, $p(2 \times 2)$ and $c(4 \times 2)$ structures are much more stable than $p(2 \times 1)$ symmetric and $p(1 \times 1)$. Since the energy difference is small, surface structures in reality shift among $p(2 \times 1)$ asymmetric, $p(2 \times 2)$, and $c(4 \times 2)$ at finite temperatures due to temporal and spatial fluctuations of dimers.

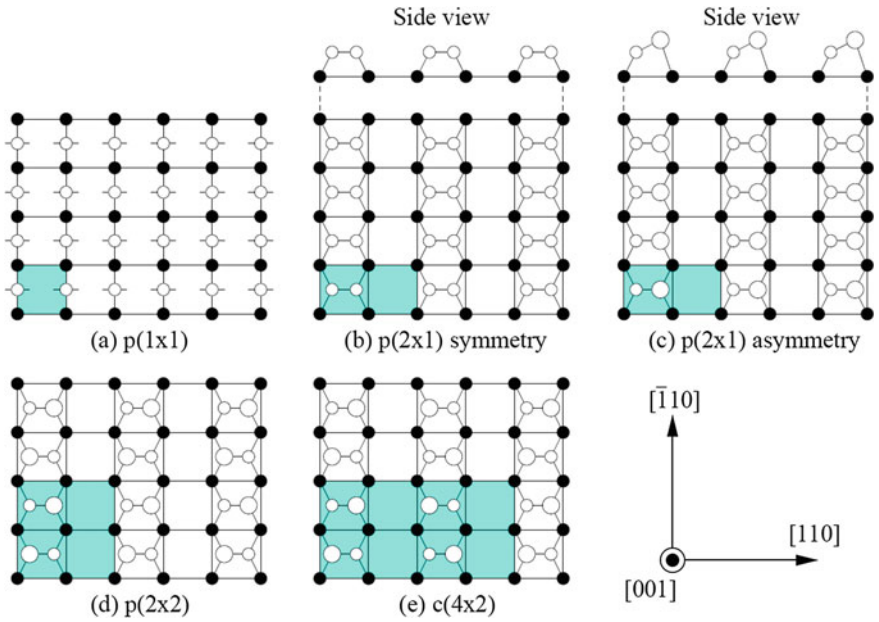

Fig. 3.3 Si(001) surface structures: **a** unreconstructed, **b**–**e** reconstructed. *White* and *black circles* are the first and second layer atoms, respectively. The *blue-shaded* areas indicate the surface unit cells [5]

As an example of first-principles calculations for the ideal structural strength of two-dimensional periodic nanostructures, we here introduce a tensile deformation simulation of Si(001) thin films [5, 25]. In what follows, the "(*hkl*) film" denotes a thin film with surface orientations of (*hkl*). Figure 3.4 schematically shows the Si (001) thin film. The dashed lines indicate the simulation cell used in the calculations, which is a slab model with a thin film sandwiched by vacuum layers. The periodic boundary conditions are applied to the three directions to realize a thin

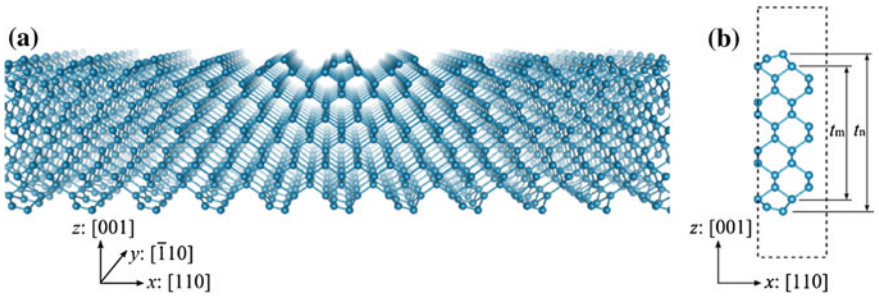

Fig. 3.4 **a** Schematic and **b** unit cell of Si(001) nanofilm with 14 layers and $p(2 \times 1)$ asymmetric surfaces [5]

film. The thin-film model in Fig. 3.4 contains 14 atomic layers and possesses the $p(2 \times 1)$ asymmetric surface structures. The x-, y-, and z-axes are along [110], [$\bar{1}$10], and [001] orientations, respectively, and the dimers are arranged parallel to [110]. To discuss the ideal structural strength of thin films, it is important to clarify the definition of stress, as was in Sect. 3.3. Because of the lack of internal inhomogeneous structure, it is valid to adopt the usage of global stress for the entire thin film. However, caveats should be noted for the calculation of the stress. In the conventional first-principles calculations, stress is obtained as the global one for the entire simulation cells. Thus, the stress actually attained by the film must be calculated by subtracting the vacuum layers in the simulation cell. For example, the normal stress along the x direction attained by a thin film is calculated by $\sigma_x = \sigma_x^{\text{cell}} \times L_z^{\text{cell}}/t$ where σ_x^{cell} is the global normal stress along the x direction for the entire simulation cell, L_z^{cell} is the cell dimension perpendicular to the film surfaces (cell size in the z direction), and t is the thickness of the film. The following example of a Si(001) thin film under [110] uniaxial tension demonstrates that special care is needed for the definition of the film thickness. Figure 3.5 shows the stress–strain relationships of 6-, 10-, and 14-layer films and bulk: (a) using the thickness t_m excluding the dimer layers and (b) using the thickness t_n including the dimer layers. While the Young's modulus of the film (the slope of the $\sigma_x - \varepsilon_x$ curve at $\varepsilon_x = 0$) is close to that of the bulk in Fig. 3.5a, the modulus becomes smaller as the film thickness decreases in Fig. 3.5b. In addition, the amount of reduction in the maximum stress by decreasing in the thickness is different between Fig. 3.5a, b. As is presented above, for the discussion of the mechanical property of such low-dimensional structures one should realize that the definition of the dimension of the structure makes a significant effect on the stress–strain relationship.

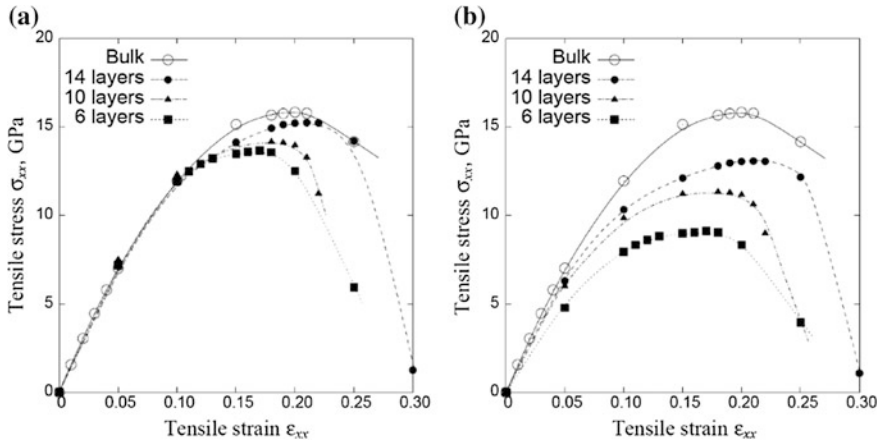

Fig. 3.5 Tensile stress–strain curves of Si bulk and (001) nanofilms in [110] uniaxial tension. As the film thickness, t_m and t_n shown in Fig. 3.4 are used in (**a**) and (**b**), respectively. Reprinted with permission from Ref. [25]. Copyright 2005 by American Physical Society

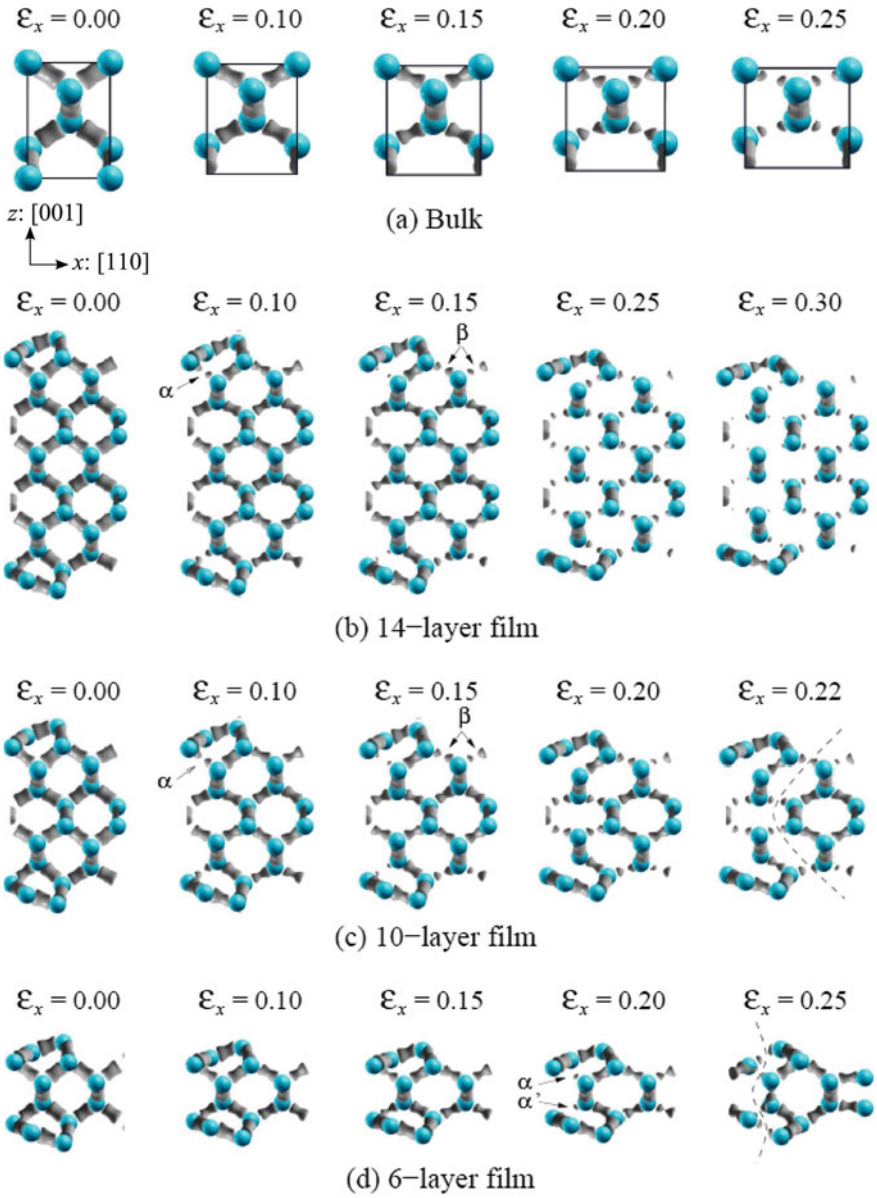

Fig. 3.6 Change in atomic configuration and charge density distribution of Si bulk and (001) nanofilms under [110] uniaxial tension [5, 25]. Reprinted with permission from Ref. [25]. Copyright 2005 by American Physical Society

Figure 3.6 shows the changes in the atom arrangement and charge distribution of the Si(001) thin film and bulk during [110] uniaxial tension. Blue spheres represent Si atoms and interatomic bonds are depicted by the charge density isosurface of

$0.425 \times 10^3\,\mathrm{nm}^{-3}$. At $\varepsilon_x < 0.10$, the atom arrangement and charge distribution in the thin film except for the dimer layers are close to those in the bulk. Here, the stress attained by the thin film calculated using the thickness t_m (excluding the dimer layers), based on the assumption that the dimer layers do not account for the tensile stress, should be almost the same as that by the bulk. Indeed, the stress–strain curves of the film and the bulk in Fig. 3.5a are close to each other at $\varepsilon_x < 0.10$, indicating that the assumption is valid. In other words, the initial resistance against deformation of the Si(001) thin film is little affected by the surface.

On the other hand, the ideal structural strength and the instability modes of the Si (001) thin film are affected by the surface. In the case of tensile deformation in a ductile material such as Si, its instability mode is likely to become that of cleavage due to bond rupture, and the ideal structural strength is given by the maximum value of the stress–strain curve. To be rigorous, of course, the onset of instability and its deformation mode must be evaluated by the mechanical stability analysis. Since Si thin films exhibit inhomogeneous deformation as will be presented later, a scheme that can treat inhomogeneous deformation (e.g., the method by Kitamura et al. [8, 9]) must be used. As shown in Fig. 3.6, cleavage fracture in thin films occurs from bond α located beneath a dimer, which is presumably because bond α becomes weaker than the other bonds due to charge transfer to form the dimer. Indeed, it is found that bond α is slightly thinner than other bonds at $\varepsilon_x = 0.00$. In the 14-layer model, bond β breaks subsequently after the break of bond α, before bonds inside the film break (see Fig. 3.6b, $\varepsilon_x = 0.15$). Then, the bonds inside the film undergo stretch uniformly until they break nearly simultaneously to cause cleavage fracture (Fig. 3.6b, $\varepsilon_x = 0.25$). The fracture behavior of the 10-layer model is similar to that of the 14-layer model only before the break of bond β. The inside bonds stretch inhomogeneously and part of them break to bring about cleavage fracture along the dashed line in Fig. 3.6c ($\varepsilon_x = 0.22$). The 6-layer model exhibits a cleavage fracture process different from those of the above two models. After the break of bond α, the dimers instead of bond β break before cleavage fracture (Fig. 3.6d, $\varepsilon_x = 0.25$). Assuming that the tensile load is carried just before the cleavage fracture by the atomic layers that are parallel to the tensile direction and are not broken yet, the tensile stress should be distributed to six layers (four subsurface and two surface layers) in the 14-layer model, four in the 10-layer model, and two in the 6-layer model. This explains the fact that the ideal structural strength decreases with decreasing thickness as shown in Fig. 3.5; that is, the fraction of the number of layers where the stress is attained is reduced as the film becomes thinner.

Investigation has also been done for the ideal structural strength and the instability modes and their dependence on the crystal orientation, the combination of stress components, and temperature in Si(001) thin films. According to stress–strain curves of Si(001) thin films with 14 atomic layers under [110] and [$\bar{1}$10] uniaxial tensions obtained by first-principles calculations, where thickness t_m (excluding the dimer layers) is used for the calculation of the tensile stress, the film exhibits nearly the same ideal structural strength under the two tensile directions, 15.2 and

15.4 GPa for [110] and [$\bar{1}$10] directions, respectively. In contrast, there is larger difference in the critical stress, which is 0.21 for the [110] tension and 0.18 for [$\bar{1}$10]. In addition, a sharp drop in the stress is seen after the critical strain in the latter case.

Change in the atom arrangement and charge distribution of the 14-layer model under the [$\bar{1}$10] uniaxial tension was also investigated. From $\varepsilon_y = 0.18$–0.20, the dimer structure changes from $p(2 \times 1)$ asymmetric to $p(2 \times 1)$ symmetric, before reaching cleavage fracture by bond breaking. The sharp drop in stress after the critical strain is caused by stress relaxation due to the change in the surface structure. The instability mode under the [$\bar{1}$10] uniaxial tension is therefore different from that under the [110] tension in Fig. 3.6. Change in the structure of the Si(001) surface due to tension was also discussed in Ref. [26], which reported that the $c(4 \times 2)$ structure alters to $p(2 \times 1)$ symmetric or a structure with symmetric and asymmetric dimers arranged alternately when the surface undergoes the [$\bar{1}$10] uniaxial tension or the [110]/[$\bar{1}$10] biaxial tension. An analysis by first-principles calculations of biaxial tension of Si(001) thin films reported reduction of the biaxial elastic coefficient and the maximum biaxial stress with increasing film thickness [27], where the surface reconstruction was ignored and the surface structure was assumed to stay $p(1 \times 1)$. The details of instability modes and the critical point for unstable deformation were not examined. A molecular dynamics study using the Stillinger–Weber potential reported that the (001) thin film exhibits not cleavage but slip deformation along {111} planes in the [100] uniaxial tension at 300 K [28]. However, as was discussed in Ref. [28], plastic deformation due to slip in Si at room temperature has barely been experimentally observed. In addition, simulation results also differ between brittle and ductile behaviors depending on the empirical potential function used. The above simulation result therefore has room for further discussion.

It is also well known that the Si(111) surface is reconstructed. The most well-known reconstruction of the Si(111) surface structure is the 7×7 dimer-adatom-stacking-fault (DAS), the detail of which can be found in Ref. [24] for example. Besides, the 2×1, 3×3 DAS, and 5×5 DAS structures are known to exist. A first-principles calculation study showed that the Si(111) surface structure can change under in-plane biaxial stretch and compression [29]. Although there has been no study of the ideal structural strength and instability modes of Si (111) thin films, unstable deformation is expected to occur at or near nonuniform region at the surface (e.g., the dimer and adatom in the DAS structure). The ideal structural strength and instability modes of Si(111) thin films are expected to depend on the crystal orientation, the combination of stress components and temperature, similarly to the Si(001) thin films.

A first-principles calculation study of tension of Cu(111) thin films [5] is introduced here as an example of investigation of the ideal structural strength of nanofilms without surface reconstruction. In this analysis, a thin film is represented by a slab model having vacuum layers with the periodic boundary condition, as is in the case of Si thin films in Fig. 3.4. The thickness of the films was defined as the

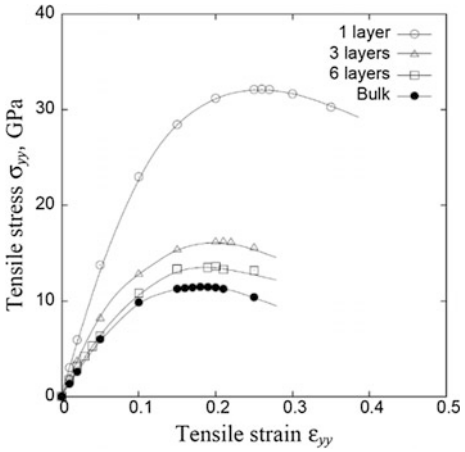

Fig. 3.7 Tensile stress–strain curves of Cu bulk and (111) nanofilms in $[11\bar{2}]$ uniaxial tension [5]

number of atomic planes times the separation between (111) planes in bulk. This definition should be valid because the Cu(111) surface is not reconstructed and the interlayer spacing in the films is nearly equal to that in bulk. Figure 3.7 shows the stress–strain curves of the Cu(111) thin films under $[11\bar{2}]$ uniaxial tension, where the results of the one-, three- and six-layer models and bulk are presented. It is interesting to find that the maximum stress increases with decreasing thickness in contrast to the case of Si. The dependence of the maximum stress on the film thickness should stem from the nature of electronic structure since there is no surface reconstruction and no difference in the atom arrangement. Figure 3.8 presents change in the atom arrangement and charge distribution during the $[11\bar{2}]$ uniaxial tension in the one-layer film model and bulk. Comparison of the charge distribution at $\varepsilon_y = 0.00$ finds higher charge density in the (111) plane of the thin film than that of bulk, which is due to charge transfer from between the (111) planes to the surface, leading to stronger interatomic bonding in the surface layer. This explains the increasing strength with decreasing thickness in terms of the fraction of surface layers in the thin films.

Whether or not the maximum stress in Fig. 3.7 represents the ideal structural strength is not clear because mechanical instability analysis was not performed. It should therefore be pointed out the possibility that the system may become mechanically unstable before the maximum point of the stress–strain curve. However, the possibility of slip deformation due to $\{111\}\ \langle 11\bar{2}\rangle$ shear instability is excluded in the $[11\bar{2}]$ tension of the Cu(111) thin film due to the following reason: The shear stress on a slip system τ is related to the uniaxial tension σ as

$$\tau = \sigma \cos\phi \cos\lambda$$

where ϕ is the angle between the tensile direction and the normal of the slip plane, and λ is the angle between the tensile direction and the slip direction. Assuming that shear instability occurs at $\tau = \tau_c$, the required uniaxial tension σ_c becomes

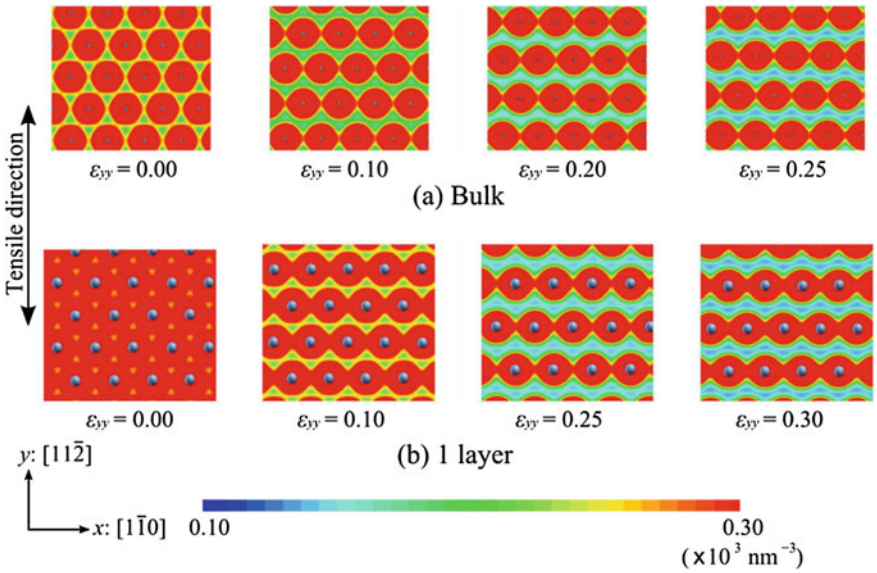

Fig. 3.8 Change in atomic configuration and charge density distribution of Cu bulk and 1-layer (111) nanofilm under [11$\bar{2}$] uniaxial tension [5]

$$\sigma_c = \frac{\tau_c}{\cos\phi \cos\lambda}$$

In the Cu(111) thin film under the [11$\bar{2}$] tension, therefore, the infinitely large tensile stress is required to cause $\{111\}\ \langle 11\bar{2}\rangle$ shear instability because $\cos\phi \cos\lambda = 0$. By the way, this approach was also applied to the ideal strength of the fcc perfect crystal under $\langle 100\rangle$, $\langle 110\rangle$, and $\langle 111\rangle$ uniaxial tensions [30].

In the [100] uniaxial tension of Cu(001) thin films, $\{111\}\ \langle 11\bar{2}\rangle$ shear instability can occur since σ_c takes a finite value. Indeed, classical molecular dynamics study revealed that the instability is caused by the nucleation of partial dislocations from a surface at finite temperatures [31, 32]. Although instability modes at 0 K have yet to be analyzed, a Cu(001) thin film with a sufficient thickness may exhibit a deformation bifurcation from fcc to bcc structures with the Born instability under the [100] uniaxial tension at 0 K, which can be conjectured by the fact that the Born instability occurs in the perfect crystal under the [100] uniaxial tension and the Cu (001) surface is not reconstructed. This also indicates the possibility of emergence of other instability modes than slip or bifurcation since the structure of the one-layer Cu(001) film is no longer fcc. That is, the one-layer thin film is unique and deserves to be in a special category among the two-dimensional periodic nanostructures.

It is graphene that has been most well known and most actively studied as the one-layer film. Here, we introduce a first-principles analysis of the ideal structural strength and instability modes of graphene under uniaxial tension [33]. Graphene

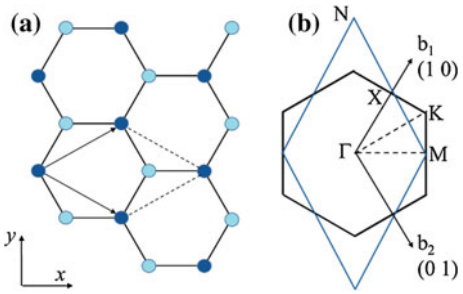

Fig. 3.9 **a** Schematic of graphene with translation vectors (*arrows*) and two-atom primitive cell (*dashed line*), **b** primitive cell of the reciprocal lattice (*green line*) and the first Brillouin zone (*black line*)

structure is schematically shown in Fig. 3.9a, where the arrows and the dashed lines indicate the translation vectors and the primitive unit cell, respectively. Figure 3.9b shows the primitive unit cell and the first Brillouin zone in the reciprocal space. Since the primitive unit cell contains two atoms, the phonon dispersion diagram has six curves. As was explained in Sect. 3.2, the system is unstable when the square of the phonon frequency $\omega(\mathbf{q}, s)$ is negative (phonon soft mode).

Figure 3.10 shows change in tensile stress and Poisson's ratio of graphene under uniaxial tension along the x and y directions. Here, the graphene structure was realized by a slab model with vacuum layers in the periodic boundary condition. The simulation cell dimension normal to the graphene $L_z = 0.8$ nm, which is more than twice as thick as the interlayer distance of graphite, $d = 0.334$ nm, should be large enough to represent the one-layer graphene structure. The equivalent stress shown in Fig. 3.10 was obtained by the stress on the entire simulation cell multiplied with L_z/d, meaning that the stress was calculated regarding the graphene as a thin plate with a thickness of 0.334 nm. The graphene exhibits isotropic elasticity in the small strain region ($\varepsilon = 0 - 0.05$) where the Young's modulus and Poisson's ratio are nearly the same between the two uniaxial tensile directions (the x- and y-axes). On the other hand, the stress–strain curves obviously differ in the large strain region ($\varepsilon > 0.15$), exhibiting dependence on the tensile orientation. While the maximum stress and the corresponding strain in tension along the x-axis are 110 GPa and 0.194, respectively (red circles in Fig. 3.10), those for the y-axis tension are 121 GPa and 0.266 (red triangles in Fig. 3.10), showing the larger strength along the x direction than along the y direction.

The phonon dispersion curves at $\varepsilon_{xx} = 0.194$ are shown in Fig. 3.11a, where the imaginary phonon frequency ($\omega^2(\mathbf{q}, s) < 0$) is plotted as the negative value for convenience. It is seen that a phonon soft mode appears near the Γ point. Besides the two points near the Γ point, the phonon frequency is real in the Brillouin zone. Therefore, the phonon soft mode that appears near the Γ point at $\varepsilon_{xx} = 0.194$ is the first instability mode. Figure 3.11c shows the eigenvector corresponding to the phonon soft mode. The eigenvector is parallel to the tensile direction (the x-axis), implying that the graphene exhibits cleavage fracture associated with bond breaking under the uniaxial tension along the x-axis. As shown in Fig. 3.11d–f, the phonon property of the graphene under uniaxial tension along the y-axis is similar to that in the case of the x-axis tension. A phonon soft mode appears near the Γ point at

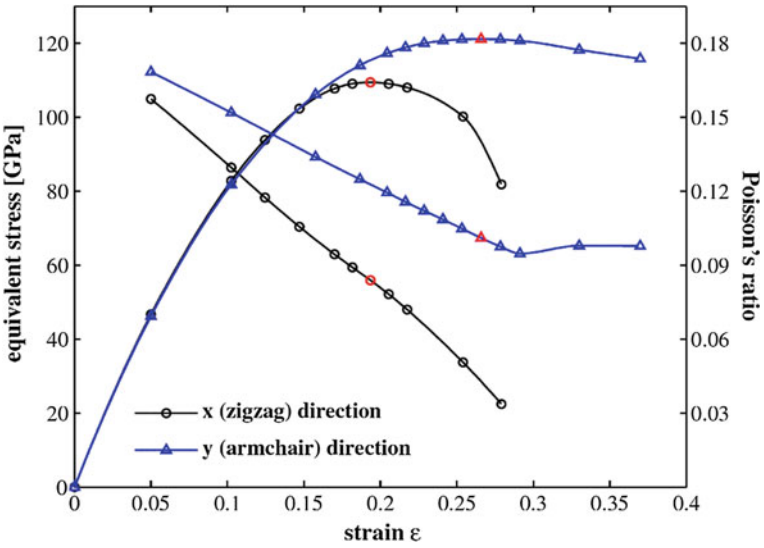

Fig. 3.10 Uniaxial tensile stress–strain relations of graphene (the *curves* connected to the origin) and the Poisson's ratios as functions of uniaxial strains (the *lines* with initially negative slopes). Reprinted with permission from Ref. [33]. Copyright 2007 by American Physical Society

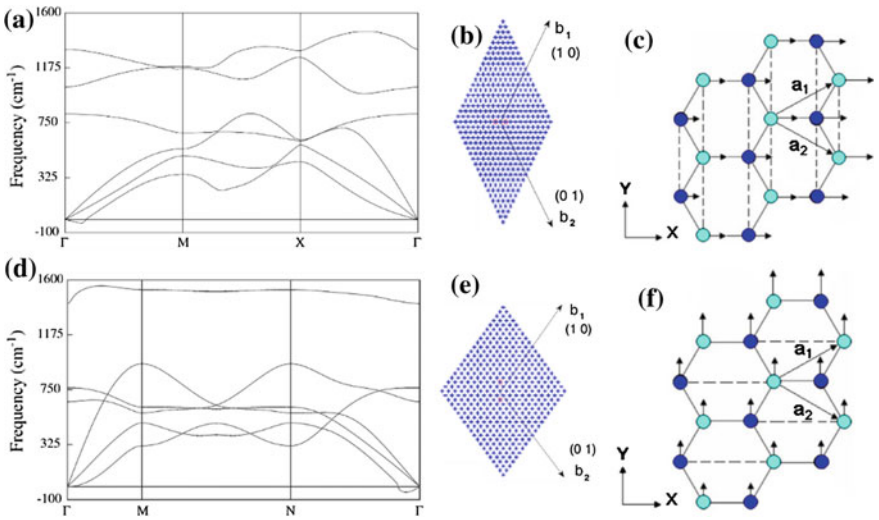

Fig. 3.11 **a**, **d** Phonon dispersion curves of graphene, **b**, **e** scan of the entire Brillouin zone, and **c**, **f** the unstable eigenvector corresponding to the soft mode, at **a**–**c** $\varepsilon_{xx} = 0.194$ and **d**–**f** $\varepsilon_{yy} = 0.266$. Reprinted with permission from Ref. [33]. Copyright 2007 by American Physical Society

$\varepsilon_{xx} = 0.266$ and its eigenvector is parallel to the tensile direction. However, in contrast to the x-axis tension where the interatomic bonds attaining the load are parallel to the tensile direction, not only stretch but also rotation is applied to the bonds in the y-axis tension where the bonds are oblique with respect to the tensile direction by 30°. As the instability mode in the graphene subject to the y-axis tension, brittle bond breaking occurs at 0 K while ductile Stone–Wales (SW) transformation associated with bond rotation may arise depending on temperature, which will be discussed in more detail in Sect. 3.3.2.

As seen above, the ideal structural strength and instability modes of graphene depend not only on the crystal orientation but also on temperature and the stress condition. A first-principles approach with DFT calculations combined with the quasi-harmonic approximation (QHA) [34] and classical molecular dynamics simulations [35–37] revealed with slight quantitative difference that the ideal structural strength of graphene decreases nearly linearly with increasing temperature. In graphene under biaxial tension, a phonon soft mode appears at the K point instead of the Γ point and unstable deformation called K_1 mode emerges according to a first-principles study [38]. The study also investigated various conditions of biaxial tension to clarify that the cleavage fracture (K_1) mode is likely to occur under stress conditions close to the uniaxial (biaxial) tension.

In this subsection, we explained the ideal structural strength and instability modes of two-dimensional periodic nanostructures through the examples of the Si thin film, Cu thin film, and graphene. Of course, there are a substantial number of two-dimensional periodic nanostructures. Among them, especially thin films consisting of one atomic layer (monolayer) like graphene are attracting much attention, such as silicone, phosphorene, graphene, silicane, BN monolayer (boron nitrene), and MoS_2 monolayer. Unlike graphene, silicone and phosphorene show a corrugated structure due to in-plane buckling at equilibrium. Graphane and silicane are graphene and silicone with the adsorption of hydrogen atoms, respectively. BN (MoS_2) monolayer is the graphene structure substituting C atoms with B and N (Mo and S). The ideal structural strength and instability modes of these monolayers are being intensively studied [39–43].

3.3.2 *One-Dimensional Nanostructures*

A nature that is unique to one-dimensional periodic structures is the presence of the junction between surfaces (edge). In one-dimensional periodic structures that have relatively large fraction of edge region in the entire volume, the effect of edges cannot be ignored. A typical example of the one-dimensional periodic structure having edges is the nanowire. Figure 3.12 shows first-principles calculation results of Si⟨100⟩, ⟨110⟩, and ⟨111⟩ nanowires with diameters of $2 - 5$ nm [44]. The top figures are for the cross section and the bottom for the side. Here, uvw denotes the direction of the nanowire axis. The {100} surfaces in the Si⟨100⟩ nanowires have the dimer structure eliminating dangling bonds (red atoms in Fig. 3.12a).

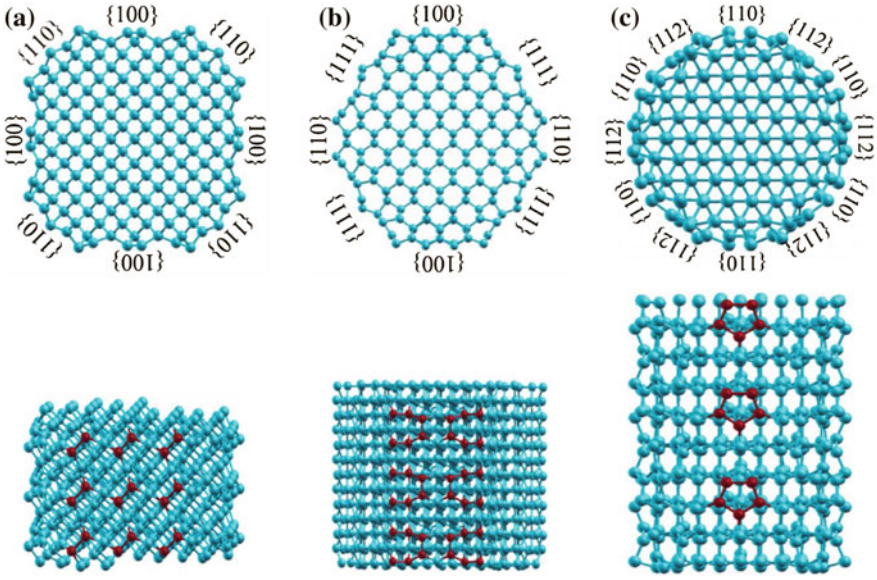

Fig. 3.12 Si nanowires with **a** ⟨100⟩, **b** ⟨110⟩, and **c** ⟨111⟩ orientations. *Top* and *bottom* panels show the cross-sectional and lateral views, respectively. Reprinted with permission from Ref. [44]. Copyright 2009, American Institute of Physics

The surface structure is $c(2 \times 2)$, instead of $p(2 \times 2)$ or $c(4 \times 2)$ that have lower energies, because the $p(2 \times 2)$ or $c(4 \times 2)$ structure adjoining an edge would require atoms having two dangling bonds to appear on the edge, resulting in a high-energy configuration. First-principles calculations showed that Si⟨100⟩ nanowires with diameters of several nanometers should possess the $c(2 \times 2)$ surface structure because the energy gain by $c(2 \times 2)$ surfaces eliminating dangling bonds on edges exceeds that by surface reconstruction to $p(2 \times 2)$ or $c(4 \times 2)$ [45]. In the Si⟨110⟩ nanowires, step structures appear on the {100} surfaces due to the effect of edges between {100} and {111} surfaces (red atoms in Fig. 3.12b). In the Si⟨111⟩ nanowires, five-membered ring structures appear on the edges as the edges with two {110} surfaces are reconstructed to small {112} facets (red atoms in Fig. 3.12c). These results demonstrate that the edge can not only affect the surface reconstruction but also reconstruct itself into a structure different from the surface and subsurface structures.

The ideal structural strength and instability modes of nanowires are intermingled with the edge. Molecular dynamics simulations using the Stillinger–Weber potential for uniaxial tension of Si⟨100⟩ nanowires at 300 K investigated the effect of the edge by comparing nanowires with square and round cross sections [46]. In both nanowires, {111} slip deformation associated with the nucleation and propagation of dislocations occurs as the instability mode. The dislocation is nucleated at the edge in the nanowire with the square cross section, which is because the energy required for the dislocation nucleation is lower at the edge than at the surface. As a

result, the nanowire with the square cross section shows the earlier onset of slip deformation and the lower ideal structural strength than the round-shaped nanowire. Through molecular dynamics calculations of Si nanowires with different interatomic potentials (Stillinger–Weber and MEAM); different crystal orientations (⟨100⟩ and ⟨123⟩); stress conditions (tension and compression); and different temperatures (300 and 1200 K), it was shown that a crack or an amorphousized shear band instead of {111} slip can occur as the instability mode depending on the condition [47]. However, it remains the same that the instability mode occurs at the edge in any case.

Even in nanowires that show no reconstruction at the surface or the edge, the ideal structural strength and instability modes are influenced by the edge. For example, molecular dynamics simulations of Cu nanowires with fcc structures using the EAM potential revealed that unstable deformation occurs from the edge in both Cu⟨100⟩ and ⟨110⟩ nanowires under tension at finite temperatures [48–51]. The instability mode differs between the crystal orientations. More specifically, {111} slip associated with dislocation nucleation and propagation from the edge occurs in the Cu⟨100⟩ nanowire [48, 49] while lattice reorientation from ⟨110⟩ to ⟨100⟩ occurs in the Cu⟨110⟩ nanowire [50, 51]. In the latter, after a twin boundary is formed by the nucleation of 1/6 ⟨112⟩ partial dislocations at the edge and their propagation on {111} planes, the twin boundary moves along the wire axis, resulting in the reorientation from ⟨110⟩ to ⟨100⟩. The stress causing the lattice reorientation (i.e., the ideal structural strength) and the wire shape after the reorientation depend on the initial surface and edge of the nanowire. The rhombus cross section of the Cu⟨110⟩ nanowire consisting of four {111} surfaces becomes a rectangle with four {100} surfaces due to lattice reorientation triggered by partial dislocation emission from the acute edges [50]. On the other hand, the rectangular cross section of the Cu⟨110⟩ nanowire consisting of two {111} and two {112} surfaces transforms to a rhombus with two {100} and two {112} surfaces [51]. The difference in the edge angle leads to different ideal structural strengths because the energy required for partial dislocation nucleation changes.

Unlike the Cu nanowires with the fcc structure, nanowires with the side dimension less than 1 nm can transform to another structure by reconstruction in the entire volume. For example, a first-principles study clarified that Cu⟨100⟩ fcc nanowires having four {110} surfaces with the side dimension of 0.260, 0.520, and 0.779 nm are reconstructed to the bct (body-centered tetragonal) structure at equilibrium [52]. When tension is applied to the bct nanowire, elastic deformation as the bct structure appears first, followed by unstable deformation to the fct (face-centered tetragonal) structure. Then, the nanowire exhibits elastic deformation as the fct structure until fracture [52].

There are nanowires having peculiar structures rather than the original crystal structures as a result of the reconstruction of the entire structure. An example is core–shell nanowires consisting of a core in the center and a surrounding shell. Cu core–shell nanowires consisting of a fcc ⟨110⟩ chain (core) and a fcc {111} sheet (shell) were examined in Refs. [5, 53]. The 5-1 nanowire consists of a shell with five atoms along its circumference and a chain. The 10-5-1 nanowire consists of a 5-1 nanowire

wrapped with a shell with 10 atoms with its circumference. The solid lines in the left figure indicate simulation cells in the following first-principles calculations. Vacuum regions are set in the x and y directions normal to the wire axis, and the boundaries are periodic to mimic isolated nanowires. The tensile stress σ_z was calculated as $\sigma_z = \sigma_z^{\text{cell}} \times L_x^{\text{cell}} \times L_y^{\text{cell}} / \pi r^2$ where σ_z^{cell} is the stress along the z direction for the entire cell; L_x^{cell} and L_y^{cell} are the cell size in the x and y directions, respectively; and r is the wire radius. It should be noted here that the definition of the wire radius has arbitrariness similarly to the thickness of nanofilms (see Sect. 3.3.1).

Figure 3.13 shows stress–strain curves of the 5-1 and 10-5-1 Cu nanowires under tension obtained by first-principles calculations [5, 53]. Using the separation between {111} layers in the Cu perfect crystal (a_0), the radius of the 5-1 nanowire was defined as the distance between the chain and the 5-shell + $a_0/2$, that of the 10-5-1 nanowire as the distance between the chain and the 5-shell + the distance between the 5-shell and the 10-shell + $a_0/2$. The results of the isolated chain, 5-shell, sheet, and bulk are also shown for comparison. The 5-1 and 10-5-1 nanowires show the Young's modulus and the maximum stress lower than those of the isolated chain and shell. The calculated Young's moduli of the 5-1 and 10-5-1 nanowires, 178 and 140 GPa, respectively, are significantly different from an estimation by the simple combination law (344 and 322 GPa, respectively). The discrepancy implies the mechanical properties of the core–shell nanowires differs from those of the isolated components owing to the difference in their electronic structure. Figure 3.14 shows the charge density distribution of the 5-1 nanowire subtracted by that of the chain and the 5-shell. The charge density decreases in the core and shell while it increases between the core and shell, indicating charge transfer from the core and shell to the region in between. The weakening of the bonds in the core and shell due to the charge transfer explains the lower Young's modulus and maximum stress of the core–shell nanowires.

Since the core–shell nanowires possess no fcc structure, the instability mode is expected to be cleavage fracture due to bond breaking rather than {111} slip

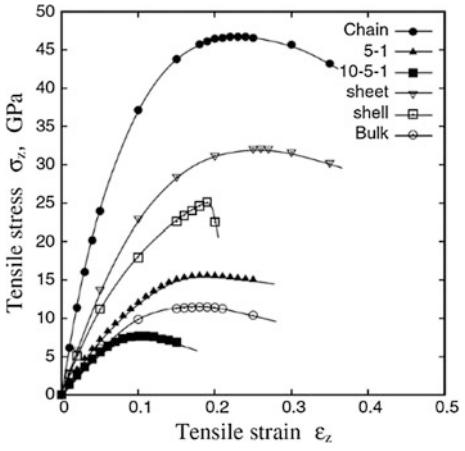

Fig. 3.13 Tensile stress–strain curves of Cu monatomic chain, 5-1 and 10-5-1 nanowires, sheet, shell (curled sheet), and bulk. Reprinted with permission from Ref. [53]. Copyright 2006 by Institute of Physics

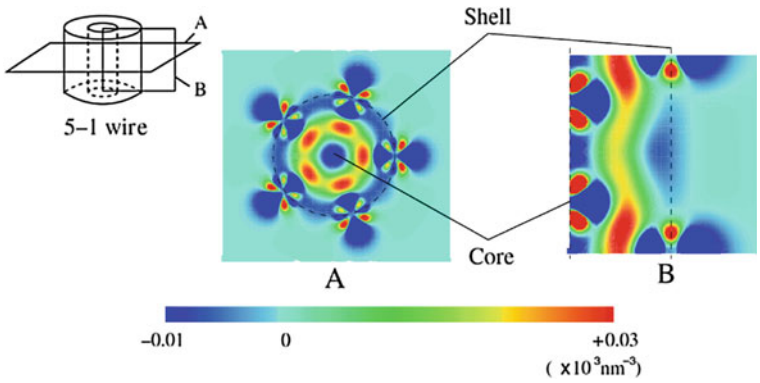

Fig. 3.14 Difference in charge density distribution between the Cu 5-1 nanowire and freestanding core and shell. Reprinted with permission from Ref. [53]. Copyright 2006 by Institute of Physics

deformation or lattice reorientation. Indeed, this is supported by the change in charge distribution during tension (Fig. 3.15). In the 5-1 nanowire, bonds in the 5-shell break first followed by the break of the core before cleavage fracture (Fig. 3.15a). In the 10-5-1 nanowire, similarly, the bond breaking occurs from the 10-shell leading to cleavage fracture (Fig. 3.15b). The critical strain of the 10-5-1 nanowire is lower than that of the 5-1 nanowire, which is because inhomogeneous charge distribution in the 10-5-1 creates weak bonds already at vanishing strain. Deformation concentration on the weak bonds causes the smaller critical strain. Bond breaking occurs from the shell because the chain is stronger than the shell as shown in Fig. 3.13.

The chain is an ultimate form of one-dimensional periodic nanostructures because it has the smallest possible diameter. The chain form is the strongest among structures consisting of the same atom species (Fig. 3.13) owing to its electronic structure. Since charge density between Cu atoms is higher in the chain than in the bulk, the chain possesses stronger interatomic bonds. The high charge density in the chain is due to the smaller coordinate number; that is, the reduction in the coordinate number leads to a larger number of electrons per bond, causing stronger interatomic bonds. As the bond breaking is the only possible instability mode in the chain, the maximum point in the stress–strain curve gives the ideal structural strength of the chain.

The nanotube is a one-dimensional periodic nanostructure that is distinct from the nanowire and chain explained above. Carbon nanotubes (CNTs), the most well-known one-dimensional nanomaterials, consist of rolled-up graphene sheets, whose structure is denoted by the chiral vector (n, m) that represents the circumference of the tube. Using two translation vectors $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$ shown in Fig. 3.9a, the chiral vector ($\boldsymbol{C}$) of a (n, m) CNT is denoted as $n\boldsymbol{a}_1 + m\boldsymbol{a}_2$. The $(n, 0)$ and (n, n) CNTs are called zigzag and armchair types, respectively, while the other CNTs are called chiral types. The angle between $\boldsymbol{a}_1$ and $\boldsymbol{C}$, or chiral angle χ, is $\chi = 0^\circ$ in zigzag, $\chi = 30^\circ$ in armchair and $0 < \chi < 30^\circ$ in chiral CNTs. Stress attained by a

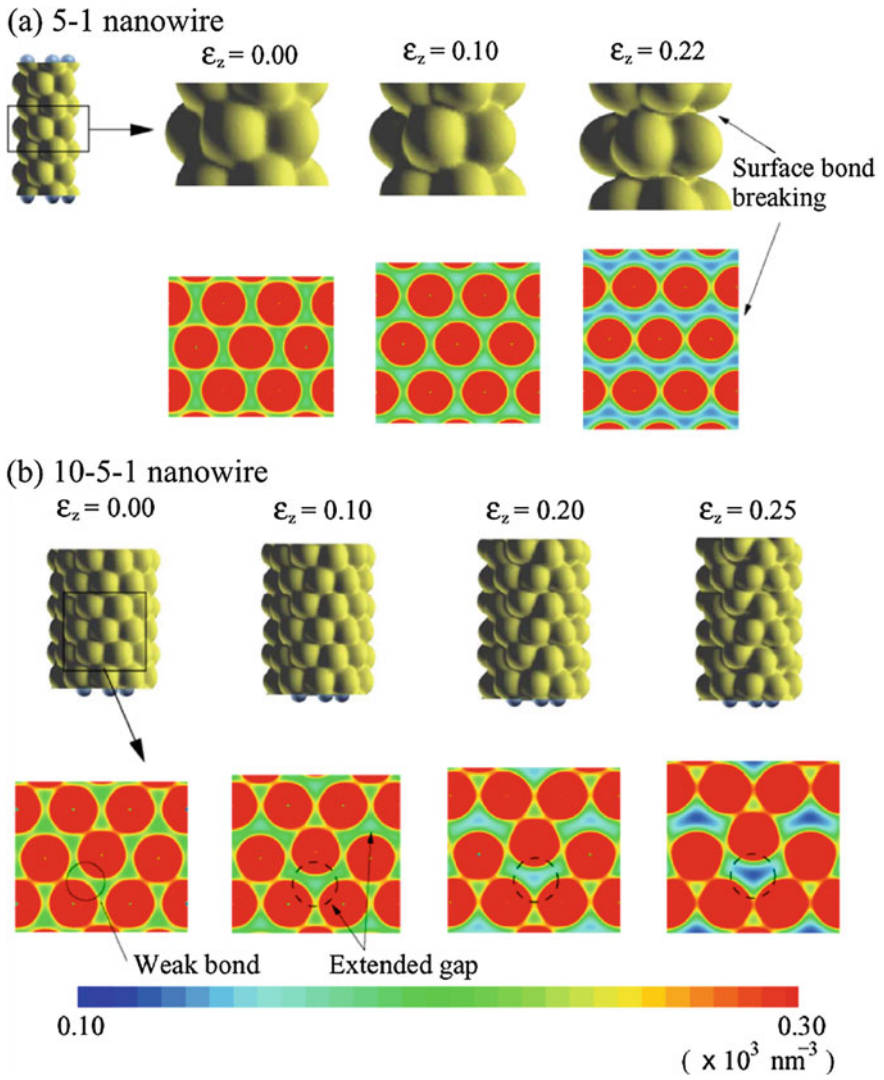

Fig. 3.15 Change in atomic configuration and charge density distribution of Cu **a** 5-1 and **b** 10-5-1 nanowires under uniaxial tension. Reprinted with permission from Ref. [53]. Copyright 2006 by Institute of Physics

CNT is often calculated assuming the CNT as a hollow cylinder with a wall thickness of 0.34 nm, which is the interlayer spacing in graphite.

Unstable deformation in a CNT occurs due to brittle cleavage fracture associated with bond breaking or the ductile Stone–Wales (SW) transformation with bond rotation [54], as schematically shown in Fig. 3.16. Configurations with breaking of some bonds shown in blue in Fig. 3.16a are depicted in Fig. 3.16b–d. Figure 3.16e

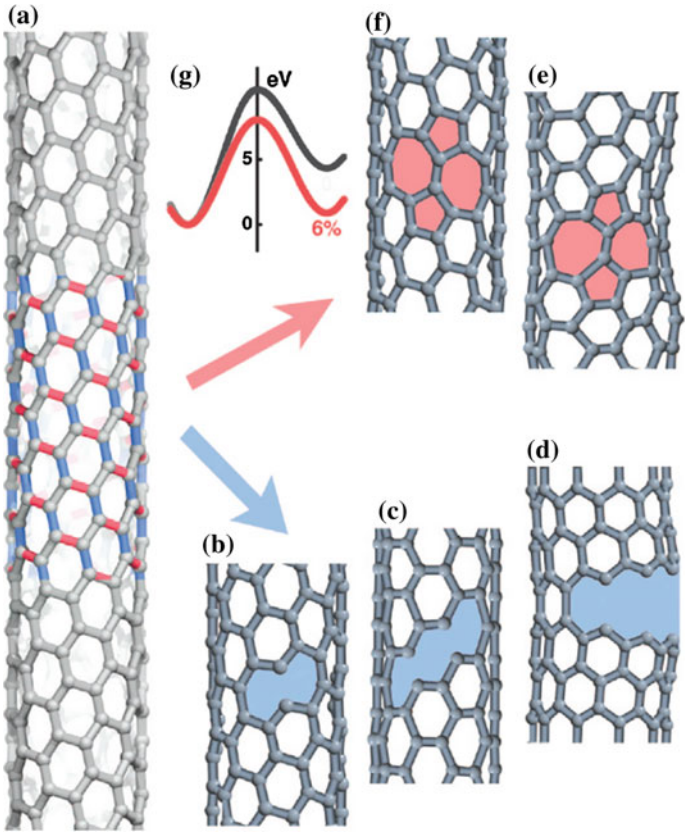

Fig. 3.16 Two instability modes in CNTs. **a** A CNT under uniaxial deformation, resulting in **b**–**d** brittle cleavage associated with bond breaking or **e**, **f** ductile SW transformation associated with bond rotation. **g** SW activation barrier at the temperature of 0 K and uniaxial strain of 0% (*black*) or 6% (*red*). Reprinted with permission from Ref. [54]. Copyright 2006 by National Academy of Sciences

and f shows configurations where rotation of a bond marked with red led to the combination of 5-membered ring and 7-membered rings. The 5/7/7/5 structure created by the bond rotation is called the Stone–Wales defect. Figure 3.16g presents the activation energy barrier for the SW transformation in a chiral CNT shown in Fig. 3.16a at 0 K, where the black and red curves are for strains of 0 and 6% in uniaxial tension, respectively. According to first-principles calculations of uniaxial tension of (8,0), (9,0), (10,0) zigzag, and (8,8) armchair CNTs, the critical stress σ_{bb} and strain ε_{bb} for bond breaking in armchair CNTs are larger than those in zigzag [55], being consistent with the orientation dependence found in the uniaxial tension of graphene (Fig. 3.10). More generally, ε_{bb} becomes larger as the chiral angle χ is larger (closer to that of armchair). In contrast, the critical strain ε_{SW} for the SW transformation becomes smaller as χ is larger according to first-principles

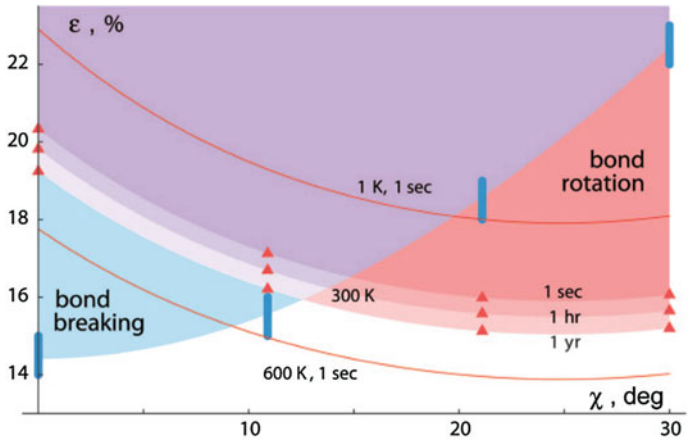

Fig. 3.17 Critical uniaxial strain of CNTs as a function of chiral angle (χ) at various load durations (t) and temperatures (T). As χ increases, the critical strain for bond breaking (*blue marks*) increases, while the critical strain for bond rotation (*red marks* and *lines*) decreases. The crossover between bond breaking and rotation shifts left or right depending on t and T. Reprinted with permission from Ref. [54]. Copyright 2006 by National Academy of Sciences

calculations combined with the reaction rate theory [54, 56]. While ε_{bb} is barely influenced by the tube radius, temperature, and strain rate, ε_{SW} is reduced with smaller radius, higher temperature, and slower strain rate, as schematically summarized in Fig. 3.17. The critical point of whether bond fracture or bond rotation occurs is determined by the intersection between the $\varepsilon_{bb} - \chi$ curve (interpolation of blue bars) and the $\varepsilon_{SW} - \chi$ curve (interpolation of red triangles for 300 K or red curves for 1 and 600 K).

In this subsection, the ideal structural strength and instability modes of one-dimensional periodic nanostructures were explained with the examples of Si nanowires, Cu nanowires, and CNTs. There have been a lot of types of nanotubes and nanowires fabricated thus far besides these examples and more will undoubtedly be realized. The existence of the peculiar nanostructures such as core–shell nanowires and atomic chains suggests the production feasibility of a lot more types of one-dimensional periodic nanostructures than bulk materials that exist in reality. Considering various factors to explore including the material species, crystal orientation, loading condition, and temperature, the ideal structural strength of one-dimensional periodic structures has been only little understood thus far.

3.3.3 Zero-Dimensional Nanostructures

Zero-dimensional nanostructures such as nanoparticles and nanoclusters lie out of the definition of the ideal shape of low-dimensional periodic structures. In addition,

it is extremely difficult to evaluate the critical stress for instability under homogeneous stress or strain because loading on a zero-dimensional nanostructure can only be achieved by a point load, which does not realize a homogenous loading state. Since zero-dimensional nanostructures are therefore out of the scope of this section, we only briefly refer to some of experiments and simulations that evaluated the mechanical properties of such nanostructures.

The nanoparticle is a particle with a diameter in the order of nanometer with the crystal structure identical with bulk material. In situ compression test with HRTEM and classical molecular dynamics calculations have revealed that Ag nanoparticles at room temperature can be largely deformed like a liquid drop by compression and return to their initial shape by unloading [57]. The study confirmed that this quasi-elastic deformation process is not associated with dislocation activities or melting inside the particles where the crystal structure is maintained. The liquid-like deformation behavior of Ag nanoparticles is essentially different from that of bulk and is brought not by dislocations or melting but by diffusion of surface atoms.

The nanocluster should be categorized separately from the nanoparticle because it possesses atom arrangements distinct from that in bulk. The C_{60} fullerene is a soccer-ball-shaped nanocluster consisting of 5-membered and 6-membered rings of carbon atoms and can be regarded as an ideal zero-dimensional nanostructure due to its high symmetry. First-principles calculations were performed to obtain strain energy of C_{60} and $X@C_{60}$ encapsulating atom X (= He, Ne, Ar or Kr), which revealed that while the tensile strain energies of C_{60} and $X@C_{60}$ are nearly identical with each other, the compressive strain energy of $X@C_{60}$ is larger than that of C_{60} [58–61]. The result indicates the possibility of enhancing the strength against compression of C_{60} by doping inert gas atoms. Deformation behavior of Si nanoclusters was investigated by first-principles calculations, which clarified their spring-like behavior when displacements in the opposite directions are applied to the top and bottom atoms [62].

Little investigation of the mechanical properties of zero-dimensional nanostructures has been done by far compared to those of one- and two-dimensional nanostructures because of the practical difficulty of the investigation. Efforts of such investigation are emerging and further development is expected.

3.4 Conclusion

The definitions of the ideal strength, ideal understructure, and ideal shape were explained. Theoretical approaches for the ideal strength and structural instability were explained, and then theoretical investigations of ideal strength of various nanostructures with first-principles calculations were introduced. As theoretical methods to investigate the ideal strength and instability modes of materials have been established, they can be in principle applied to low-dimensional nanostructures.

Some caveats in such analyses (e.g., the definition of film thickness) were explained. Further development in the investigation of low-dimensional nanostructures, especially for one- and zero-dimensional structures, is awaited to achieve comprehensive understanding of the mechanical properties of nanomaterials.

References

1. M. Born, K. Huang, *Dynamic Theory of Crystal Lattices* (Oxford, UP, 1954)
2. G. Grimvall, B.M. Kope, V. Ozolins, K.A. Perrsson, Rev. Mod. Phys. **84**, 945 (2012)
3. J. Pokluda, M. Cerny, M. Sob, Y. Umeno, Prog. Mater Sci. **73**, 127 (2015)
4. S. Yip, *Handbook of Materials Modeling*, Springer (2005)
5. A. Kushima, Ph.D. thesis, Kyoto University, 2007
6. D.M. Clatterbudk, C.R. Krenn, M.L. Cohen, J.W. Morris Jr., Phys. Rev. Lett. **91**, 135501 (2003)
7. K. Yashiro, K. Yamagami, K. Kubo, Y. Tomita, JSME Int. J. A **49**, 100 (2006)
8. T. Kitamura, Y. Umeno, N. Tsuji, Comp. Mater. Sci. **29**, 499 (2004)
9. T. Kitamura, Y. Umeno, R. Fushino, Mater. Sci. Eng., A **379**, 229 (2004)
10. R.C. Cammarata, Prog. Surf. Sci. **46**, 1 (1994)
11. Y. Shiihara, M. Kohyama, S. Ishibashi, Phys. Rev. B **81**, 075441 (2010)
12. S. Giusepponi, M. Celino, Nucl. Instr. Meth. Phys. Res. B **342**, 70 (2015)
13. D. Hull and D. J. Bacon, *Introduction to Dislocations*, fifth ed., Butterworth-Heinemann (2011)
14. B. Joos, Q. Ren, M.S. Duesbery, Phys. Rev. B **50**, 5890 (1994)
15. R. Janisch, N. Ahmed, A. Hartmaier, Phys. Rev. B **81**, 184108 (2010)
16. X. Pang, N. Ahmed, R. Janisch, A. Hartmaier, J. Appl. Phys. **112**, 023503 (2012)
17. Y. Yan, T. Kondo, T. Shimada, T. Sumigawa, T. Kitamura, Mater. Sci. Eng., A **534**, 681 (2012)
18. D. Farkas, Curr. Opin. Solid State Mater. Sci. **17**, 284 (2013)
19. Y.L. Liu, H.B. Zhou, S. Jin, Y. Zhang, G.H. Lu, Chin. Phys. Lett. **27**, 127101 (2010)
20. S. Giusepponi, M. Celino, J. Nucl. Mater. **435**, 52 (2013)
21. Y. Kinoshita, Ph.D. thesis, Kyoto University, 2008
22. S.K. Yadav, R. Ramprasad, J. Wang, A. Misra, X.Y. Liu, Modelling Simul. Mater. Sci. Eng. **22**, 035020 (2014)
23. A. Ramstad, G. Brocks, P.J. Kelly, Phys. Rev. B **51**, 14504 (1995)
24. K.D. Brommer, M. Needels, B.E. Larson, J.D. Joannopoulos, Phys. Rev. Lett. **68**, 1355 (1992)
25. Y. Umeno, A. Kushima, T. Kitamura, P. Gumbsch, J. Li, Phys. Rev. B **72**, 165431 (2005)
26. Y.S. Kim, S.M. Lee, Phys. Rev. B **75**, 165304 (2007)
27. X.Y. Zhou, H. Ren, B.L. Huang, T.Y. Zhang, Phys. Lett. A **379**, 471 (2015)
28. P. Lin, R.I. Babicheva, M. Xue, H.S. Zhang, H. Xu, B. Liu, K. Zhou, Comp. Mater. Sci. **96**, 295 (2015)
29. R. Zhachuk, S. Teys, J. Coutinho, J. Chem. Phys. **138**, 224702 (2013)
30. M. Cerny, J. Pokluda, J. Phys.: Condens. Matter **21**, 145406 (2009)
31. Q.N. Guo, X.D. Yue, S.E. Yang, Y.P. Huo, Comp. Mater. Sci. **50**, 319 (2010)
32. S. Xu, Y.F. Guo, Z.D. Wang, Comp. Mater. Sci. **67**, 140 (2013)
33. F. Liu, P. Ming, J. Li, Phys. Rev. B **76**, 064120 (2007)
34. T. Shao, B. Wen, R. Melnik, S. Yao, Y. Kawazoe, Y. Tian, J. Chem. Phys. **137**, 194901 (2012)
35. H. Zhao, N.R. Aluru, J. Appl. Phys. **108**, 064321 (2010)
36. K. Min, N.R. Aluru, Appl. Phys. Lett. **98**, 013113 (2011)

37. M.A.N. Dewapriya, A.S. Phani, R.K.N.D. Rajapakse, Modelling Simul. Mater. Sci. Eng. **21**, 065017 (2013)
38. C.A. Marianetti, H.G. Yevick, Phys. Rev. Lett. **105**, 245502 (2010)
39. C. Yang, Z. Yu, P. Lu, Y. Liu, H. Ye, T. Gao, Comp. Mater. Sci. **95**, 420 (2014)
40. Q. Wei, X. Peng, Appl. Phys. Lett. **104**, 251915 (2014)
41. E.B. Isaacs, C.A. Marianetti, Phys. Rev. B **89**, 184111 (2014)
42. Q. Peng, S. De, Nanoscale **6**, 12071 (2014)
43. T. Li, Phys. Rev. B **85**, 235407 (2012)
44. D.B. Migas, V.E. Borisenko, J. Appl. Phys. **105**, 104316 (2009)
45. S. Ismail-Beigi, T. Arias, Phys. Rev. B **57**, 11923 (1998)
46. Z. Yang, Z. Lu, Y.P. Zhao, J. Appl. Phys. **106**, 023537 (2009)
47. J. Guénolé, J. Godet, S. Brochard, Modelling Simul. Mater. Sci. Eng. **19**, 074003 (2011)
48. H.A. Hu, Euro. J. Mech. A/Solids **25**, 370 (2006)
49. A. Cao, Y. Wei, E. Ma, Phys. Rev. B **77**, 195429 (2008)
50. W. Liang, M. Zhou, F. Ke, Nano Lett. **5**, 2039 (2005)
51. A. Cao, Y. Wei, Phys. Rev. B **74**, 214108 (2006)
52. N.T. Hung, D.V. Truong, Surf. Sci. **641**, 1 (2015)
53. A. Kushima, Y. Umeno, T. Kitamura, Modelling Simul. Mater. Sci. Eng. **14**, 1031 (2006)
54. T. Dumitrica, M. Hua, B.I. Yakobson, Proc. Natl. Acad. Sci. U.S.A. **103**, 6105 (2006)
55. S. Ogata, Y. Shibutani, Phys. Rev. B **68**, 165409 (2003)
56. T. Dumitrica, B.I. Yakobson, Appl. Phys. Lett. **84**, 2775 (2004)
57. J. Sun, L. He, Y.C. Lo, T. Xu, H. Bi, L. Sun, Z. Zhang, S.X. Mao, J. Li, Nature Mater. **13**, 1007 (2014)
58. Z.Y. Wang, K.H. Su, H.Q. Fan, L.D. Hu, X. Wang, Y.L. Li, Z.Y. Wen, Comp. Mater. Sci. **40**, 537 (2007)
59. Z. Wang, D. Liu, K. Su, H. Fan, Y. Li, Z. Wen, Chem. Phys. **331**, 309 (2007)
60. Z.Y. Wang, K.H. Su, H.Q. Fan, Y.L. Li, Z.Y. Wen, Mol. Phys. **106**, 703 (2008)
61. Z.Y. Wang, K.H. Su, X.P. Yao, Y.L. Li, F. Wang, Mater. Chem. Phys. **119**, 406 (2010)
62. T. Kitamura, Y. Umeno, A. Kushima, Mater. Sci. Forum **482**, 25 (2005)

Chapter 4
Strain Engineering on Nanosemiconductors

Abstract A lot of scientific investigations have been dedicated to strain engineering on semiconductors to alter and tune band gap energies and electronic structures by the application of strain. It also attracts much attention with broad interests in industrial applications. After starting with a brief coverage over strain engineering on bulk semiconductors, we review studies of strain effect on electronic structures around the Fermi energy in two-, one-, and zero-dimensional nanostructures, including metallic-semiconducting transition in carbon nanotubes.

Keywords Strain engineering · Semiconductor · Band gap energy

4.1 Strain Engineering on Semiconductors

It is well known that various properties of materials can be tuned when they undergo deformation or strain due to change in their electronic structures. In other words, a number of materials exhibit strong coupling between the physicochemical and mechanical properties. While the coupling may arise as a negative factor deteriorating the functional aspects of materials, there have been extensive trials to utilize the effect and realize devices with improved properties, as is often referred to as "strain engineering."

The coupling with deformation or strain is known to exist with a wide range of properties, e.g., electronic, magnetic, and optical properties. The recent trend of scientific investigation is reviewed, for example, in Ref. [1]. Among such fields, strain engineering on semiconductor devices has enjoyed abundant scientific investigation and is seen to be most promising for industrial applications. In the following sections, we review strain engineering on nanosemiconductors, while effect of strain on other material properties is covered in other chapters.

Y. Umeno et al., *Multiphysics in Nanostructures*, Nanostructure Science and Technology, DOI 10.1007/978-4-431-56573-4_4

4.2 Bulk Semiconductors

4.2.1 Bulk Semiconductors Subjected to External Strain

Figure 4.1 presents the energy (electronic) band structure of Si, which is most widely used for semiconductor devices [2]. The dashed and solid lines in Fig. 4.1a represent the calculation results by LDA and GWA, respectively. Comparing the GWA results to the LDA, the valence band energies are shifted to lower energies by a small amount. In contrast, the conduction band states rise in energy nearly rigidly, resulting in the band gap energy of 1.31 eV by GWA, while the prediction with LDA is 0.56 eV. The GWA result shows a good agreement with the experimental value of 1.17 eV, while LDA tends to underestimate the band gap energy as explained in Chap. 2. A more detailed comparison between the GWA result and experiment is shown in Fig. 4.1b. The experimental data have been obtained by photoemission and inverse photoemission. As shown in Fig. 4.1b, the GWA band structure results are in a good agreement with experimental data.

As for the so-called beyond-DFT methods, calculations with the DFT-HF hybrid functional have also been performed. Figure 4.2 compares band gap energies of various materials calculated with the Perdew-Burke-Ernzerhof (PBE) functional within the GWA approach and two DFT-HF hybrid functionals: the Heyd-Scuseria-Ernzerhof (HSE) and long-range corrected (LC) functionals [3]. ω is a factor representing the magnitude of HF hybridization. For HSE and LC, we show the calculated results with the optimum ω values: $\omega = 0.1 a_B^{-1}$ (a_B denotes the Bohr radius) for HSE and $\omega = 0.2\ a_B^{-1}$ for LC. It is clearly shown that the calculated band

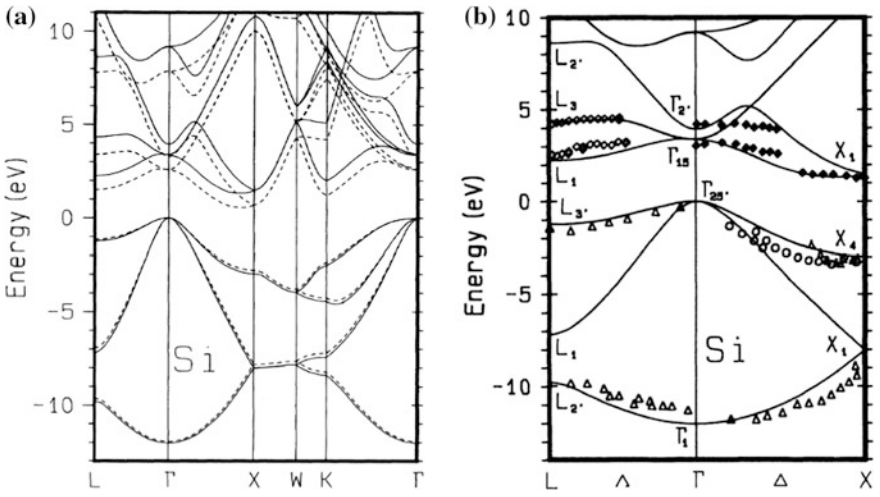

Fig. 4.1 Electronic band structures of Si. **a** LDA (*dashed lines*) and GWA (*solid lines*) results. **b** GWA results (*solid lines*) and experimental data (*symbols*). Reprinted with permission from Ref. [2]. Copyright 1993 by American Physical Society

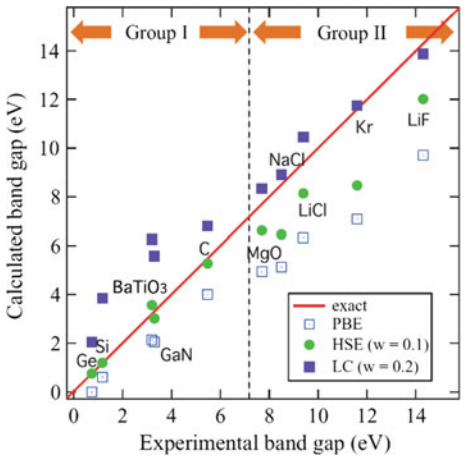

Fig. 4.2 Calculated band gaps obtained from different exchange–correlation functionals: PBE (*blank squares*), HSE (*green dots*), and LC (*purple squares*), plotted against experimental band gaps. Reprinted with permission from Ref. [3]. Copyright 2011 by American Physical Society

gaps by hybrid functionals, HSE and LC, are in a better agreement with the experimental values than the PBE, indicating the promising possibility of the hybrid functionals. The appropriate choice of ω with each exchange–correlation functional provides reasonable agreement with the calculated band gaps for a rather wide range of materials: HSE with $\omega \sim 0.1a_B^{-1}$ for materials with band gaps smaller than 7 eV, and LC with $\omega \sim 0.20.1a_B^{-1}$ for materials with band gaps larger than 7 eV. This information certainly provides a practical recipe to obtain reliable band gaps for various materials [3].

As was explained in Chap. 2 and shown in Figs. 4.1 and 4.2, it is necessary to employ beyond-DFT methods for quantitative evaluation of the energy band structure in semiconductor materials. Nevertheless, calculations using the LDA or GGA method can serve for qualitative discussion as exemplified below. Figure 4.3 shows the energy band structure of Si at [110] uniaxial strains of 0 and 2.5% calculated by LDA and GWA [4]. For uniaxial tensile strain, the sixfold degenerate conduction band minimum splits into two valleys (Δ_2) with lower energy along the Γ–Z direction and four others (Δ_4) with higher energy along the Γ–X direction. The degeneration is seen in both the LDA and GWA calculations, indicating that the effect of strain on the energy band structure can be qualitatively grabbed by LDA. This is also supported by the relationship between band gap energy and strain as shown in Fig. 4.4, which presents the change in band gap energy of Si in [110] uniaxial tension and $[110] - [1\bar{1}0]$ biaxial tension [4]. It is found that LDA can qualitatively reproduce the reduction (and reduction rate) of band gap energy by uniaxial and biaxial tensions. In both uniaxial and biaxial cases, the band gap decreases with strain, but the splitting between the Δ_2 and Δ_4 valleys is more pronounced for uniaxial strain. Furthermore, for uniaxial strain, the splitting grows with increasing strain, which reduces the scattering rate because only the two lower occupied Δ_2 valleys affect the electron mobility. In the case of biaxial strain, the splitting between the Δ_4 and Δ_2 valleys is much smaller than that for uniaxial strain, and the energetic ordering in fact changes for strain above $\varepsilon_{\text{biax}} = 3.25\%$, where the

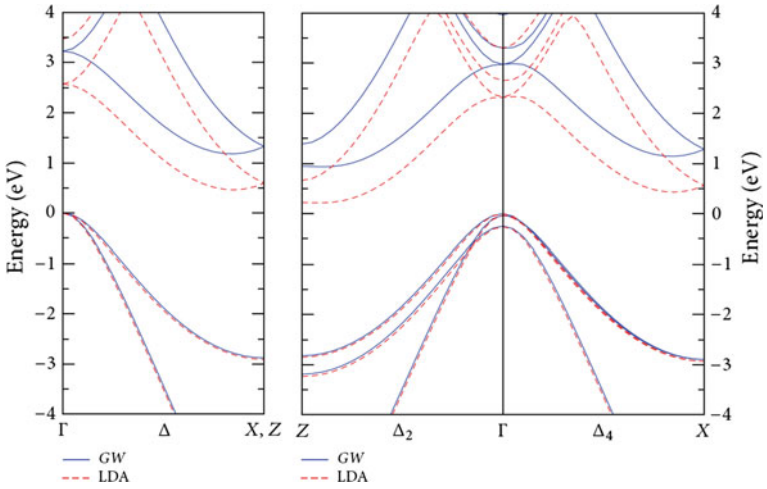

Fig. 4.3 Energy band structures of Si within LDA (*dashed lines*) and GWA (*solid lines*) without strain (*left*) and 2.5% uniaxial strain in the [110] direction (*right*) [4]

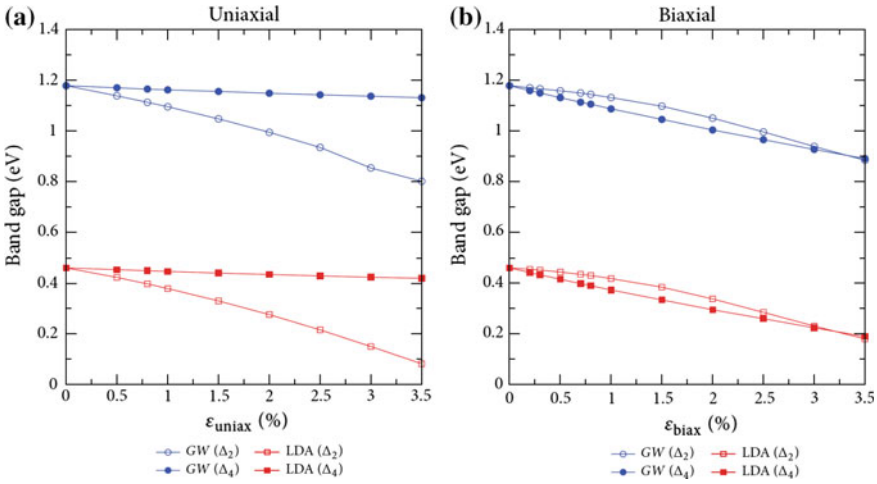

Fig. 4.4 Band gap energy of Si as a function of **a** [110] uniaxial and **b** $[110] - [1\bar{1}0]$ biaxial strains [4]

Δ_2 valleys are lower in energy. This regime of large tensile biaxial strain in the (110) plane might hence be interesting for an enhancement of the electron mobility.

The energy band structure can alter not only by tensile strain but also by shear strain. Change in density of states of Si during $\{111\}\langle 110\rangle$ shear was calculated by LDA [5]. The band gap energy decreases with increasing shear strain and vanishes at a strain of 0.2. It should be noted that the band gap vanishes before reaching the

critical strain of 0.3 for dislocation nucleation. This suggests that the electrical deterioration may appear before the disorder of atomic configuration introduced by the dislocation nucleation.

In this section, we exemplified that LDA or GGA calculations are sufficient for qualitative evaluation of change in band structures with respect to strain in semiconductors although "beyond-DFT" schemes are required for quantitative discussion. In the following sections in this chapter, we will introduce calculation results with LDA or GGA unless otherwise stated.

4.2.2 Bulk Semiconductors with Internal Strain Fields

The change in energy band structures shown in the previous subsection was one due to external strain, which is induced by, for example, constraint by a substrate or application of deformation by external load. On the other hand, band structures can change due to atomistic or electronic strain that occurs inside materials. Lattice defects are typical structures that cause internal strain and are closely related to the electronic properties in semiconductor materials. Thus, the behavior of lattice defects in semiconductors has been investigated in terms of experiments, theoretical approaches, and numerical simulations for decades. The history is compiled in detail in textbooks such as Refs. [6–8]. This section gives a brief overview of the effect of lattice defects, namely the vacancy (zero dimension), the dislocation (one), and the grain boundary (two), on the energy band structure of Si.

A LCAO model explaining the change in the energy band structure of Si due to a vacancy is schematically shown in Fig. 4.5. When a Si atom is removed from the lattice, four dangling bonds, directed toward the center of the vacancy, are formed. These dangling bonds hybridize so that their totally symmetric combination, an *s*-type state, lies in energy within the bulk valence band. Three different combinations with nodal planes are *p* type, and they form deep levels in the energy gap. In the doubly positive charge state of the defect, the deep levels are empty and the ionic relaxation preserves the T_d-point symmetry of the perfect lattice. In the singly positive and neutral charge states, one of the deep levels is occupied by one and two electrons, respectively. The Jahn-Teller effect lowers the point symmetry to D_{2d} and breaks the degeneracy of the deep levels. This symmetry is deduced from the EPR (electron paramagnetic resonance) measurement. Actually, the introduction of the second electron to the deep level results in a strong ionic relaxation and a lowering of the total energy, overcoming the coulombic repulsion between the localized electrons. Therefore, it is energetically favorable that the charge state changes directly from the doubly positive to the neutral one. This is the famous negative-effective-U effect first predicted for the Si vacancy by Baraff et al. and experimentally confirmed by Watkins and Troxel. For the negative and doubly negative charge states, the symmetry is further lowered and the point symmetry derived from the EPR and ENDOR (electron-nuclear double resonance) measurements is C_{2v} [9].

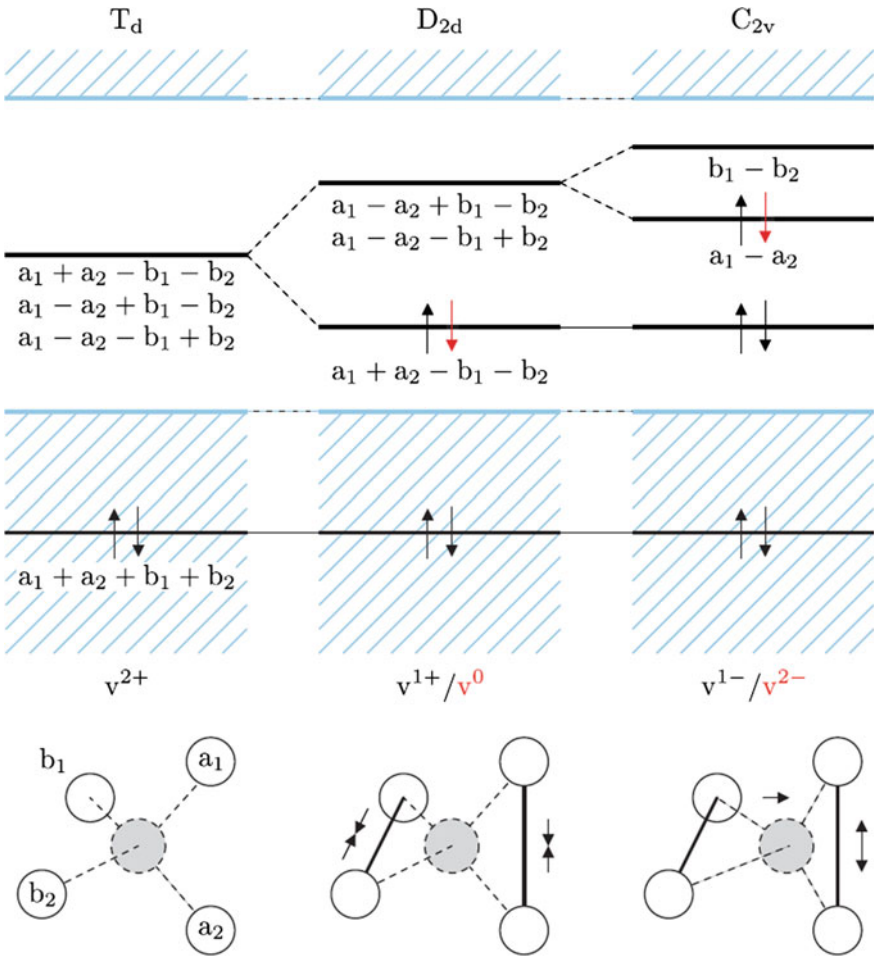

Fig. 4.5 Watkin's LCAO model for Si vacancy. Reprinted with permission from Ref. [12]. Copyright 2011 by American Physical Society

There have been a number of first-principles calculations for Si containing vacancies. While the formation energy of a vacancy, structural symmetry, volume, and ionization energy have been investigated, the calculation results have large scatter because of the difference in calculation conditions. As was explained in Chap. 2, a calculation of a material with a defect requires a sufficiently large simulation cell. We need to minimize the interaction between the defect and its own periodic images. For insufficiently large supercells, there will be an appreciable overlap between the defect and its own images, resulting in an error in the overall charge density of the system, and hence the total energy and the forces on the atoms. The obvious solution to this is to repeat the defect formation energy calculation in different sized supercells, using an equivalent sized basis set (e.g., same

plane wave cutoff energy) and same Brillouin-zone sampling density [10]. Calculation results may be influenced not only by the cell size but also by the cell shape. It was clarified that vacancy formation energy of Si converges against cell size faster in a bcc supercell than in a sc or fcc supercell [10]. The best way to improve a calculation is not just to increase the supercell size, but to do so in an appropriate manner bearing in mind the interaction of the supercell symmetry with the defect. Besides, the exchange–correlation functional (e.g., LDA vs. GGA) and k-point sampling (e.g., Γ-only vs. Monkhorst-Pack) should also been cautiously determined [9–13].

The effect of dislocations on the energy band structure in Si has been examined by first-principles and tight-binding calculations, using a dipole model (a periodic simulation cell containing two dislocations having Burgers vectors with opposite signs) and a cluster model (a simulation cell containing dislocations with its boundaries terminated by hydrogen atoms) [14–16]. According to the calculations, only shallow gap states appear and no deep midgap states emerge in 90° partial dislocations. On the other hand, the main effect of ⟨110⟩ edge dislocations is to push the top of the valence band into the gap of pure bulk Si, and this effect is enhanced as the dipole separation increases. At the Γ point, the shift in the top of the valence band changes from 0.01 eV at the one-ring separation to 0.18 eV for the five-ring separation. The enhancement of the band-edge shifts correlates well with the increasing localization of the structural distortions in the dislocation cores. In the core, the atoms retain tetrahedral coordination, so this effect is not due to the presence of dangling bonds, but instead appears to be associated with the compressed bonds in the core [16].

Calculations of the electronic structure of Si having a tilt or twist grain boundary have been performed, mainly with the tight-binding model [17]. It has been shown that coincidence tilt boundaries frequently observed in Si or Ge are stably constructed by the periodic arrangement of the structural units without any dangling bonds or large bond distortions. Thus, such boundaries have relatively small boundary energies and have no deep or shallow electronic states inside the minimum band gap. These features should be caused by the essential nature of covalent materials that fourfold coordinated configurations tend to be reconstructed. In Ref. [17], Kohyama performed a detailed analysis of the electronic structure of various twist boundaries in Si using the transferable TB method. They obtained the following general relation between the electronic structure and each kind of local structural disorder at grain boundaries, which should be common to amorphous Si. Odd-membered rings induce the changes of shapes of the local density of states. Bond distortions generate states at the top of the valence band and at the bottom of the conduction band, and greatly stretched bonds generate so-called weak bond states inside the band gap. Three-coordinated defects generate deep states, of which the wave functions are sharply localized at the defect sites. Five-coordinated defects generate both deep and shallow states. The deep state is localized at the neighboring atoms except the five-coordinated site, and the shallow states exist among the five-coordinated atom and neighboring atoms.

4.3 Nanosemiconductors

4.3.1 Two-Dimensional Nanosemiconductors

A main feature of the two-dimensional nanosemiconductor is the large fraction of surface area with respect to the entire volume. The energy band structure of two-dimensional nanosemiconductors is therefore closely related to their surface structure and its electronic state. First-principles calculations of surface electronic structure are usually performed with a slab model (e.g., Fig. 3.4b in Chap. 3) with one surface terminated by hydrogen atoms and fixed, while the other surface and a few subsurface layers are relaxed. In this subsection, we introduce first-principles calculations of the electronic state of Si surfaces [18–22].

Figure 4.6 shows the energy band structure of the $p(2 \times 1)$ asym, $p(2 \times 2)$, and $c(4 \times 2)$ reconstructions of the Si(001) surface. We illustrate that, for the $p(2 \times 2)$ and $c(4 \times 2)$ reconstructions, the dispersion curves of the occupied dangling bonds are similar to the folded branches of the $p(2 \times 1)$ surface, while significant changes

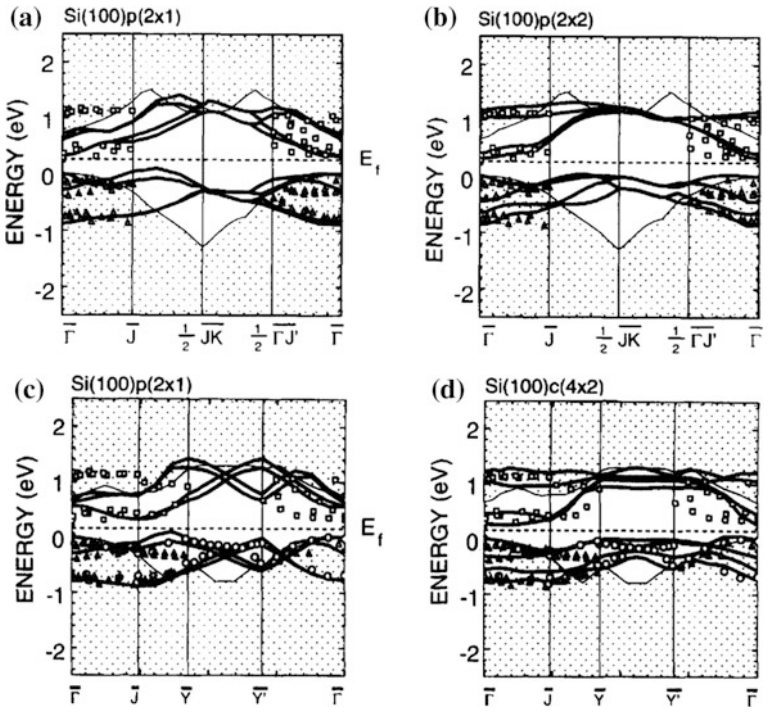

Fig. 4.6 Dispersion of dangling bond states of **a** $p(2 \times 1)$ (folded in accordance with $p(2 \times 2)$ symmetry), **b** $p(2 \times 2)$, **c** $p(2 \times 1)$ (folded in accordance with $c(4 \times 2)$ symmetry), and **d** $c(4 \times 2)$ surface configurations. Reprinted from Ref. [19], Copyright 1995, with permission from Elsevier

occur for the unoccupied surface states. We find a much better agreement of the calculated dispersion curves for the $p(2 \times 2)$ and $c(4 \times 2)$ phases with the photoemission spectra than for Si(001) $p(2 \times 1)$. In particular, the experimentally observed splitting of about 0.7 eV between the two empty surface states with an almost flat dispersion in the $\Gamma - J$ direction is well reproduced by our calculations for the $p(2 \times 2)$ and $c(4 \times 2)$ structures, while the $p(2 \times 1)$ geometry cannot explain the dispersion of these states. Moreover, the bandwidths of both the filled and empty surface states are reduced by about 0.1 eV in the $p(2 \times 2)$ structure and by about 0.2 eV in the $c(4 \times 2)$ structure with respect to the $p(2 \times 1)$ arrangement. The best agreement with the measured bandwidths of 0.70 eV for the occupied surface states and 0.80 eV for the empty surface bands is achieved for the $c(4 \times 2)$ phase, for which we obtain 0.79 and 0.93 eV, respectively. The good agreement of the calculated band structure of the $p(2 \times 2)$ and $c(4 \times 2)$ geometries with the experimental data indicates a short-range correlation of the dimers in Si(001) with at least a local $c(4 \times 2)$ or $p(2 \times 2)$ symmetry at room temperature.

The branches of the folded dangling bond states of the $p(2 \times 1)$ surface lying above the Fermi level are split by about 0.7 eV along the $\Gamma - \mathrm{J}$ direction in the $p(2 \times 2)$ as well as in the $c(4 \times 2)$ reconstruction, because of the correlation of the dimers along a row. As the branches are almost flat along the Γ–J direction, we conclude from our results that the dimer correlation between neighboring rows is weak compared to the coupling along the rows.

The results for the energy gain with respect to the unrelaxed (111) surface are listed in Table 4.1 for the three basic reconstructions 2×1 (p-bonded chain model), $c(2 \times 8)$ (adatom model), and 7×7 (DAS model) of the C, Si, and Ge(111) surfaces [22]. The results agree with the experimental findings. The most stable structures among the considered ones are the 2×1 chain reconstruction for diamond, the 7×7 translational symmetry of the complicated dimer-adatom-stacking fault surface with corner holes for silicon, and the $c(2 \times 8)$ adatom surface for germanium.

Figure 4.7 shows the energy band structure of the (111) surface [22]. In the 2×1 case, the bands are presented for the unbuckled chain reconstruction (C) and the chain-left isomer (Si, Ge). One observes a clear chemical trend of the positions of the π- and π*-chain bands in the fundamental gap with respect to the bulk valence band maximum (VBM) as well as of the surface-state gap at J. Along the

Table 4.1 Calculated reconstruction-induced energy gain (in eV) per (1×1) surface unit cell for three group-IV semiconductors. Gains due to relaxation are given for comparison. Reprinted with permission from Ref. [22]. Copyright 2001 by American Physical Society

Reconstruction	C	Si	Ge
Relaxed	0.57	0.06	0.01
2×1	1.37	0.30	0.23
$c(2 \times 8)$	0.39	0.33	0.27
7×7	0.34	0.36	0.27

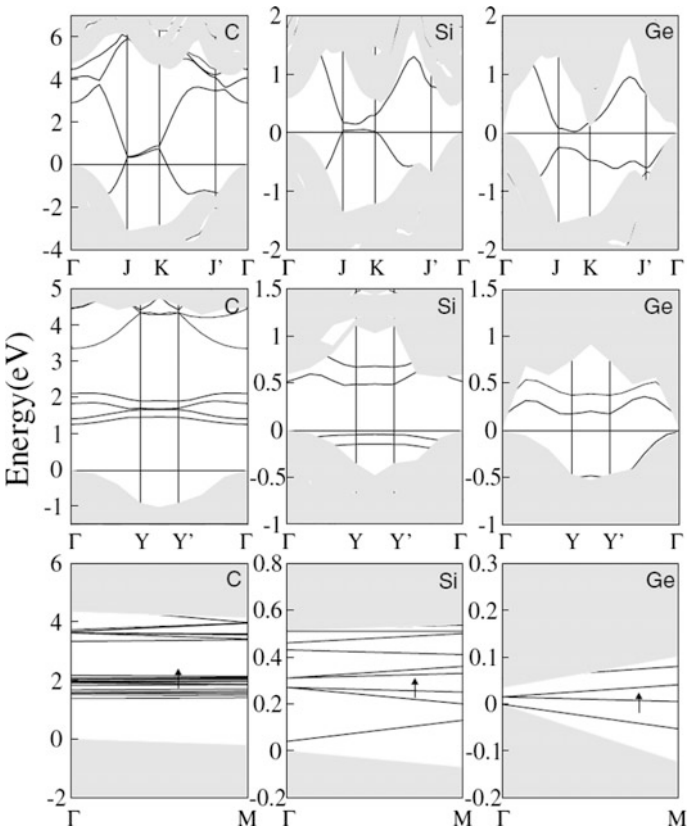

Fig. 4.7 Surface band structures versus high-symmetry directions in the corresponding two-dimensional Brillouin zone. The projected bulk band structures around the fundamental gap are presented as shaded areas. *Upper panels* (2×1) π-bonded chain model; *middle panels* $c(2 \times 8)$ adatom model; and *lower panels* (7×7) DAS model. In the latter case, the half-occupied band is indicated by an arrow. Reprinted with permission from Ref. [22]. Copyright 2001 by American Physical Society

row C -> Si -> Ge, the gap is opened as a consequence of the chain buckling. The occupied π bands are shifted below the VBM, indicating an energy gain due to the band structure contribution. Nevertheless, this gain does not cause an energetical favorization of Ge $(111) - c(2 \times 1)$ versus Ge $(111) - c(2 \times 8)$.

The stabilization of Ge $(111) - c(2 \times 8)$ becomes much clearer from the band structures in the middle panels of Fig. 4.7. Essentially, the dangling bonds belonging to the two adatoms and the two rest atoms appear in the fundamental gap region in the projected bulk band structure. The four bands are clearly observable for diamond because of the weak interaction of the dangling bonds and the similarities in the adatom and rest atom bonding to the underlying atomic layer. There is only a vanishing surface-state gap. In the silicon case, the adatom dangling bonds

become more p_z-like, whereas the rest atom dangling bonds increase the *s* character. As a consequence, the occupied surface bands belonging to the rest atoms are close to the VBM. In the Ge case, the occupied rest atom bands are shifted farther into the projected bulk valence bands. The accompanying energy gain via the band structure energy explains why the $c(2 \times 8)$ reconstruction is energetically more favorable than the π-bonded chain 2×1 reconstruction as well as why this happens in particular for germanium. The surface bands are clearly related to the geometry discussed above. Whereas the adatom structure is similar for C, Si, and Ge, there is an increase of the vertical distance of the rest atoms to the atomic layer beneath, 0.25, 0.47, and 0.50 d_{bulk}. It is accompanied by a dehybridization from four sp^3- to p_x-, p_y-, p_z-, and *s* orbitals and, hence, a downshift of the surface bands related to the occupied rest atom dangling bonds.

The lower panels in Fig. 4.7 give an idea about the stabilization of the 7×7 DAS surface in the case of the larger group-IV atoms, in particular Si, with respect to the diamond case. The dangling bonds of the rest atoms, adatoms, and corner-hole atoms give rise to many bands in the fundamental gap of the bulk band structure projected onto the small BZ of the 7×7 surface. The rest atoms dominate the occupied bands just below the Fermi level. It is pinned by the half-filled band at about 1.9, 0.3, and 0.02 eV, which is essentially formed by corner-hole states. According to their geometry, the rest of the atoms behave similarly as discussed for $c(2 \times 8)$. Adatom dangling bonds give remarkable contributions to the unoccupied surface bands, e.g., near the conduction band minimum of diamond. However, center adatoms also contribute to bands close to E_F. The stronger localization of C dangling bonds results in less dispersive bands, whereas the stronger surface band dispersion tends to smear out the Si and Ge gaps. One observes a clear chemical trend along the row C -> Si -> Ge of shifting the occupied surface bands closer to the VBM or below and, hence, stabilizing the 7×7 surface. This trend in particular follows our observation of the variation in the rest atom bonding.

Strain may change the surface structure and concomitantly the energy band structure of the surface. Figure 4.8 presents the effect of in-plane biaxial tension on the energy band structure for Si(001) surfaces with symmetric dimers (SD), asymmetric dimers (AD), and alternate arrangement of SD and AD (ADSD) [23]. The AD surface is semiconducting without strain (see Fig. 4.8a). The surface bands from *up* and *down* atoms of an asymmetric dimer are marked as D_{up} and D_{down}, respectively. The occupied D_{up} and unoccupied D_{down} bands are separated by about 0.3 eV of the surface band gap. When the strain is applied, the D_{up} band moves up, emerging out of the bulk valence band. The surface band gap becomes smaller and finally reaches zero as shown in Fig. 4.8b. In the transition state of 6.2% of tensile strain, the empty D_{down} band becomes occupied and electrons are transferred from the *up* atom to the *down* atom. The electron transfer decreases the ionicity and increases the covalency of the asymmetric dimer. It makes the asymmetric structure unstable and tends to symmetrize the dimer structures. However, the SD structure in the transition state is found to be still metastable. Calculated band structures of the SD structure in the $p(2 \times 1)$ phase without and with tensile strain are depicted in Fig. 4.8c, d respectively. Although the overlap of π and π^* bands is reduced as the

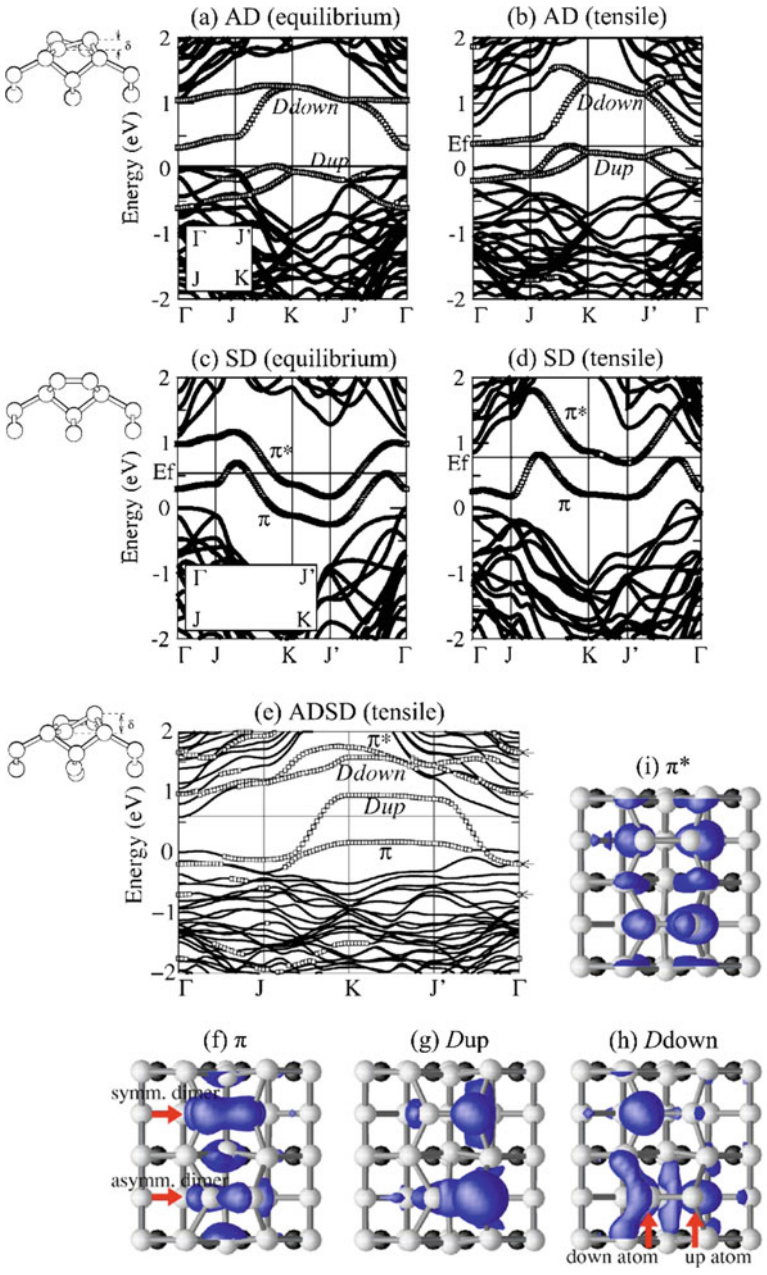

Fig. 4.8 Band structures for (**a**, **b**) $p(2 \times 2)$ AD, (**c**, **d**) $p(2 \times 1)$ SD, and (**e**) $p(2 \times 2)$ ADSD surface structures and (**f–i**) charge densities for the ADSD surface states. Reprinted with permission from Ref. [23]. Copyright 2007 by American Physical Society

tensile strain increases, the SD surface is still metallic with partial occupancy of π^* band.

Meanwhile, the ADSD structure becomes more stable than the AD or SD structure, in the ranges of 6.2–16% of biaxial strain. Figure 4.8e shows the band structure and the charge densities of the surface states for the $p(2 \times 2)$ ADSD structure on a 6.2% biaxially strained surface. The surface bands are distinguished from the bulk bands, and they are classified as π, π^*, D_{down}, and D_{up} bands. The π band is fully occupied, and it corresponds to the π-bonding state of a symmetric dimer, of which the charge density is shown in Fig. 4.8f for the Γ point. The wave function of the D_{up} band is related to the *up* atom of the asymmetric dimer, as shown in Fig. 4.8g. This band is partially occupied and overlaps with the bulk conduction bands. The D_{down} and π^* bands are fully unoccupied inside the conduction bands and correspond to the *down* atom of the asymmetric dimer and the π^* antibonding state of the symmetric dimer, respectively, as shown in Fig. 4.8h, i. It should be noted that, although the band structure of the ADSD structure shows a surface metallic character, the Fermi level is expected to be at the bulk conduction band minimum. There are surface holes in D_{up} states, while the electron carriers remain in the bulk conduction bands. The D_{down} and π^* surface bands are deep inside the conduction bands, and thus, they are fully unoccupied. The unoccupied D_{down} and π^* bands may induce the stabilization of the ADSD structure.

When the thickness of a two-dimensional nanosemiconductor becomes down to the nanometer level, interaction of two surfaces on the opposite sides of the film may change the energy band structure. In other words, the energy band structure of nanofilms is affected not only by the surface structure and strain but also by the film thickness. Change in band gap energy in a Si nanofilm having (001) $p(2 \times 1)$ asym. surfaces (Fig. 3.4 in Chap. 3) subjected to [110] uniaxial tension was examined in Ref. [24]. At the initial states, the band gap energy, E_{gap}, of the nanofilms is lower than that of the single crystal, and the thinner film has the lower gap. E_{gap} decreases with increasing tensile strain in all the models, and the strain at which the gap becomes null is higher in the thinner film. It is worth noting that the band gap in the thin film sustains up to higher tensile strain in spite of the lower band gap at the initial state than the bulk. This may be explained as follows: While the ideal (001) surface has no band gap, dimerization by 2×1 relaxation causes Peierls distortion that opens up a gap between π and π^*. As the film gets thinner, the two dimerized surfaces start to interfere with each other electronically, resulting in incomplete Peierls distortion and a smaller gap. The vanishing of the band gap under tension stands for the conduction band descending to reach the Fermi energy level. In consequence, the filling of the antibonding states takes place, which leads to weakening the stability of the bonding. This phenomenon, change in band configuration affecting the mechanical properties, has been discussed for other materials in some reports. In the case of silicon, it is expected that a discontinuity in the stress–strain relation would occur when the band gap vanishes because the occupancy of the conduction band suddenly becomes 1. However, it is not obvious in the present work and the stress–strain curves are rather smooth (see Fig. 3.5 in Chap. 3). In the case of a Si single crystal under shear, a smooth stress–strain

relation has been reported. In the calculation, the occupancy changes rather smoothly due to the smearing band occupancy around the Fermi level to avoid divergence in the self-consistent loop, which can cover up a slight cusp. In any case, a strong cusp in stress–strain curves is not expected.

The ultimate form of ultrathin films is the monolayer, which consists of a one atom-thick layer and comprises surfaces only. Graphene is a typical monolayer structure consisting of carbon atoms. Graphene under no strain possesses the energy band structure with cone-shaped valence and conduction bands adjoined with each other at the tip of the cones (Dirac's cone), and therefore becomes semiconducting or half-metallic with a vanishing band gap energy. Since the peculiar energy band structure originates in the sixfold symmetry of graphene, the band gap does not open under isotropic biaxial strain because the symmetry does not break and the Dirac point does not move [25, 26]. On the other hand, the sixfold symmetry breaks under uniaxial strain, resulting in the elimination of one of the two Dirac points that would be located at K and K' points with no strain. Nevertheless, the other Dirac point changes its location only, making the band gap remain closed [26–28]. As explained above, neither biaxial nor uniaxial strain can open the band gap in graphene. However, recent first-principles calculations revealed that the band gap can be formed in graphene by a devised application of strain. For example [29], it was reported that the combination of two normal strain components in biaxial strain changes the band gap energy and that a gap up to about 1.0 eV can be achieved when tensile and compressive strains are applied to the y and x axes, respectively. Another calculation reported that the band gap energy can be altered by the magnitude and direction of local strain [30].

4.3.2 One-Dimensional Nanosemiconductors

Nanowire—effect of shape on semiconducting property

The energy band structure of one-dimensional nanosemiconductors is influenced not only by the surface but also by the junction of surfaces (edge). The smaller the radius of a one-dimensional nanosemiconductor (e.g. nanowire) is, the more the ratio of the edge length to the entire volume becomes. When the radius is down to several nanometers, the effect of edges appears clearly. For example, as was explained in Sect. 3.3.2, the $\{100\}$ surface of a Si$\langle 100 \rangle$ nanowire with a radius of several nanometers becomes $c(2 \times 2)$, which is energetically higher than (2×1) and $c(4 \times 2)$, due to the presence of edges [31]. This is because the energy gain by eliminating dangling bonds at edges of $c(2 \times 2)$ surfaces exceeds that by surface reconstruction to (2×1) or $c(4 \times 2)$. Here, we introduce first-principles calculations about the effect of edges on the energy band structure in Si nanowires [31, 32].

Figure 4.9 shows eigenenergies of electronic states in Si$\langle 100 \rangle$ nanowire with the $\{100\}$ surfaces of (2×1) and $c(2 \times 2)$ [31]. Comparing the electronic structures of

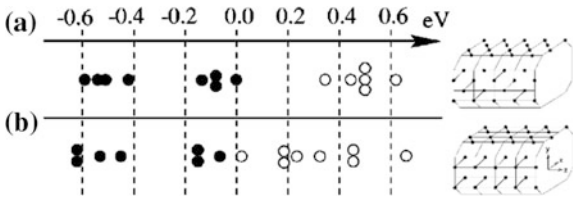

Fig. 4.9 Eigenenergies of the electronic states for **a** 2×1 and **b** $c(2 \times 2)$ reconstructions. Filled and empty states are denoted by filled and empty circles, respectively. The top and bottom right schematics depict Si⟨100⟩ nanowires with 2×1 and $c(2 \times 2)$ reconstructions, respectively. Reprinted with permission from Ref. [31]. Copyright 2001 by American Physical Society

the two reconstructions, it was found that the 2×1 configuration has a gap of 0.35 eV across the Fermi level. Furthermore, the topmost filled states consist of four nearly degenerate states, which are localized on the four edges of the bars, and this cluster is separated from states below it in energy by a gap of 0.30 eV. This nearly symmetric placement with sizeable gaps as well as the spatial localization of the edge states leads us to conclude that the 2×1 bar is insulating. On the other hand, the electronic structure of the $c(2 \times 2)$ configuration is more subtle. The states in the vicinity of the Fermi level are localized on the surfaces of the bars, and the gap across the Fermi level is only 0.09 eV. It should therefore be reasonable to believe the $c(2 \times 2)$ configuration is a small-gap semiconductor or even perhaps metallic.

The energy band structure of a Si⟨110⟩ nanowire in a hexagonal prism shape with four {111} and two {110} surfaces was also calculated [32]. Unlike the previous case of the Si⟨100⟩ nanowire with {100} surfaces, the neighboring facets can match through a smooth angle and they are not significantly distorted with respect to their homologous infinite surfaces. Hence, the {111} facets that dominate the reconstruction follow the π-bonded chain model of Si{111} surfaces, conferring to the wire a semiconducting character. The band gaps are found to be smaller than the bulk value as an effect of the formation of the π-surface states.

It is known that, even without reconstruction of surfaces or edges, the energy band structure of a nanowire can be changed when its thickness becomes very small. For example, it was found that the band gap energy of a Si nanowire with H-terminated surfaces and edges having no reconstruction increases with decreasing diameter due to the quantum confinement effect [33–35]. Change in the energy band structure of nanowires subject to axial strain was confirmed by first-principles calculations [36–41]. In what follows we introduce examples of Si⟨110⟩ and ⟨111⟩ nanowires.

Figure 4.10 shows results of first-principles calculations of the energy band structure in Si⟨110⟩ and ⟨111⟩ nanowires with a diameter of about 2 nm under axial strain [38]. We start with a discussion of the band structure of these unstrained nanowires. The general features of the conduction bands can be understood from effective mass theory and the band structure of bulk Si. The conduction band edge of bulk Si consists of six equivalent anisotropic Δ valleys, located at about

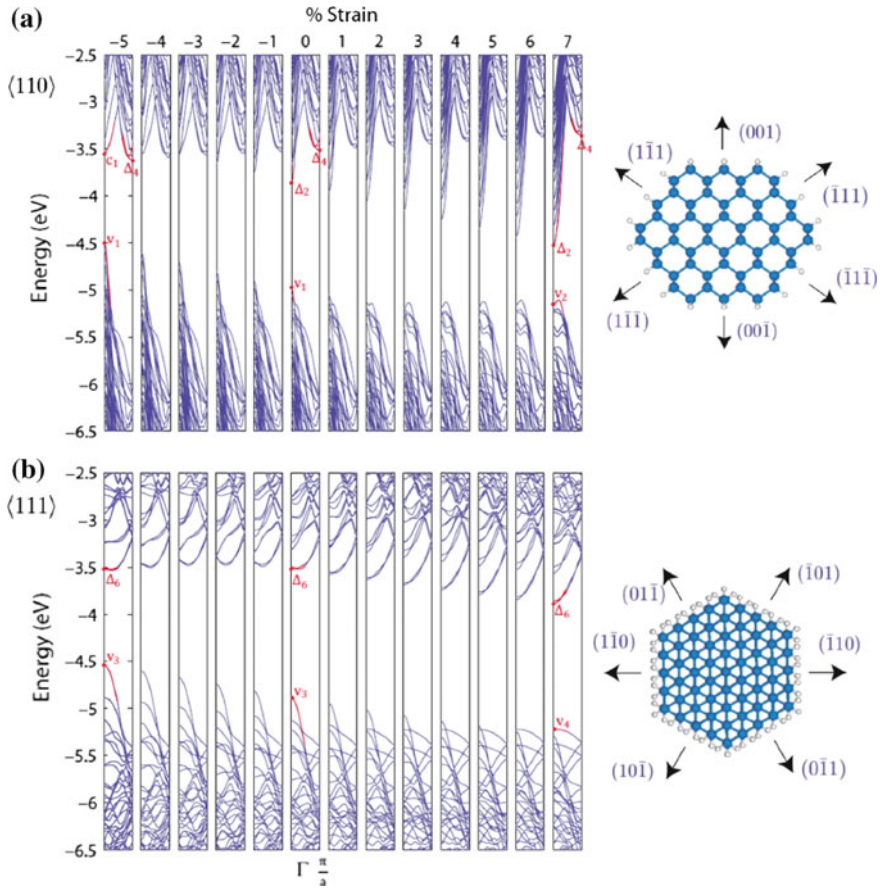

Fig. 4.10 Change in the energy band structures of Si **a** ⟨110⟩ and **b** ⟨111⟩ nanowires under uniaxial strain. The top and bottom right schematics depict the cross section of the nanowires. Reprinted with permission from Ref. [38]. Copyright 2008 by American Physical Society

$0.85 \times 2\pi/a$ along the ± [100], ± [010], and ± [001] directions. Further in the text, the six valleys are referred to as [± a, 0, 0], [0, ± a, 0], and [0, 0, ± a].

For the ⟨110⟩ nanowire, the conduction band minimum at Γ is formed from the two Δ valleys [0, 0, ± a] of bulk Si and is thus labeled Δ_2. Another minimum Δ_4 is formed from the bulk Si valleys at [± a, 0, 0] and [0, ± a, 0] and folded close to ± $0.5\pi/a$. According to this effective mass picture, ⟨110⟩ wires with reasonably circular cross sections (no high aspect ratio) should be direct band gap. Indeed, the conduction minimum Δ_2 at the Γ point has lower confinement energy than the conduction band minimum Δ_4. This is because the energy of Δ_2 is determined by the large longitudinal mass and smaller transverse mass in the confinement directions, while the energy of Δ_4 is determined by the smaller transverse mass. Similar analysis can be applied to nanowires grown in other directions. Thus, the

conduction band minimum of ⟨111⟩ nanowires is formed from the Δ valleys [± a, 0, 0], [0, ± a, 0], and [0, 0, ± a] of bulk Si and labeled Δ_6. The minimum is located away from the Γ point, and the wire is, in general, indirect band gap. The large effective mass in the growth direction results in a conduction band which is very flat. The conduction band is also very sensitive to the morphology of the nanowire, and the deviation of wire cross section from circular geometry lifts the degeneracy of the Δ_6 minimum. At last, the valence band maxima of the ⟨110⟩ and ⟨111⟩ nanowires, positioned near the Γ point, are similar to those of bulk Si. There is, however, a significant mixing and splitting of the valence band, which cannot be explained by effective mass theory.

When the ⟨110⟩ nanowire is compressed, the energy of Δ_2 increases and eventually exceeds the energy of the Δ_4 state. We will refer to this transition as a direct-to-indirect band gap transition. Under tensile strain, the Δ_2 state is pulled down together with a subset of higher subbands at the Γ point. Remarkably, Δ_4 and another energetically close state at Γ, which we call c_1, remain fairly constant under tension or compression. At a tensile strain of 7%, the c_1 state is the 14th subband at Γ, while it is the lowest and the second lowest at a strain of −7 and 0%, respectively. We note that the situation is somewhat different in bulk Si under uniaxial tensile stress in the ⟨110⟩ direction: While the splitting of Δ_2 and Δ_4 has the same sign, it is the Δ_4 valleys that are raised, while Δ_2 valleys are less affected. The valence band edge of ⟨110⟩ nanowires is determined by the states v_1 and v_2. The v_1 state is the highest at zero strain and is shifted significantly with strain, rising under compression and falling under tension, while the v_2 state remains fairly constant. Therefore, under the tensile strain, when the v_1 state is pulled down, the nanowire becomes indirect band gap and the hole properties are determined by the v_2 state.

Applying strain to the ⟨111⟩ nanowire tended to lower the energy of Δ_6 under tension and raise it slightly under compression. As expected in the effective mass picture, there is no splitting of the conduction band minimum because all valleys are affected equally by strain. Consequently, the conduction band behavior under strain in ⟨111⟩ Si nanowires is analogous to that of bulk Si under uniaxial strain in the ⟨111⟩ direction. The valence band behavior of ⟨111⟩ nanowires is very similar to that of ⟨110⟩ nanowires: The valence band state v_3 was raised under compression and lowered under tension, while v_4 remained fairly constant under strain. The lowering of the v_3 state below the v_4 state occurs at much larger tensile strain than the direct-to-indirect band gap transition (v_1 below the $v2$) in the ⟨110⟩ nanowire.

The behavior of the band gap is a direct consequence of the change in band structure described above. At zero strain, ⟨111⟩ wires tend to have a larger band gap than ⟨110⟩ wires of the same size due to smaller bulk Si effective masses in the confinement directions. This trend continues to hold in strained wires. In ⟨110⟩ nanowires, compressive strain tended to first increase the band gap, due to rising of the Δ_2 state. However, the band gap increase is followed by a decrease, due to the v_1 state rising, while the bottom of conduction band Δ_4 remains fairly constant. Tensile strain causes a monotonous closing of the band gap, due to lowering of the Δ_2 state, while the top of the valence band v_2 remains constant. In strained ⟨111⟩ nanowires, the bottom of the conduction band Δ_6 remains almost nearly constant.

Thus, compressive strain tended to decrease the band gap, due to rising of the v_3 state, while tensile strain tends to keep the band gap constant, as the energies of Δ_6 and v_4 do not change significantly.

Nanotube and nanoribbon—more pronounced, peculiar effect of shape

The nanotube is another form of one-dimensional nanosemiconductors. Carbon nanotubes (CNTs), rolled-up structures of a graphene sheet, can be conducting (metallic) and semiconducting depending on the chirality vector (m, n), while unstrained graphene is a semimetal; i.e., conducting when $m - n = 3q$ and semiconducting when $m - n = 3q + 1$ or $m - n = 3q + 2$, where q is an integer. This rule, however, no longer holds under strain. In fact, first-principles and tight-binding calculations have revealed that the energy band structure of CNTs can shift from conducting to semiconducting (and vice versa) under axial [42–44], radial [45, 46], and torsional [47, 48] deformations. In what follows we introduce a study of systematic investigation about the relation between the energy band structure and axial tension in CNTs with various chiral vectors [44].

Figure 4.11 shows the band gap energy E_{gap} and axial tensile strain ε_z calculated by the tight-binding method, where the validity of the employed tight-binding potential was confirmed by comparison with first-principles calculation results. Here, armchair CNTs $(m - n = 0)$ are excluded because they remained conducting under any tensile strains. In CNTs with $m-n = 3q$ (Fig. 4.11a), E_{gap} is 0 at the initial state, $\varepsilon_z = 0$. When tensile strain is applied, a band gap appears in each CNT. The magnitude of tensile strain at which the band gap appears varies from 0.03 to 0.09, and as the tensile strain increases, E_{gap} temporarily increases before decreasing. E_{gap} becomes 0 again at high tensile strain. From the above results, we conclude that the CNTs change their electronic properties from metallic -> semiconducting -> metallic, as the strain increases. Moreover, the change appears at lower strain for CNTs with smaller diameters. In CNTs with $m-n = 3q + 1$ (Fig. 4.11b), the band gap at $\varepsilon_z = 0$ is positive and the tube with a smaller diameter has a larger band gap energy. E_{gap} increases in the region of low strain; however, as the strain increases, the magnitude of E_{gap} decreases, finally resulting in the band gap's elimination. A CNT with a larger initial band gap possesses higher critical strain where the gap disappears, and the CNTs become semiconducting again as the tensile strain increases. In CNTs with $m-n = 3q + 2$ (Fig. 4.11c), the property in $m - n = 3q + 2$ is similar to that in $m-n = 3q + 1$. The value of E_{gap} at $\varepsilon_z = 0$ is positive and its magnitude depends on the diameter. However, E_{gap} is reduced as the tensile strain increases and the band gap is eliminated. The transient strain is dependent on the diameter and is relatively small compared to that of $m - n = 3q + 1$. The tubes become semiconductors again under further tension.

Nanoribbons can also be characteristic one-dimensional nanosemiconductors. The energy band structure of a graphene nanoribbon (GNR) can vary depending on its structure, i.e., how it is cut out from a graphene sheet, just as a CNT can become metallic or semiconducting depending on how it is rolled up from a graphene sheet. For example, the energy band structure of a zigzag CNR (ZGNR), which is cut out

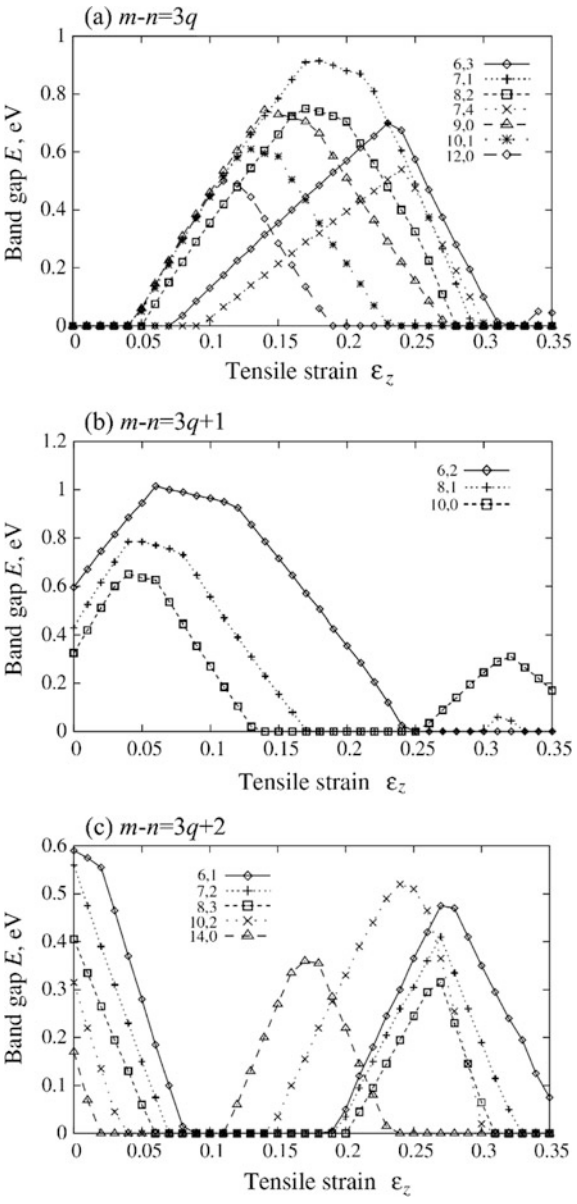

Fig. 4.11 Change in the band gap energy of (m, n) SWCNTs with the chirality of **a** $m-n = 3q$, **b** $m-n = 3q + 1$, and **c** $m-n = 3q + 2$. Reprinted from Ref. [44], Copyright 2004, with permission from Elsevier

from a graphene sheet in such a way that the y axis corresponds to the longitudinal direction, exhibits little change against strain [49, 50]. On the other hand, an armchair GNR (AGNR) with the x axis along the longitudinal direction shows significant change in the band gap energy according to strain [49–51]. In addition, the energy band structure of AGNRs can be classified into three types by the

number of atom layers in the width of the nanoribbons, W; i.e., $W = 3n$, $3n + 1$, $3n + 2$ (n is an integer) [52]. Thus far, first-principles calculations have been done for tensile and compressive strain along the longitudinal direction [49, 50], in-plane shear [50], antiplane bending [51], and longitudinal torsion [51] in AGNRs. In what follows we introduce calculation results for tensile and compressive strain along the longitudinal direction, ε [49].

The electronic structures for AGNR with width $W = 13$ under a series of uniaxial strains were first investigated, as shown in Fig. 4.12a, where $\varepsilon = -4.0$, -0.8, 0.0, 3.0, 7.3, and 10.0% were labeled as A, B, C, D, E, and F, respectively. They all exhibit direct band gaps at the Γ point. It is clear that the subband spacings and energy gap of the deformed AGNRs are tunable with uniaxial strain. The obvious difference among band structures of the deformed AGNRs under six different given ε values is the energy positions of two upmost valence subbands (v_1 and v_2) and two lowest conduction subbands (c_1 and c_2). Since the energy of a specific band depends on the C–C overlap integrals and the C–C bond geometries, it is expected that these subbands will respond to the applied uniaxial strains differently. Four subbands (v_1, v_2, c_1, and c_2) separate at the Γ point when the applied strain (ε) is -4.0% (case A). It was observed that two valence subbands (v_1 and v_2) and two conduction subbands (c_1 and c_2) nearly degenerate and the gap reaches the maximal energy gap of 1.0 eV when the compressive deformation is -0.8% (case B). The subbands separate gradually and the gap decreases almost linearly as the tensile strain increases (case C–D). When the strain increases continuously to 7.3% (case E), the subband separation between v_1 and c_1 decreases to 0.03 eV, which corresponds to the minimal energy gap case. When ε further increases to 10.0% (case F), four subbands will separate again.

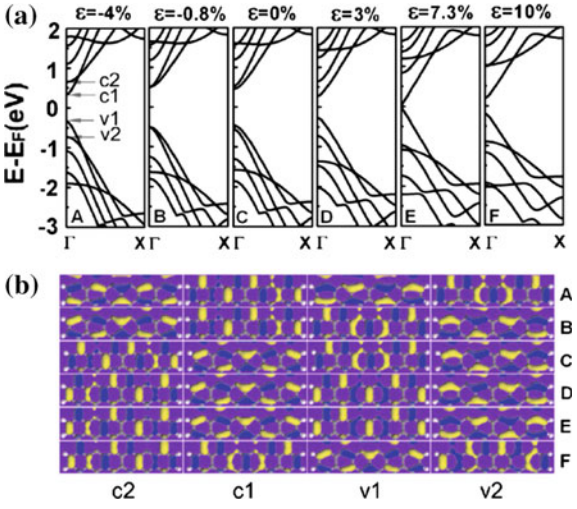

Fig. 4.12 **a** Change in the energy band structures of AGNRs with width $W = 13$ under uniaxial strain and **b** the spatial distributions and energy variations of these subbands (v_1, v_2, c_1, and c_2) near the Fermi surface of the deformed AGNRs with different strains. Reprinted with permission from Ref. [49]. Copyright 2008, American Institute of Physics

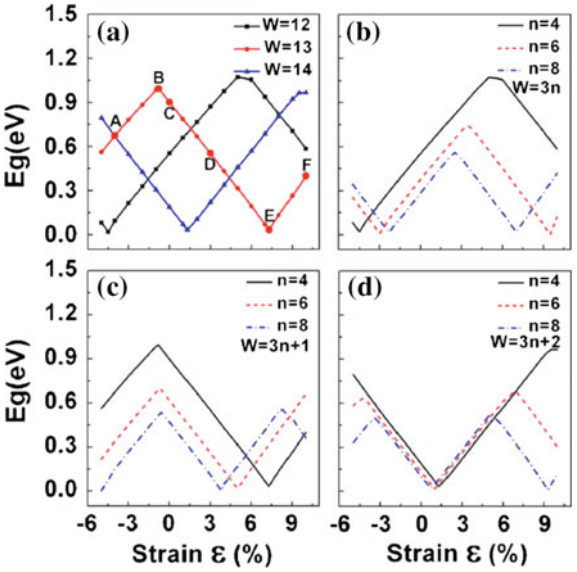

Fig. 4.13 Relationships between the band gap and uniaxial strain for the AGNRs with **a** $W = 12$, 13, and 14, **b** $W = 3n$, **c** $W = 3n + 1$, and **d** $W = 3n + 2$. Reprinted with permission from Ref. [49]. Copyright 2008, American Institute of Physics

Figure 4.13b presents the spatial distributions of wave functions literally corresponding to the four subbands (v_2, v_1, c_1, and c_2) under six different uniaxial strains at the Γ point. These subbands can be classified into two kinds of states: the vertical bonds along the periodic direction (denoted as VB) and the parallel bonds perpendicular to the periodic direction (denoted as PB). Two subbands c_1 and v_2 are VB states, two subbands c_2 and v_1 are PB states for the AGNRs with strain $\varepsilon = -4.0\%$ (case A), while the AGNRs is elongated ($\varepsilon = 3.0\%$, case D), c_2 and v_1 are tuned to be VB states, and c_1 and v_2 have PB features. Clearly, the spatial distributions of frontier subbands can be significantly manipulated by the external strain.

To present clearly the energy gap modification, the variations of energy gap of AGNRs with widths $W = 12$, 13, and 14 as a function of ε are shown in Fig. 4.13a with filled square, circle, and triangle symbol lines, respectively. The calculated maximal values of E_g for the AGNR with widths $W = 12$, 13, and 14 are 1.07, 1.00, and 0.97 eV appearing at $\varepsilon = 5.0$, -0.8, and 9.5%, respectively, while the minimal values of E_g are 0.02, 0.03, and 0.03 eV, which occur at $\varepsilon = -4.5$, 7.3, and 1.3%, respectively. The shapes of the calculated curves display a zigzag feature for three different ribbon widths, and the energy gaps change almost linearly between two neighboring turning points by changing the ε. It is clear that the energy gap is sensitive to the external strain (ε). This delicate electromechanical coupling of AGNRs suggests one interesting application of AGNRs as strain sensors.

The variations of the energy gaps of three family structures with different widths ($W = 3n$, $3n + 1$, and $3n + 2$, where $n = 4$, 5, and 6) as a function of ε are shown in Fig. 4.13b–d, respectively. Obviously, the following are several common features for all AGNRs. (1) The energy gap decreases when the width of AGNRs increases

without external strain. (2) The minimal energy gaps of all deformed AGNRs are of the order of several meV, while the maximal energy gaps are sensitive to the width of the deformed AGNRs and the values decrease as the width of AGNRs increases. For example, the maximal values of energy gaps are 1.07, 0.74, and 0.56 eV for the deformed AGNRs with width $W = 3n$ ($n = 4$, 6, and 8), which appear at $\varepsilon = 9.5$, 6.6, and 4.8%, respectively, as shown in Fig. 4.13b. (3) The zigzag feature here is universal for all deformed AGNRs. (4) Clearly, the separation of strains between two turning points becomes shorter when the width of AGNR increases, which suggests that the possibility to tune the energy gap becomes weaker for the wider AGNRs. (5) The previous TB studies reported that the uniaxial strain could lead to the semiconductor-metal transition for the AGNRs. Here, it should be pointed out that the exact semiconductor-to-metal transition is not achieved from DFT calculations by either elongating or compressing the AGNRs. To observe such a semiconductor-to-metal transition, the crossing of subbands v_1 and c_1 is required. However, the special geometry of AGNR and orbital interaction prevent this type of interband crossing.

As the width of a nanoribbon is reduced, ultimately it becomes the atomic chain structure, which has the potential to be the ultimate one-dimensional nanosemiconductor, as is represented by the carbyne: carbon atomic chain. Recently, first-principles calculations were reported for change in the band gap energy of carbynes under tension [53–56], torsion [53], and bending [57]. In what follows we introduce results for the electronic properties of carbynes under tension [54].

Atomic chain—ultimate one-dimensional semiconductor

The electronic structure of a cumulene chain (with double bonds throughout the chain, … = C=C = C=C = C=…) exhibits degenerate π bands crossing the Fermi energy (Fig. 4.14a), thus resulting in a quantum conductance of two quanta G_0 ($G_0 = 2e^2/h$) and a corresponding current of 15 μA at 0.1 eV bias (Fig. 4.14b). Within this highly symmetric configuration, the π orbitals are homogeneously distributed along the chain, forming equivalent bonds (Fig. 4.14e, top). As mentioned previously, such an ideal 1D system is subject to structural distortion. Indeed, Peierls theorem predicts that a 1D system of equally spaced atomic sites with one electron per atom is unstable. A lattice distortion will cause the electrons to be at a lower energy than they would be in a perfect crystal, thus inducing the well-known Peierls dimerization in the chain as can be observed from the plot of the electronic distribution of π orbitals shown in Fig. 4.14e. This new structure with bond-length alternation (Fig. 4.14e, middle), becomes energetically favorable and, due to the loss of symmetry, exhibits an electronic band gap (E_{Gap}). The LDA (GGA) ab initio electronic structure of this polyyne chain (with alternating single and triple bonds, …–C ≡ C–C ≡ C–C ≡ …) presents an E_{Gap} of 0.28 (0.34) eV (Fig. 4.14c). However, although the DFT calculations usually give an efficient and accurate description of ground state properties (total energy, lattice constants, atomic structures, phonon spectra, etc.), the Kohn-Sham band structure systematically underestimates the band gap (often by more than 50%). In order to address excited-state properties and to calculate the band structure including electron-

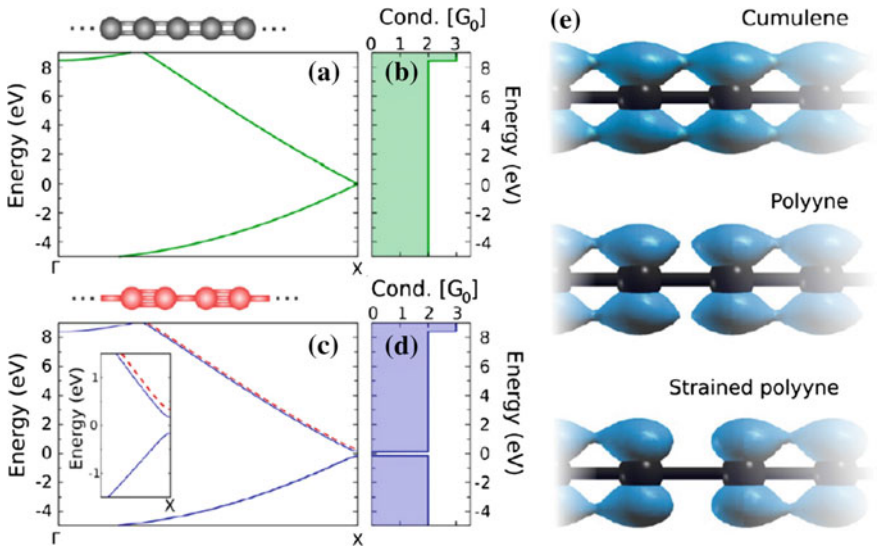

Fig. 4.14 Energy bands and quantum conductances of **a**, **b** cumulene and **c**, **d** polyyne carbon chains and **e** electronic distribution (π orbitals in *blue*) perpendicular to the chain. The *blue* and *red lines* in (c) illustrate the band structures of polyyne within DFT/LDA and GW, respectively. Reprinted with the permission from Ref. [54]. Copyright 2013 American Chemical Society

electron interactions (many body effects), MBPT calculations within the GW approximation for the self-energy are frequently used to provide corrected values for E_{Gap} in better agreement with experiments. The value of E_{Gap} for the polyyne chain estimated after the GW corrections to DFT/LDA is 0.407 eV (Fig. 4.14c, dotted red line). Such an absence of electronic states is also clearly visible in the quantum conductance of the polyyne chain (Fig. 4.14d).

Figure 4.15a illustrates the strain dependence of the electronic band gap (in both DFT and MBPT approaches) for an infinite polyyne chain. Note that 5% strain is enough to induce a 1 eV GW gap. In addition, as the strain increases, the dimerization (ΔL: difference of lengths between the triple and the single bonds in the polyyne chain) is more pronounced (Fig. 4.15b), thus stabilizing this specific configuration under strain. Indeed, the electronic distribution perpendicular to the strained polyyne chain depicts that π-orbitals are even more localized on the shorter bond length than for the pristine polyyne case (Fig. 4.14e, bottom).

A 1D system constituted by a polyyne segment embedded into two perfectly conducting semi-infinite cumulene chains (e.g., 10 carbon atoms as depicted in Fig. 4.15) is used as a realistic atomic model of a polyyne chain with seamless contacts. Such a model is used to highlight the effect of strain that could appear locally in the carbon chain. Indeed, a local constraint should stabilize a polyyne-type configuration locally (finite-size segment), thus perturbing transport along the chain. Recall that the DFT conductance does not reflect the predicted GW gap for an infinite polyyne chain under strain. To go beyond this DFT problem, a

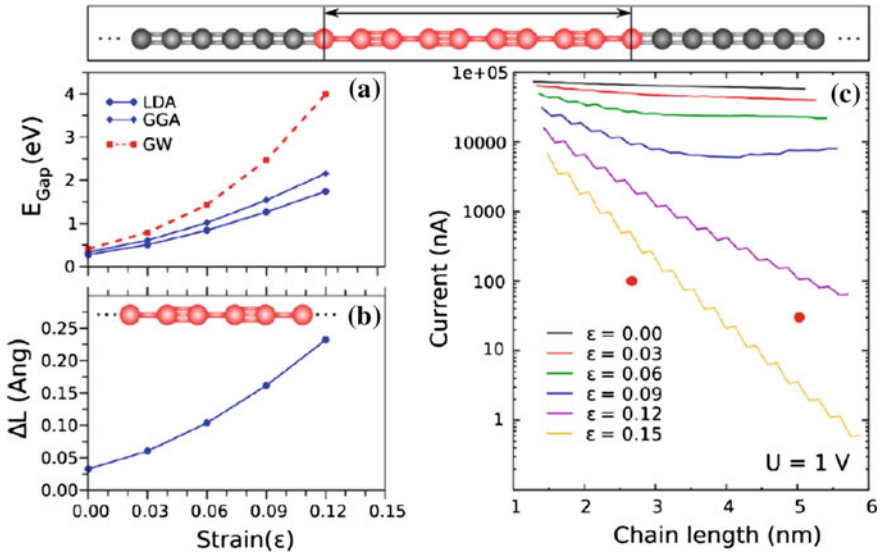

Fig. 4.15 Strain dependence of the **a** electronic band gap, **b** atomic structure, and **c** electron current in polyyne chains. In **b**, ΔL is the length of the longer bond minus length of the shorter one. In **c**, the two *red points* are experimental measurements of the current at different lengths of the chain. Reprinted with the permission from Ref. [54]. Copyright 2013 American Chemical Society

π–π^* distance dependent tight-binding approach that reproduces the GW band gaps for different strains was used. The quantum conductance for the polyyne segment embedded on semi-infinite cumulene leads was calculated for different chain lengths and strain. An estimation of the current flowing through these systems was obtained by the integration of the quantum conductance from the charge neutrality point to 1 eV considering the charge density at zero bias (Fig. 4.15c). The fluctuations of the current are due to the difference in conductance between an even and odd number of atoms in the atomic chain. The current is found to drop by several orders of magnitude for long chains when the strain increases from 0 to 15%.

4.3.3 Zero-Dimensional Nanosemiconductors

Nanoparticles and nanoclusters have the potential to be a zero-dimensional nanosemiconductor having a diameter of the nanometer order. Here, we differentiate the nanoparticle and the nanocluster depending on whether it has the bulk crystal structure. For example, the fullerene consisting of carbon atoms with the soccer-ball structure is defined as a nanocluster since the structure is different from bulk structures such as diamond and graphite. First-principles calculation studies for the strain effect on the electronic properties of zero-dimensional

nanosemiconductors are by far limited compared to those for two- or one-dimensional nanosemiconductors. In this subsection, we introduce recent studies for Si nanoparticles.

The band gap energy of Si nanoparticles is reduced as the diameter becomes smaller due to the quantum confinement effect [58]. In addition, first-principles calculations revealed that the band gap energy changes with hydrostatic [58, 59], biaxial [59], and shear [59] strains. While uniaxial, biaxial, and shear strains reduce the band gap energy [59, 60], the effect of hydrostatic strain differs depending on the particle diameter [58, 59]. The change of band gap as a function of strain in silicon is shown in Fig. 4.16a–d. One notices that the bulk indirect band gap increases almost linearly with increasing strain—this result is in good agreement with earlier studies and corresponds to a constant deformation potential parameter. In contrast, the strain dependence of the smallest cluster (Si_5H_{12}) shows the opposite dependence. Clusters which are 1–2 nm in size show intermediate

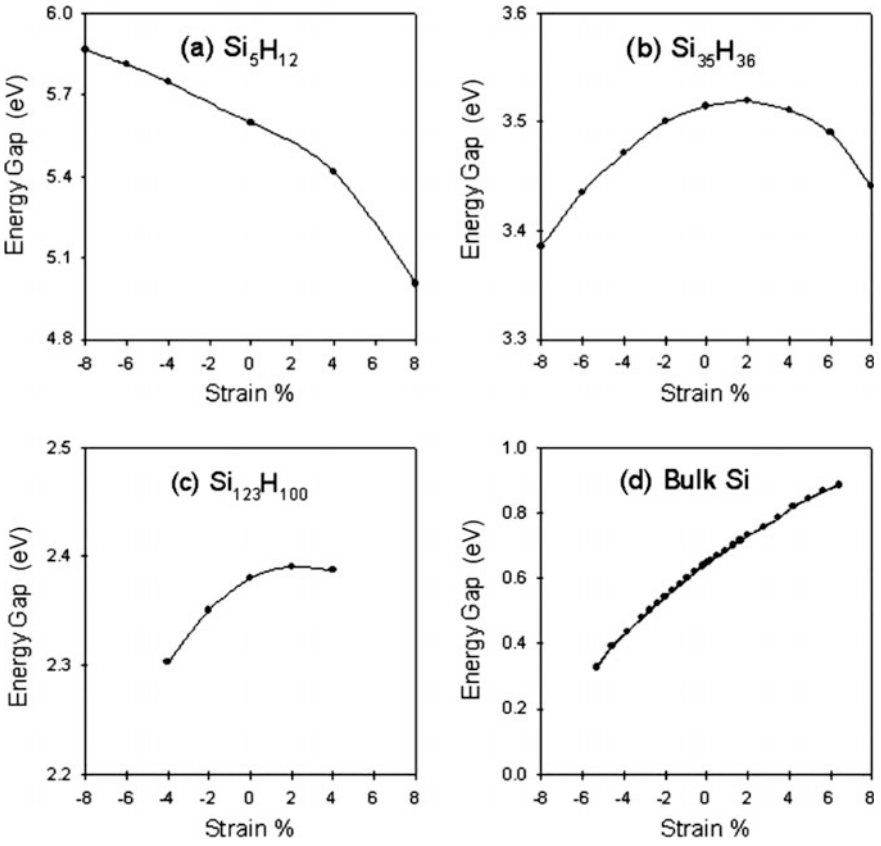

Fig. 4.16 HOMO-LUMO gap as a function of hydrostatic strain for Si nanoparticles. Reprinted with permission from Ref. [58]. Copyright 2006 by American Physical Society

behavior. Note that the overall change in the E_G with strain is rather small in this intermediate regime. Additionally, the E_G decreases both with compressive and expansive strains, exhibiting an approximately parabolic behavior.

Replotting the data in Fig. 4.16 yields three different regimes of strain-E_G dependence (Fig. 4.17). In regime I (clusters smaller than 1 nm), tensile strain results in a significant reduction of the E_G. Compression yields an opposite trend. In regime II (cluster size $\sim$1–2 nm), the E_G is effectively insensitive to strain. Finally, in regime III (>2 nm), bulk-like behavior results—tension increases the E_G. These results are understood as follows. From a continuum perspective, the predominant effect of hydrostatic strain is to shift energy levels and these changes are related to the deformation potentials. The band gap of a strained semiconductor can be expressed as $E_G \sim \Delta a^* \varepsilon$, where $\Delta a = ac - av$. ε is the trace of the strain tensor and a_c and a_v correspond to the deformation potentials for the conduction and the valence

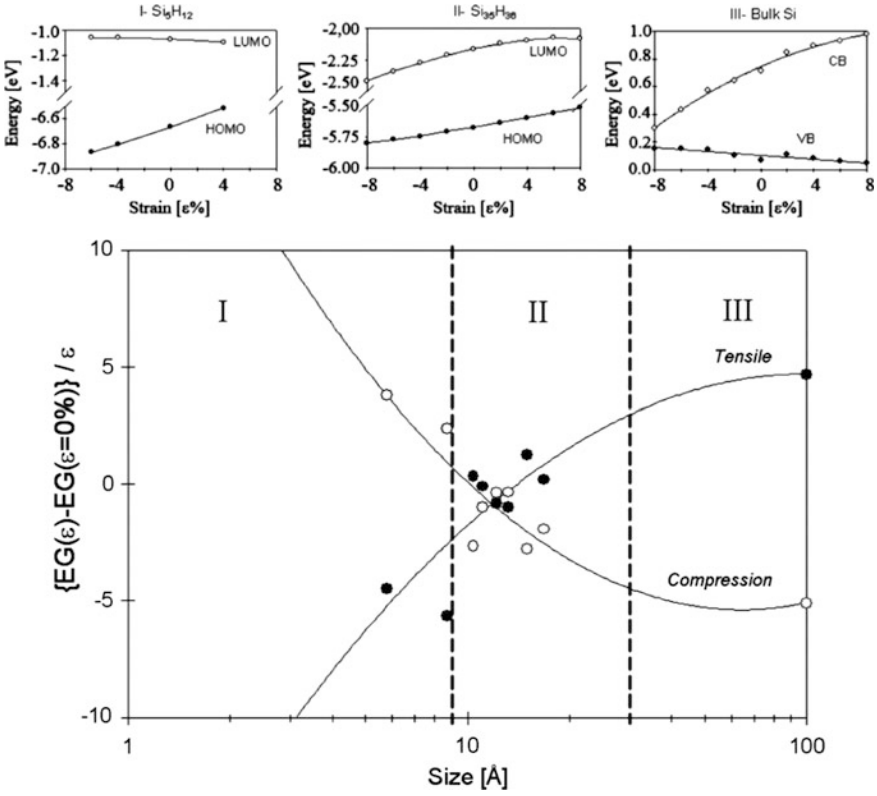

Fig. 4.17 Change in the HOMO and LUMO energies of Si nanoparticles under hydrostatic strain (*top*) and the relationship among the energy gap, hydrostatic strain, and nanoparticle size (*bottom*). In the bottom figure, filled and open symbols represent 4% tensile and compressive strains, respectively. Reprinted with permission from Ref. [58]. Copyright 2006 by American Physical Society

band, respectively. Although the concept of a deformation potential is ill defined for such small sizes, it is nevertheless a useful concept to interpret our results. As shown in Fig. 4.17, the a_v (the deformation potential for HOMO) is large compared to a_c (the deformation potential for LUMO) for small clusters. The electron density contour plots of HOMO and LUMO for Si_5H_{12} at a fixed value (Fig. 4.18) illustrate that the HOMO has a bonding character; that is, the electron cloud is mainly located in the intermediate regions shared by silicon atoms. In contrast, the LUMO has an antibonding character, i.e., the charge density primarily distributes in the vicinity of atoms. The reduction of Si–Si bond lengths on compression makes the electron cloud of HOMO more efficiently shared by Si atoms. This results in an appreciable

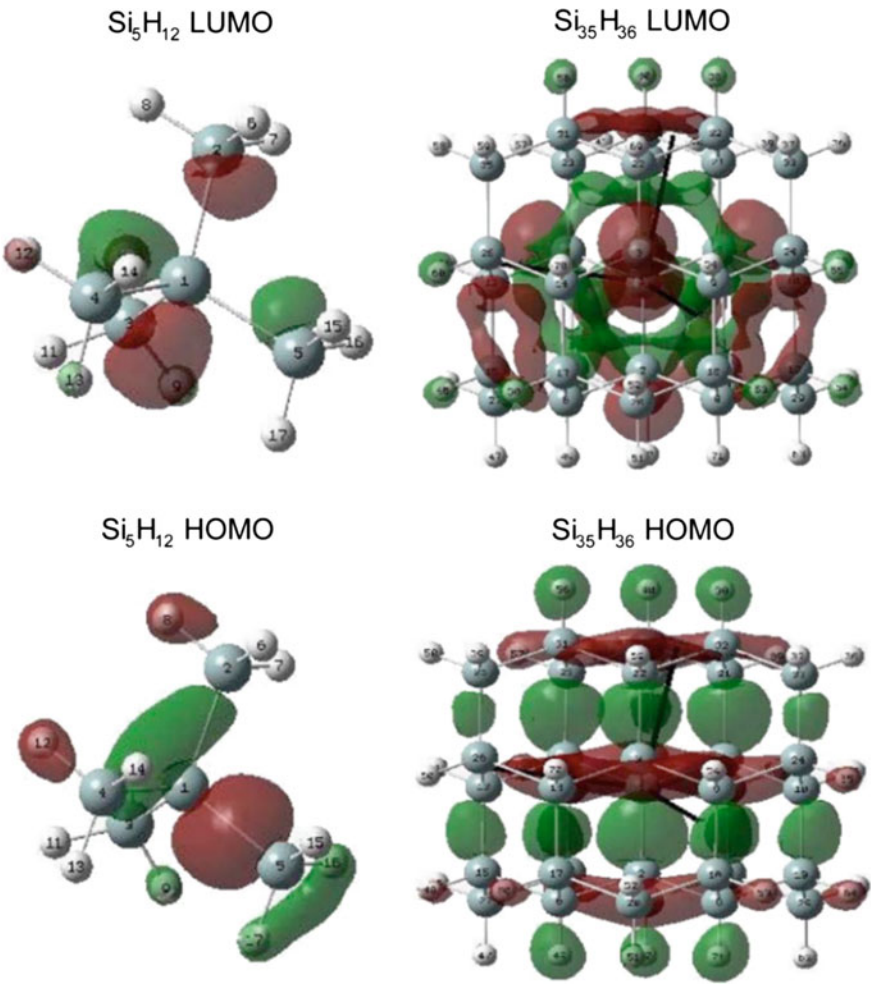

Fig. 4.18 HOMO and LUMO electron density of Si nanoparticles under no hydrostatic strain. Reprinted with permission from Ref. [58]. Copyright 2006 by American Physical Society

decrease of the HOMO energy due to the increased electron-nucleus attraction (the change in the electron-electron repulsion energy is relatively small). In the case of LUMO, the strain effect is small as electrons are already localized in the neighborhood of the atoms. (Note that the effect of nuclei–nuclei interaction is taken as a constant shift in the total energy—this is not included in the calculation of electronic orbital energies.) In intermediate size clusters, LUMO also acquires a level of bonding character in addition to its inherent antibonding nature (see Fig. 4.18). Consequently, both LUMO and HOMO are similarly affected by strain (Fig. 4.17), thus resulting in a negligible change in the E_G. The bulk-like behavior is fully expected. As empirical models (such as the standard 8×8 *kp* approach) indicate, the strain band gap coupling is linear (i.e., the deformation potential parameter is a constant). The valence band, due to its *p*-type symmetry, does not exhibit a major shift, thus contributing little to strain-induced band gap changes. On the other hand, the conduction band shows a large shift. This is fully reflected in the larger (empirical) conduction band deformation potential constants as compared to the valence band deformation constants.

4.4 Conclusion

Theoretical investigations about the effect of strain on the electronic properties in nanosemiconductors were reviewed to give insights to strain engineering to control band gap energies. After briefly covering bulk materials, investigations using first-principles and other methods for the effect of strain on the band gap properties were introduced. While a number of theoretical approaches have been devoted to the strain engineering of two- and one-dimensional nanosemiconductors thus far, such efforts for zero-dimensional nanostructures found in the literature are limited and further studies are expected. Such theoretical investigations will certainly contribute to driving practical applications for novel nanostructured devices.

References

1. J. Li, Z. Shan, and Evan Ma, MRS Bulletin, 39, 108 (2014)
2. M. Rohlfing, P. Krieger, J. Pollmann, Phys. Rev. B **48**, 17791 (1993)
3. Y. Matsushita, K. Nakamura, A. Oshiyama, Phys. Rev. B **84**, 075205 (2011)
4. M. Bouhassoun, A. Schindlmayr, Adv. Condens. Mat. Phys. **2015**, 453125 (2015)
5. Y. Umeno, T. Kitamura, Mater. Sci. Eng., B **88**, 79 (2002)
6. M.D. McCluskey, E.E. Haller, *Dopants and Defects in Semiconductors* (CRC Press, 2012)
7. D.B. Holt, B.G. Yacobi, *Extended Defects in Semiconductors* (Cambridge University Press, 2014)
8. D.A. Drabold, S. Estreicher, *Theory of Defects in Semiconductors* (Springer, 2006)

9. M.J. Puska, S. Pöykkö, M. Pesola, R.M. Nieminen, Phys. Rev. B **58**, 1318 (1998)
10. M.I.J. Probert, M.C. Payne, Phys. Rev. B **67**, 075204 (2003)
11. A.F. Wright, Phys. Rev. B **74**, 165116 (2006)
12. F. Corsetti, A.A. Mostofi, Phys. Rev. B **84**, 035209 (2011)
13. J. Dabrowski, G. Kissinger, Phys. Rev. B **92**, 144104 (2015)
14. J.R.K. Bigger, D.A. McInnes, A.P. Sutton, M.C. Payne, I. Stich, R.D. King-Smith, D.M. Bird, I.J. Clarke, Phys. Rev. Lett. **69**, 2224 (1992)
15. S. Öberg, P.K. Sitch, R. Jones, M.I. Heggie, Phys. Rev. B **51**, 13138 (1995)
16. F. Liu, M. Mostoller, V. Milman, M.F. Chisholm, T. Kaplan, Phys. Rev. B **51**, 17192 (1995)
17. M. Kohyama, Modelling Simul. Mater. Sci. Eng. **10**, R31 (2002)
18. J. Dabrowski, M. Scheffler, Appl. Surf. Sci. **56–58**, 15 (1992)
19. J. Fritsch, P. Pavone, Surf. Sci. **344**, 159 (1995)
20. G.B. Adam, O.F. Sankey, Phys. Rev. Lett. **67**, 867 (1991)
21. M. Zitzlsperger, R. Honke, P. Pavone, U. Schroder, Surf. Sci. **377–379**, 108 (1997)
22. F. Bechstedt, A.A. Stekolnikov, J. Furthmüller, P. Käckell, Phys. Rev. Lett. **87**, 016103 (2001)
23. Y.S. Kim, S.M. Lee, Phys. Rev. B **75**, 165304 (2007)
24. Y. Umeno, A. Kushima, T. Kitamura, P. Gumbsch, J. Li, Phys. Rev. B **72**, 165431 (2005)
25. G. Gui, J. Li, J. Zhong, Phys. Rev. B **78**, 075435 (2008)
26. J.H. Wong, B.R. Wu, M.F. Lin, J. Phys. Chem. C **116**, 8271 (2012)
27. M. Farjam, H. Rafii-Tabar, Phys. Rev. B **80**, 167401 (2009)
28. S.M. Choi, S.H. Shi, Y.W. Son, Phys. Rev. B **81**, 081407R (2010)
29. N. Kerszberg, P. Suryanarayana, RSC Adv. **5**, 43810 (2015)
30. G. Gui, D. Morgan, J. Booske, J. Zhong, Z. Ma, Appl. Phys. Lett. **106**, 053113 (2015)
31. S. Ismail-Beigi, T. Arias, Phys. Rev. B **57**, 11923 (1998)
32. R. Rurali, A. Poissier, N. Lorente, Phys. Rev. B **74**, 165324 (2006)
33. X. Zhao, C.M. Wei, L. Yang, M.Y. Chou, Phys. Rev. Lett. **92**, 236805 (2004)
34. T. Vo, J. Williamson, G. Galli, Phys. Rev. B **74**, 045116 (2006)
35. J.A. Yan, L. Yang, M.Y. Chou, Phys. Rev. B **76**, 115319 (2007)
36. G.W. Peng, Y.P. Feng, Appl. Phys. Lett. **91**, 083116 (2007)
37. A.J. Lu, R.Q. Zhang, S.T. Lee, Appl. Phys. Lett. **91**, 263107 (2007)
38. P.W. Leu, A. Svizhenko, K. Cho, Phys. Rev. B **77**, 235305 (2008)
39. L. Huang, N. Lu, J.A. Yan, M.Y. Chou, C.Z. Wang, K.M. Ho, J. Phys. Chem. C **112**, 15680 (2008)
40. J. Kim, M.V. Fischetti, J. Appl. Phys. **110**, 033716 (2011)
41. R.Q. Zhang, C. Hou, N. Gao, Z. Wen, Q. Jiang, ChemPhysChem **12**, 1302 (2011)
42. T. Ito, K. Nishidate, M. Baba, M. Hasegawa, Surf. Sci. **514**, 222 (2002)
43. S. Ogata, Y. Shibutani, Phys. Rev. B **68**, 165409 (2003)
44. Y. Umeno, T. Kitamura, A. Kushima, Comp. Mater. Sci. **31**, 33 (2004)
45. O. Gulseren, T. Yildirim, S. Ciraci, C. Kilic, Phys. Rev. B **65**, 155410 (2002)
46. Y. Umeno, T. Kitamura, A. Kushima, Comp. Mater. Sci. **30**, 283 (2004)
47. H. Jiang, G. Wu, X. Yang, J. Dong, Phys. Rev. B **70**, 125404 (2004)
48. S.M. Choi, S.H. Jhi, Carbon **46**, 773 (2008)
49. L. Sun, Q. Li, H. Ren, H. Su, Q.W. Shi, J. Yang, J. Chem. Phys. **129**, 074704 (2008)
50. Y. Li, X. Jiang, Z. Liu, Z. Liu, Nano Res. **3**, 545 (2010)
51. O. Hod, G.E. Scuseria, Nano Lett. **9**, 2619 (2009)
52. H. Zheng, Z.F. Wang, T. Luo, Q.W. Shi, J. Chen, Phys. Rev. B **75**, 165414 (2007)
53. M. Liu, V.I. Artyukhov, H. Lee, F. Xu, B.I. Yakobson, ACS Nano **7**, 10075 (2013)
54. O. Cretu, A.R. Botello-Mendez, I. Janowska, C. Pham-Huu, J.C. Charlier, F. Banhart, Nano Lett. **13**, 3487 (2013)
55. V.I. Artyukhov, M. Liu, B.I. Yakobson, Nano Lett. **14**, 4224 (2014)

56. C. Ming, F.X. Meng, X. Chen, J. Zhuang, X.J. Ning, Carbon **68**, 487 (2014)
57. Y.H. Hu, J. Phys. Chem. C **115**, 1843 (2011)
58. X.H. Peng, S. Ganti, A. Alizadeh, P. Sharma, S.K. Kumar, S.K. Nayak, Phys. Rev. B **74**, 035339 (2006)
59. X.H. Peng, A. Alizadeh, N. Bhate, K.K. Varanasi, S.K. Kumar, S.K. Nayak, J. Phys.: Condens. Matter **19**, 266212 (2007)
60. X. Jiang, J. Zhao, X. Jiang, J. Nanopart. Res. **14**, 818 (2012)

Chapter 5
Ferroelectric Nanostructures

Abstract Ferroelectricity is known to be sensitive to strain, and therefore, salient effects of nanostructures on ferroelectric properties can be expected. We introduce recent first-principles calculations of ferroelectricity in two- and one-dimensional nanostructures; peculiar polarization and domain structures found in surfaces, ferroelectric instability in ultrathin capacitors, polarization transition in nanowires under strain, etc. Effective Hamiltonian approaches for ferroelectric nanodots to investigate peculiar domain structures are also covered.

Keywords Ferroelectricity · Ferroelectric instability · Domain structure

5.1 Ferroelectricity in Bulk

5.1.1 Ferroelectric Instability and Its Response to Strain

Ferroelectricity is a property of solid materials that the material exhibits spontaneous electric polarization (polarization remains in the absence of an electric field), whose direction can be switched by the application of a substantially large electric field. Not all polar crystals are ferroelectric because of the latter requirement. There are many types of perovskites (crystals that possess ABO_3 structure) that exhibit ferroelectricity, e.g., $BaTiO_3$, $PbTiO_3$, and $KNbO_3$.

$PbTiO_3$ is a prototypical ferroelectric perovskite. Figure 5.1 shows a paraelectric (PE; having no polarization) cubic structure at high temperature and tetragonal (ferroelectric; FE) structures of $PbTiO_3$. In a tetragonal structure, the Ti atoms and oxygen sublattice are shifted in the same direction to the Pb atoms but with different distances. The energy profile along the shift (ξ) shows a typical double-well shape (Fig. 5.1). One of early studies about the mechanism of ferroelectricity can be found in a work by Cochran [1], which pointed out that lattice dynamics plays an essential role in ferroelectric phase transition. That is, the softening of TO (transverse optical) phonons at the Γ-point causes lattice instability, leading to phase transition of a displacive type. According to the model proposed by Cochran, the ferroelectric

Y. Umeno et al., *Multiphysics in Nanostructures*, Nanostructure Science and Technology, DOI 10.1007/978-4-431-56573-4_5

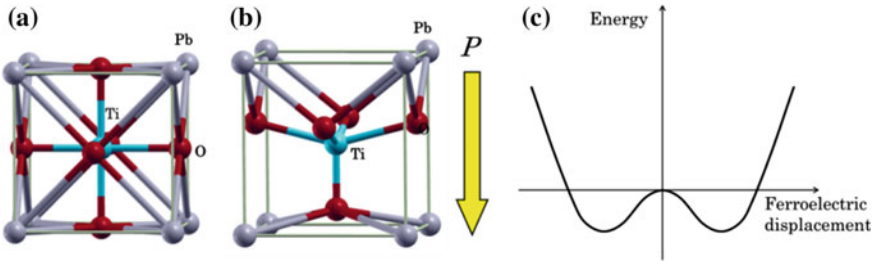

Fig. 5.1 Crystal structures of $PbTiO_3$ at paraelectric (*cubic*, **a**), ferroelectric (*tetragonal*, **b**) phases. Double-well shape of internal energy (**c**)

instability stems from a delicate balance and competition between sort-range forces and long-range dipolar forces.

The balance between the short-range and dipolar interactions can be modified by applied stress, resulting in stress (and strain) effect on the ferroelectricity. For example, Gosez et al. [2] revealed that in $BaTiO_3$, the destabilizing role of the dipolar interaction and the stabilizing short-range forces increase under hydrostatic pressure, causing the reduction in the ferroelectric instability. Tinte et al. [3] pointed out that a hydrostatic tension can induce a large tetragonal strain and an increase of the spontaneous polarization in $PbTiO_3$. Then, naturally, response of ferroelectricity to anisotropic stress or strain came to much attention. In particular, the effect of biaxial strain drew much interest as the deposition of coherent epitaxial films can impose relatively large epitaxial strains to thin-film perovskites [4–12]. Besides ab initio analysis of ground-state properties, prediction of ferroelectric phase transition at finite temperature has been carried out with the aid of a simple Landau–Ginzburg–Devonshire (LGD) theory [13, 14]. Such studies for prototypical ferroelectric crystals can be found, for example, in Refs. [15, 16]. These studies provide not only scientific significance but also practical impact because of growing interest in "strain engineering" for better performance of devices utilizing ferroelectricity of materials.

5.1.2 Domain Structure and Domain Switching

Ferroelectric materials have at least two stable states of spontaneous polarization, which can be switched from one state to another by the application of external electric or strain (stress) field. Thus, ferroelectric crystals usually contain regions that have different orientations of polarization. Such regions are called "ferroelectric domains," and boundaries between two adjacent domains are called "domain walls" (DWs). DWs are characterized by the angle between polarization orientations in the adjoining domains. Most typical in ferroelectric crystals are 90° and 180° DWs, as schematically shown in Fig. 5.2.

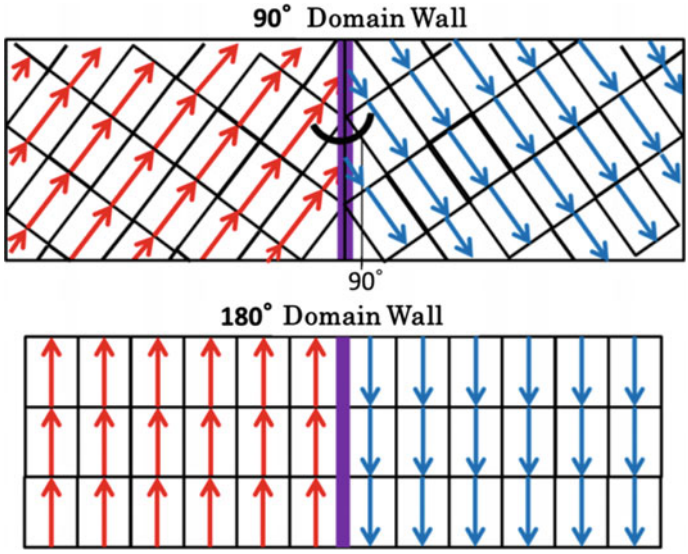

Fig. 5.2 Schematic illustrations of 90° and 180° DWs

Several ab initio analyses of DW structures have been carried out thus far. The first ab initio-based approach to the problem of DW structure in perovskites was made by Padilla et al. [17] for the 180° DW of $BaTiO_3$. Their approach was based on an effective Hamiltonian that had been determined from first-principles DFT calculations, which enables one to perform simulations at nonzero temperatures and is essential to deal with $BaTiO_3$ because its room temperature tetragonal phase is different from the ground-state (zero temperature) rhombohedral phase. The Ba-centered DW structure was found to be much energetically favorable than the Ti-centered one. The DW width was estimated to be 5.6 Å, and the free energy of the 180° DW of $BaTiO_3$ was calculated to be 4–5 mJ/m^2 at 250–260 K. It was also pointed out that the domain boundary entails a simple reversal of the ferroelectric order parameter rather than a rotation, as opposed to the case of ferromagnetic DWs, because of the much stronger strain coupling in the ferroelectric case.

A "direct" first-principles DFT calculation (i.e., without using an effective Hamiltonian approach) is possible for $PbTiO_3$ because of its phase diagram: The zero temperature and room temperature phases are the same tetragonal ferroelectric phase. The 180° DW structure of $PbTiO_3$ was investigated by Pöykkö and Chadi [18]. They performed plane wave pseudopotential calculations based on DFT within the local density approximation using norm-conserving pseudopotentials for supercells containing 30–60 atoms. The Pb-centered DW structure was found to be most stable, with an extremely narrow width of only about two lattice constants and the energy density of 0.1–0.2 J/m^2, which is significantly larger than that of $BaTiO_3$.

An ab initio study of 90° DW was first done by Meyer and Vanderbilt [19] for $PbTiO_3$, while they also investigated 180° DW in the same study. The energy of the 90° DW was estimated to be 35 mJ/m^2, which is in reasonable agreement with a rough estimate of 50 mJ/m^2 from a HRTEM experiment and is about four times lower than that of its 180° DW counterpart (132 mJ/m^2). The significant energy difference can be explained as follows: A principle difference between the 180° and 90° DWs is that the polarization reverses completely in the latter, passing through a high-energy cubic state of zero polarization in the center of the DW, whereas it merely rotates in the 90° case, passing through a more favorable orthorhombic state. The study also includes calculations of the barrier energies for the motion of 90° and 180° DWs. For the 180° DW, atomic configurations ξ_{Pb} and ξ_{Ti} of Pb- and Ti-centered walls on neighboring lattice planes were taken and the linear interpolation $\xi_{Pb} + \lambda(\xi_{Ti} - \xi_{Pb})$ was formed, where λ indicates a reaction coordinate along the path. For the 90° DW, the linear interpolation $\xi_n + \lambda(\xi_{n+1} - \xi_n)$ between the atomic configurations ξ_n and ξ_{n+1} of two fully relaxed supercells with DW positions on nearest-neighbor (101) lattice planes was formed. The energy profile for the path of the interpolation was calculated as shown in Fig. 5.3. The estimated energy barrier of the motion of 180° DW was 37 mJ/m^2, while that of 90° DW was

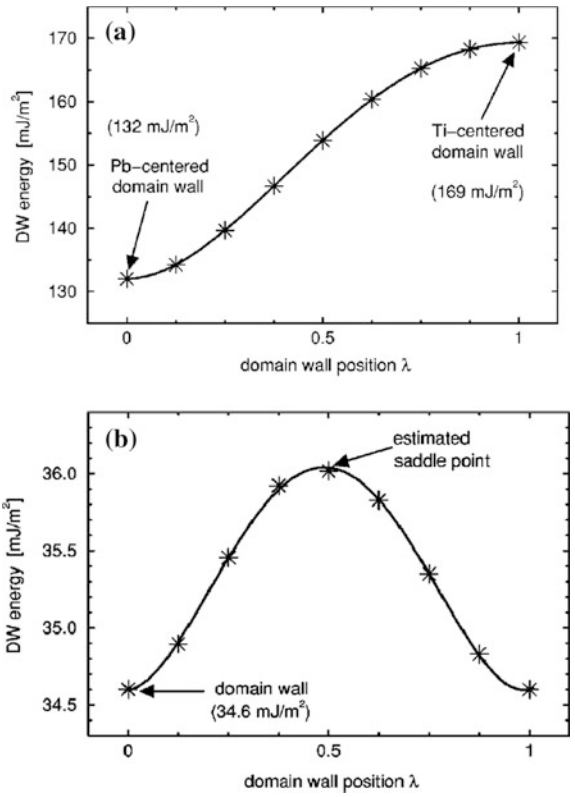

Fig. 5.3 Energy profile along the estimated path for the motion of a 180° DW **a** and 90° DW **b**, calculated with supercells of ten perovskite unit cells. Reprinted with permission from Ref. [19]. Copyright 2002 by American Physical Society

1.6 mJ/m^2, which is only about 4% of the DW formation energy, indicating that the 90° DW can be mobile at a very low temperature. Meyer and Vanderbilt also clarified that the 90° DW has the character of a dipole layer (see Fig. 5.4). Due to the fact that the polarization perpendicular to the DW is slightly reduced in the interior of the wall, a decreasing (increasing) polarization on the left (right) side of the DW leads to an accumulation of positive (negative) charge. As a consequence, the electrostatic potential should jump from one side of the domain wall to the other, leading to band offsets for both the valence and conduction bands. In the setup of a simulation cell with the periodic boundary condition, an artificial electric field should arise to compensate the band offset (see Fig. 5.4c), which must be taken into account for the evaluation of DW-related properties.

When an external electric field with a certain strength is applied, polarizations in domains alter (rotate) so that the polarization vectors become aligned to follow the electric field. This process is called "domain switching" and is often accompanied with motion of DWs. The ability of domain switching and associated DW motion is therefore one of important properties that are critical for functionality of ferroelectric crystals such as piezoelectricity. DW switching can occur also by the application of an external stress or strain. Shimada et al. [20] performed ab initio calculations to demonstrate the motion of the 90° DW by shear stress in $PbTiO_3$. The calculation was carried out using a simulation cell containing 90° DWs on the (101) plane where shear stress on the DW with increasing magnitude was applied until the DW motion occurred. It was found that during the stress-induced DW motion, a Pb–O covalent bond at the center of the DW broke and that another bond on the neighboring Pb–O site were concurrently formed with a large movement of

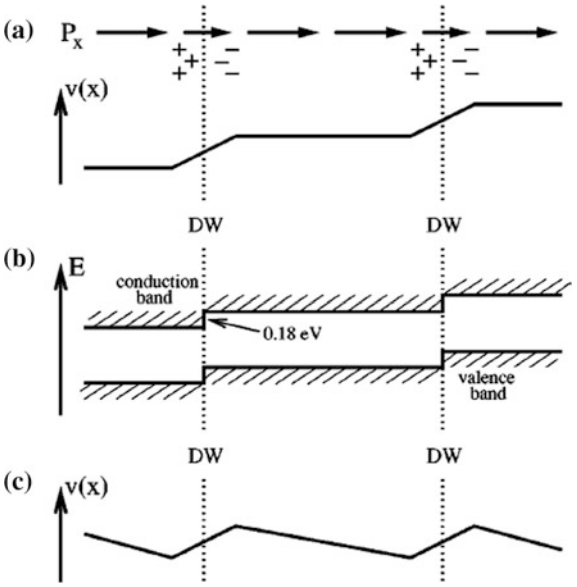

Fig. 5.4 Schematic illustration of a periodic array of 90° DWs. **a** The component of P_x of the polarization normal to DW and the induced charge and electron potential $v(x)$, assuming zero macroscopic electric field. **b** Corresponding band offsets. **c** Potential in the case of supercell calculation using periodic boundary conditions. Reprinted with permission from Ref. [19]. Copyright 2002 by American Physical Society

the Pb atom, representing the reconstruction of the Pb–O bond associated with the DW motion. The critical shear stress to cause the stress-induced DW motion was evaluated to be 152 MPa using a simulation model containing the DWs separated by 25 Å.

Besides approaches with first-principles methods, atomistic simulation methods using empirical interatomic models may have to be employed to perform more "direct" demonstrations of the effect of defects because substantially large simulation models must be used for such a purpose. For problems where remarkable size effects are expected, empirical modeling approaches can be a realistic choice if a reliable interatomic model is available. Kubo et al. investigated the 90° DW of $PbTiO_3$ using an empirical shell-model interatomic potential function, focusing on the behavior of stress-induced DW motion [21, 22]. A good agreement between the empirical model and first-principles calculations was confirmed in terms of the critical shear stress for DW motion at 0 K. Then, the effect of simulation cell size was examined to find a significant influence on shear modulus and the critical shear stress of DW motion due to a long-range interaction produced by artificial electric field caused by the periodic boundary condition. They found that the critical stress is proportional to the inverse of DW separation in the simulation model and pointed out the DFT estimation in Ref. [20] should have substantially underestimated the critical shear stress of DW motion. Molecular dynamics simulations using the shell-model interatomic potential were further carried out to investigate the effect of temperature and kink structure on the DW motion. MD simulations at finite temperatures showed that critical shear stress is reduced and that the DW structure is characterized by perovskite unit cells with a preferential polarization along the $\langle 011 \rangle$ direction. Simulations of kink structures revealed a long-range strain field caused by local disorder of the crystal lattice in the area surrounding kinks. It was demonstrated that kink structures significantly facilitate the motion of 90° DWs in $PbTiO_3$.

5.1.3 *Effect of Defects*

The effect of defects on various properties is always one of the most interesting and important issues in materials science. Attention has been drawn to how defects affect the DW properties. Since defects are expected to make remarkable influence on the mobility of the DW, it has long been a central issue to clarify the mechanism of the DW motion hindered or enhanced by the presence of defects. For example, application of ferroelectric materials is hindered by the problem of fatigue, where the magnitude of switchable polarization is reduced remarkably after repeated polarization reversals. Nevertheless, there have been a limited number of studies on this problem by means of ab initio calculations thus far because of the difficulty of employing a simulation box containing a large number of atoms; i.e., effect of internal inhomogeneous geometries (understructures) on ferroelectricity has not systematically been studied. As seen below, studies regarding the issue of understructures in ferroelectric materials have been mostly dedicated to point defects.

Pöykkö and Chadi made a first-principles study investigating the electronic and ionic structures for the oxygen-vacancy metal-impurity defect, which is a dipolar complex, in $PbTiO_3$ [23]. Their calculation estimated the binding energy of an oxygen vacancy to a Pt impurity to be about 3.0 eV. The complex is stabilized by electron capture and pins the polarization of the surrounding lattice because the binding energy is extremely large when the defect polarization is in the opposite direction of the bulk polarization. The fact that the stability of a Pt-vacancy pair is due to electron strapped on the center explains why electron injection is more crucial to fatigue than hole injection [24].

Formation energies of Pb, Ti, and O vacancies in $PbTiO_3$ were estimated using DFT calculations by Zhang et al. [25] taking into account chemical potentials altered by surrounding environment. Considering several thermo-chemistry conditions, the chemical potentials of Pb, Ti, and O atoms were calculated according to the conditions, based on which the formation energies of vacancies were estimated. The results suggest that the Ti vacancy is difficult to form due to its high formation energy under both oxidizing and reducing conditions, whereas the Pb vacancy can be formed when the oxygen partial pressure is elevated above a certain value. Under reducing atmosphere, oxygen vacancies are more stable than any other vacancies. The vacancies of the oxygen bonded with Ti in the *c*-axis are always easier to form than those bonded with Ti in the perpendicular directions, affecting the ferroelectricity of the crystal.

5.2 Thin Film and Surface Property: Two-Dimensional Structure

5.2.1 Ferroelectric Surface Structure

As a prototypical structure of two-dimensional geometry, the surface presents interesting properties due to coordinate numbers of surface atoms being different from those of bulk atoms. It is therefore worthwhile to extensively investigate surface properties of ferroelectric materials. In fact, there have been ab initio studies of surface structures and properties, ranging from clean surfaces to surfaces with defects. Among various ferroelectric surfaces, $PbTiO_3$ surfaces have been investigated most because of the easiness of applying the DFT calculation method as was mentioned in the previous section.

Meyer and Vanderbilt [26] investigated the (001) surfaces of perovskite compounds $BaTiO_3$ and $PbTiO_3$ using a first-principles USPP approach to clarify the structural properties and their response to external electric fields. In their calculations, the surfaces were represented by periodically repeated slabs (with vacuum layers in between), where the external electric field normal to the surfaces was realized by inserting a dipole layer in the vacuum region of the supercells. They pointed out the problem in simulations using a slab model that arises due to the

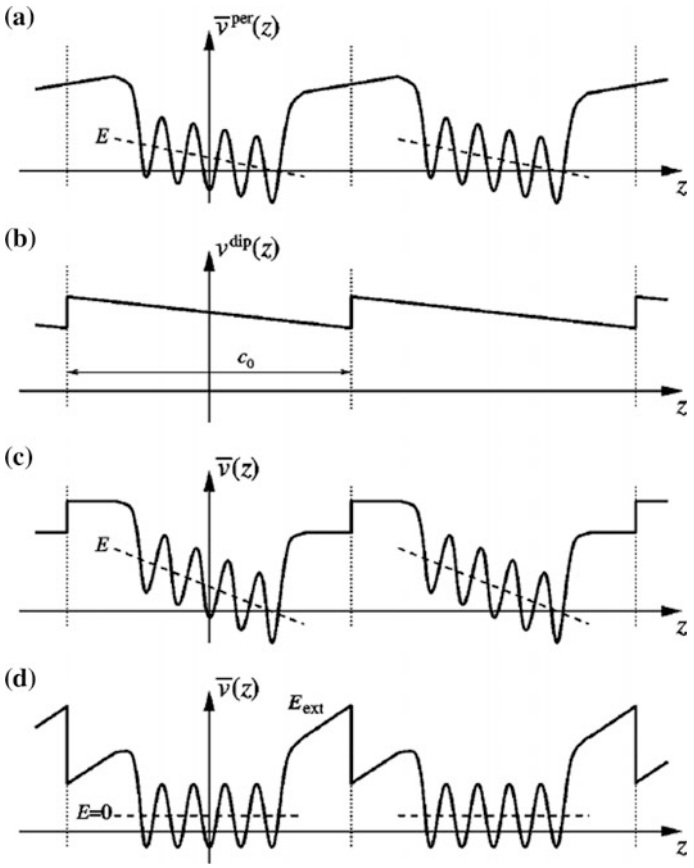

Fig. 5.5 Schematics of the planar-averaged potential $\bar{v}(z)$ for periodically repeated slabs; **a** with periodic boundary conditions, **b** potential of the dipole layer, **c** dipole-corrected slabs with vanishing external electric field, and **d** dipole-corrected slabs with vanishing internal electric field. Reprinted with permission from Ref. [26]. Copyright 2001 by American Physical Society

periodic boundary condition, especially when the slab exhibits a polarization perpendicular to the surface. The appropriate electrical boundary condition in this case is a vanishing electric field inside the slab, which is not in the ordinary periodic boundary condition because it requires continuity of the electrostatic potential in the vacuum layer. Note that a vanishing internal electric field is equivalent to the short-circuit boundary conditions. The insertion of a dipole layer in the vacuum layer addresses this problem because it can adjust external electric field (E_{ext}) as is schematically illustrated in Fig. 5.5. Calculations using this scheme enable investigation of the response of surface structures to external electric fields. The rumpling parameter was defined as the amplitude of the relative displacements between the metal and the oxygen ions normal to the surface. In the $BaTiO_3$ (001) surface, the rumpling parameter of the outermost BaO layer shows nonlinear development with

varying E_{ext} in both BaO- and TiO_2-terminated surfaces, while the other layers exhibit nearly linear response in the rumpling parameter. The $PbTiO_3$ (001) surface behaves similarly, and the rumpling parameter of the outermost PbO layer changes nonlinearly.

To distinguish whether the slab exhibits ferroelectric instability or paraelectric nature, Meyer and Vanderbilt employed a simple phenomenological picture: As illustrated in the schematics in Fig. 5.6, the Gibbs free energy (G) as a function of the dielectric displacement field $\boldsymbol{D}$ presents a roughly quadratic curve for a paraelectric slab, whereas for a ferroelectric material the $G(\boldsymbol{D})$ curve shows a double-well structure. Differentiating the Gibbs free energy immediately gives the internal electric field: $\boldsymbol{E} = \partial \mathrm{G}/\partial \mathbf{D}$. Thus, calculating $\boldsymbol{E}(\boldsymbol{D})$ directly reveals whether a slab is FE or not (see Fig. 5.6b). Figure 5.7 shows the internal electric field as a function of the applied external electric field in the seven-layer slabs of $BaTiO_3$ and $PbTiO_3$ with (001) surfaces. The BaO-terminated $BaTiO_3$ slab obviously exhibits a FE behavior because the curve shape is similar to that of the solid line in Fig. 5.6b. For the other three cases in Fig. 5.7, it is more difficult to deduce whether the slabs show FE instability or not. However, negative slopes for the $PbTiO_3$ slab models indicate that the PbO- and TiO_2-terminated $PbTiO_3$ slabs with this thickness may exhibit FE instability.

The (001) surface structure of $PbTiO_3$ with ferroelectric polarization parallel to the surface was investigated by Umeno et al. [27] by means of first-principles DFT calculations with the local density approximation. The study also investigated the effect of lateral strain (strain parallel to the surface) on the ferroelectric structure. Unlike the case of ferroelectric polarization perpendicular to the surface that requires a special treatment under the periodic boundary condition as in Ref. [26], in-plane polarization can be treated in DFT calculations relatively easily. It is known that the perovskite structure exhibits various lattice instabilities causing antiferrodistortive (AFD), antiferroelectric (AFE), and ferroelectric (FE) distortions.

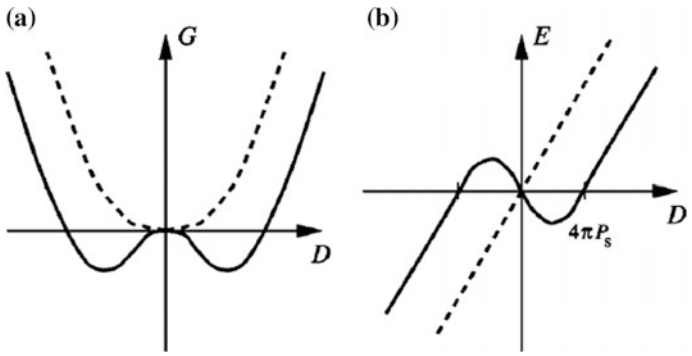

Fig. 5.6 Schematics of the Gibbs free energy G (**a**) and the internal electric field E (**b**) as a function of the dielectric displacement D, respectively, for a paraelectric (*dashed lines*) or ferroelectric (*solid*) slab. Reprinted with permission from Ref. [26]. Copyright 2001 by American Physical Society

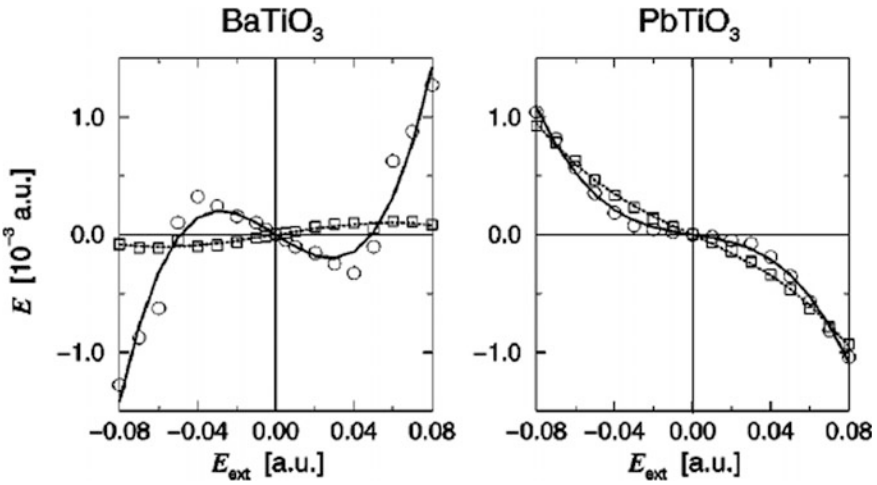

Fig. 5.7 Internal electric field E in the fully relaxed seven-layer $BaTiO_3$ and $PbTiO_3$ slabs as a function of the applied external electric field E_{ext}. Circles and Squares are for AO (BaO- or PbO-)-terminated and TiO_2-terminated slabs, respectively. Reprinted with permission from Ref. [26]. Copyright 2001 by American Physical Society

Therefore, in the study, which covered both TiO_2- and PbO-terminations, focus was placed on the latter to investigate the competition among AFD, AFE, and FE distortions at the surface using models possessing (001) surfaces with c (2×2) periodicity.

When only the FE distortion is permitted (using models with (1×1) periodicity), surface structures with FE polarizations along [100] (denoted as P[100]) and [110] (P[110]) are energetically nearly comparable, but P[110] is only slightly favored for both PbO and TiO_2-terminations. While the TiO_2-termination suppresses polarization resulting in the magnitude of polarization (δ) gradually decreasing toward the surface layer, the opposite trend is found in the PbO-termination. As for the PbO-terminated c (2×2) model, several (meta)stable structures with the coexistence of FE, AFE, and AFD are found as depicted in Fig. 5.8. In this phase, the polarization directions of unit cells are not identical; the polarizations are deviated from their average directions ([100] or [110]) clockwise or counterclockwise, which is clearly due to the contribution of AFE distortion.

Change in the structure of c (2×2) PbO-terminated surface under isotropic strain is shown in Fig. 5.9. The response of the polarization distortion to the variation of isotropic strain is similar to that in the (1×1) model; the polarization is enhanced by tension and suppressed by compression. The AFD rotation has a nature of the opposite trend. In the FE + AFD coexisted phase, FE and AFD compete with each other, resulting in the AFD rotation being more suppressed by the FE distortion under tension and the FE being more weakened by the AFD under compression.

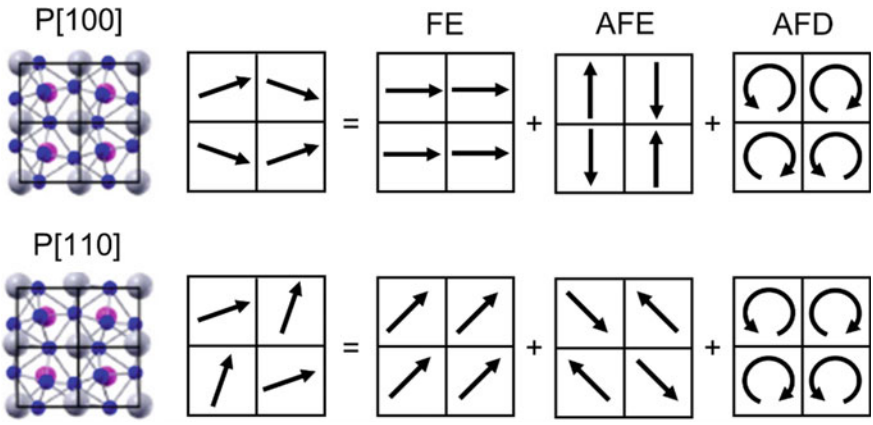

Fig. 5.8 Atomic structure of (meta)stable phases at PbO-terminated c (2×2) surface

Tensile and compressive uniaxial strain in [100] changes the aspect ratio of the $PbTiO_3$ lattices, which causes variation in the in-plane polarization direction as well as the magnitude. The polarization direction and distortion in the FE + AFD coexisted phase of the c (2×2) PbO-terminated surface under uniaxial tension in [100] (here, the transverse strain was kept zero) are shown in Fig. 5.10. Although the polarization distortion as a function of [100] strain (ε_{xx}) is quite similar to that of the (1×1) model, the FE direction behaves differently: The polarization direction (ϕ) differs among the layers. The surface layer prefers to have polarization in the [110] direction, and the adjacent layer is slightly affected. This result suggests that different polarization directions depending on the distance from the surface can be observed at the $PbTiO_3$ surface subjected to a high uniaxial strain.

Figure 5.11 shows the change in the atomistic and electronic configurations at the surface layer of the PbO-terminated models. In the (1×1) model, the center oxygen atom is shifted toward a corner Pb with a strong covalent bond (α) and two equivalent weak bonds $(\beta$ and $\gamma)$ at $\varepsilon_{xx} = 0$. When a strain is applied in the x direction, the oxygen atom immediately moves toward either of the equivalent bonds to form two strong bonds: α and β under tension or α and γ under compression. This mechanism, which is also found in the TiO_2-terminated surface, explains the polarization rotation very sensitive to the uniaxial strain in the (1×1) periodicity. On the other hand, the c (2×2) PbO-terminated surface has the different bond structure; oxygen atoms are shifted in different directions in the lattices A and B, each forming two strong covalent bonds with the nearest Pb atoms. The bond structure in the lattice A does not immediately change when a small uniaxial tension is applied. At relatively high strain, the shift of the oxygen atom occurs with the break of γ bond followed by the construction of β bond, consequently forming [100] polarization in both lattices.

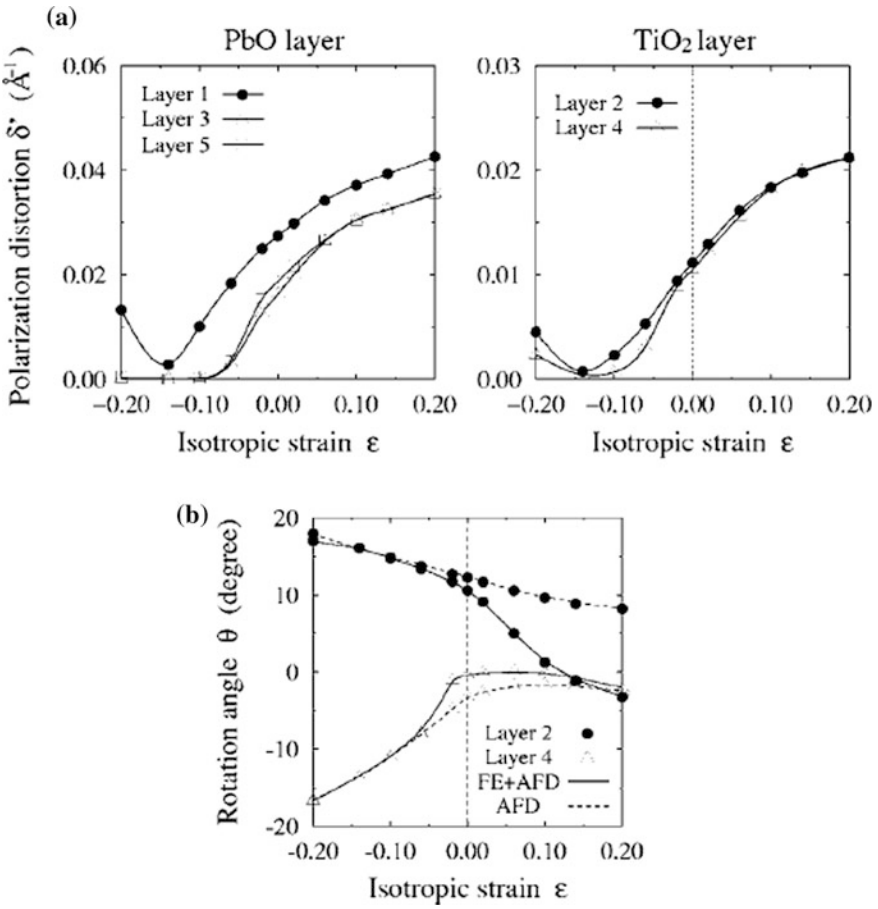

Fig. 5.9 Change in structure of c (2×2) PbO-terminated surface as a function of isotropic strain. **a** Polarization distortion per area, δ'. **b** AFD rotation angle, θ. *Dashed lines* in the bottom figure represent the results of AFD-only phase for comparison where FE polarization is artificially frozen out. Reprinted with permission from Ref. [27]. Copyright 2006 by American Physical Society

5.2.2 *Correlation Between Surface Structure and Internal Geometry*

As discussed above, for a ferroelectric thin film with a polar axis perpendicular to its surface, the termination of polarization at the surface or interface creates surface charges, which causes depolarizing field that destabilizes the ferroelectric distortion. One way to compensate for surface charges is screening by electrodes that accumulate surface charge at their interface (see next section). Another way is partitioning the system into domains as was experimentally observed by Fong et al. [28], where it was found that ferroelectricity is sustained in a $PbTiO_3$ film of three

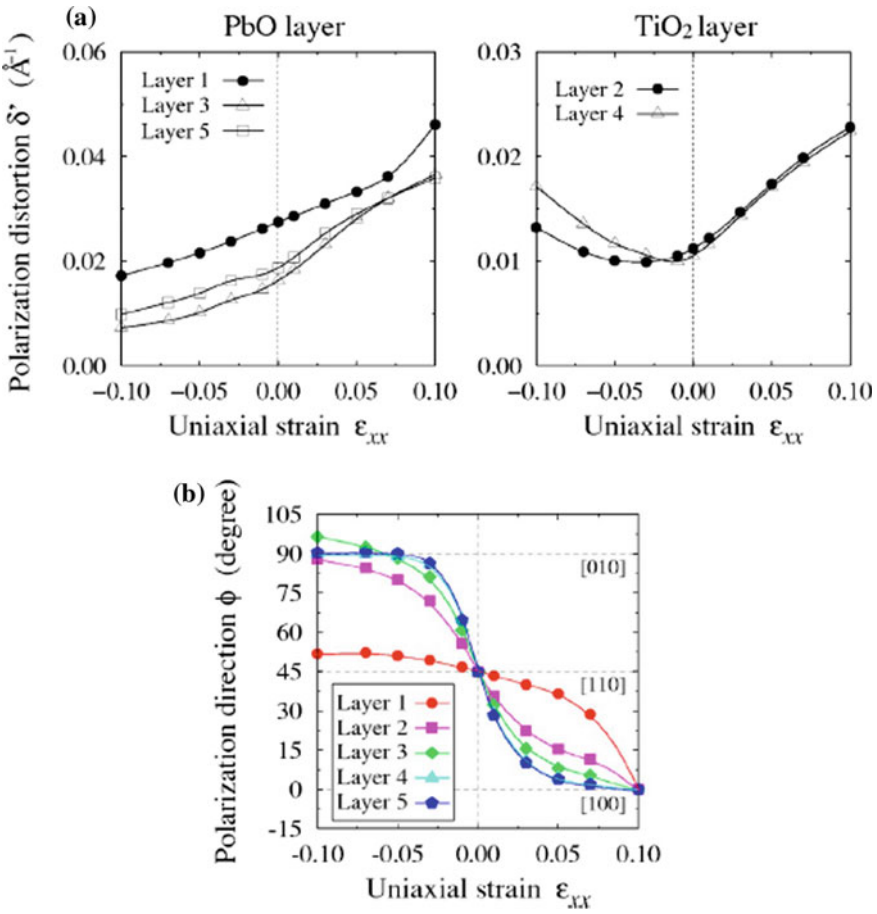

Fig. 5.10 Change in structure of c (2×2) PbO-terminated surface as a function of uniaxial strain ε_{xx}. **a** Polarization distortion per area, δ'. **b** Polarization direction, ϕ. Reprinted with permission from Ref. [27]. Copyright 2006 by American Physical Society

unit-cell thick by forming 180° stripe domains. The mechanism of forming this polydomain structure was investigated by Shimada et al. [29] using first-principles DFT calculations based on LDA. Figure 5.12 shows the local polarization distribution in a three unit-cell thick film $(m = 3)$ of a stable Ti-centered domain having a period of $N_x = 6$. A nontrivial in-plane polarization was found near the junction between the surface and domain wall. As a consequence, a closure domain structure, where polarization direction was aligned to form a closed flux, was formed in the film. Their calculations confirmed that such closure domains were also formed in thinner films of $m = 2$ and 1. In their model, the closure domain structure seems to consist of not only the 180° DW but also the 90° DW as Kittel proposed [30], which can clearly be seen in the vector field of the polarization distribution (see the

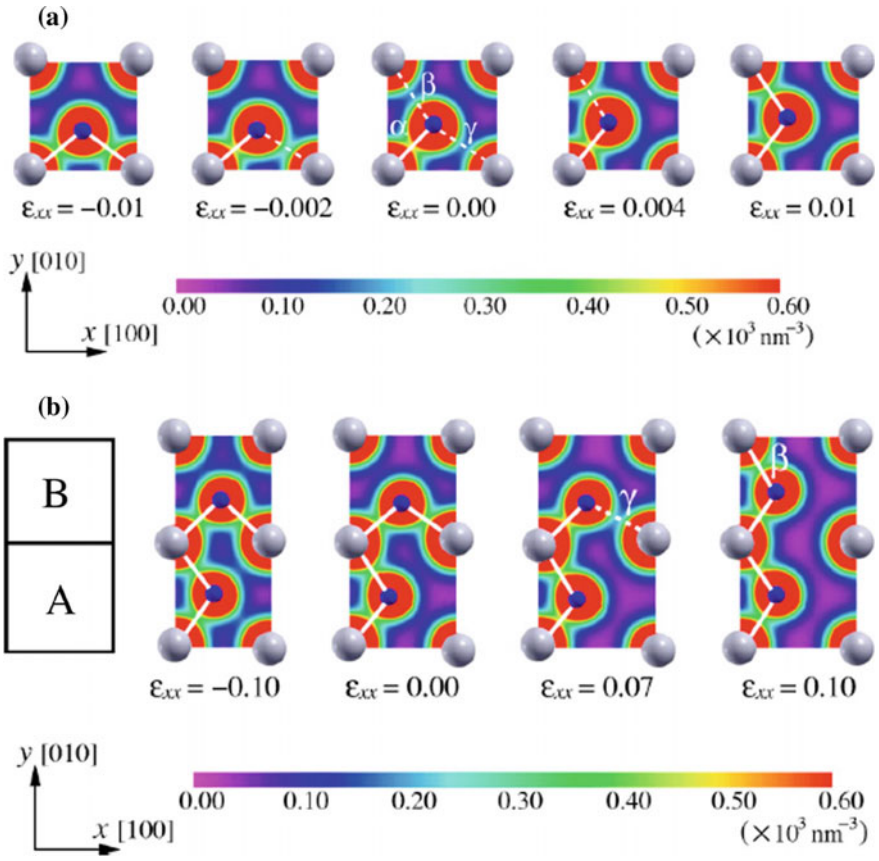

Fig. 5.11 Change in atomistic and electronic configurations at the surface PbO layer under uniaxial strain. **a** (1×1) and **b** c (2×2) models. Reprinted with permission from Ref. [27]. Copyright 2006 by American Physical Society

schematic in Fig. 5.12). The formation of the closure domain structure can considerably reduce the depolarizing field as the in-plane polarization at the surface does not produce any surface charges. Thus, more effective screening of the depolarizing field was realized by the formation of closure domains, which stabilized ferroelectric distortions in the ultrathin films.

Figure 5.13 shows the atomic displacement from the paraelectric state in a three unit-cell thick film $(m = 3)$ of a stable Ti-centered domain having a period of $N_x = 6$. The displacement of atoms forms a closure-type flux across the domain wall. This displacement pattern corresponds well to the polarization orientation of closure domains described above. A similar closed-flux displacement pattern was observed in thinner films with $m = 2$ and 1. In-plane atomic displacement was found in the first and second surface layers of the film, which aligned the polarization parallel to the surface. In other words, the in-plane displacement played a

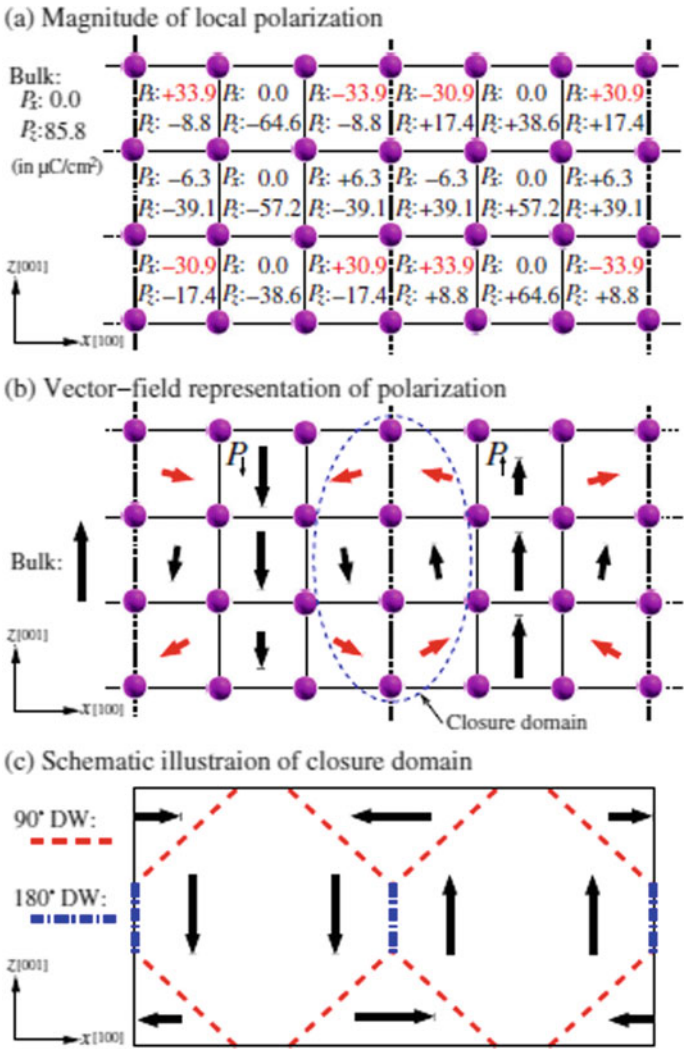

Fig. 5.12 Local polarization distribution in a three unit-cell thick film $(m = 3)$ of a stable Ti-centered domain having a period of $N_x = 6$. Reprinted with permission from Ref. [29]. Copyright 2010 by American Physical Society

significant role in the formation of closure domains, which stabilized the ferroelectric state in the film.

Shimada et al. expanded the scope of their DFT calculations to the case of a vicinal surface consisting of flat (001) terraces and surface steps [31, 32]. Surface steps have an atomically sharp edge structure consisting of (100) and (001) surfaces, which is commonly observed in perovskite oxides, such as the edge of perovskite nanowires [33], the zigzag (110) surface structure [34], and the edge of

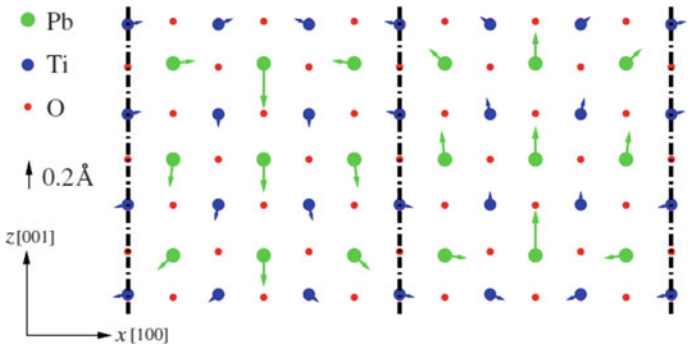

Fig. 5.13 Atomic displacement from the paraelectric state in a three unit-cell thick film ($m = 3$) of a stable Ti-centered domain having a period of $N_x = 6$. Reprinted with permission from Ref. [29]. Copyright 2010 by American Physical Society

nanoislands [35, 36]. This edge surface step is a characteristic of perovskite oxide nanostructures. Ferroelectricity at surface steps and its response to strain are therefore fundamental issues that need understanding. First-principles DFT calculations were performed to investigate ferroelectricity at $PbTiO_3$ surface steps with atomically sharp edges consisting of (100) and (001) surfaces with the polar axis in the [010] direction. The crucial role of in-plane strain along the step line ([010] direction) was also examined. In the unstrained PbO-terminated surface step, the step site exhibits polarization in the [010] direction (the y-axis) of $P_y = 108.2\,\mu C\,cm^{-2}$, which is approximately 20% greater than that of the bulk ($85.8\,\mu C\,cm^{-2}$). The local polarizations in the upper and lower terraces are nearly equal to that of the (001) surface [27]. The inner cells of the slab have a polarization comparable to that of the bulk, implying that the effect of the surface step extends to the depth of nearly one unit cell below the surface layer. In contrast, the local polarization P_y is lower than the bulk value at TiO_2-terminations, especially at the step site, where it is 19% lower than the bulk value. In addition, a nontrivial z-component (along the [001] direction) is found in step sites ($P_z = -21.9$ and $-37.1\,\mu C\,cm^{-2}$) for the PbO- and TiO_2-terminations, respectively, whereas the other sites exhibit very low values. This indicates that spontaneous polarization is locally rotated toward the inside of the film at step edges.

Figure 5.14 shows the x, y and z components of the site-by-site local polarization as a function of strain ε_{yy} for the PbO-terminated surface step. The local polarization P_y increases nearly linearly with respect to the applied tensile strain at all sites. This indicates that the tensile strain along the step line tends to enhance the ferroelectricity, because it increases the tetragonality of the $PbTiO_3$ lattice, which is strongly coupled to the ferroelectric distortions. On the other hand, P_y decreases with increasing compressive strain. The same trend is also observed for TiO_2-terminations. At a strain of $\varepsilon_{yy} = -0.05$, the magnitudes of P_x and P_z begin to increase at all sites, whereas they remain nearly unchanged under tension. This suggests that the polarization direction rotates from the y direction to the $x - z$ plane at the critical

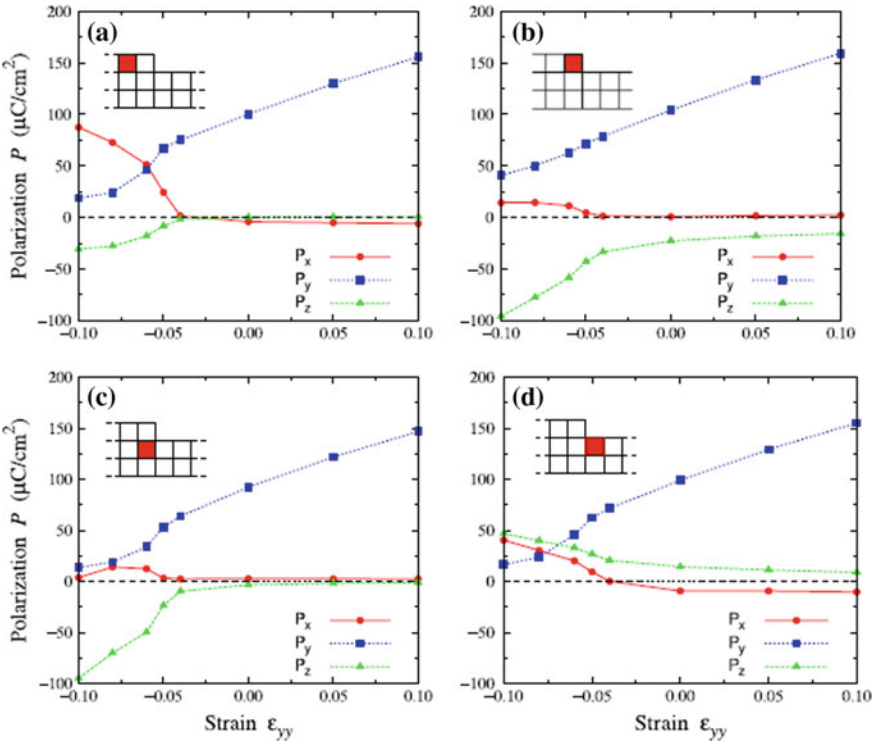

Fig. 5.14 Site-by-site local polarization as a function of strain ε_{yy} for the PbO-terminated surface step: **a** upper terrace, **b** step site, **c** inside, and **d** lower terrace. Reprinted with permission from Ref. [31]. Copyright 2010 by Institute of Physics

strain. The change in P_x and P_z differs by site, indicating that the spontaneous polarization is not uniform in the $x - z$ plane near-surface steps.

Figure 5.15 depicts the vector field of the local polarization distribution of the PbO-terminated surface step at a strain of $\varepsilon_{yy} = -0.05$. The polarization is aligned in the $-z$ direction on the left side of the surface step, whereas the opposite rotation occurs (in the $+z$ direction) on the right side. This indicates that the surface step divides into domains on the formation of a 180° domain wall along the surface step edge, as indicated by the dotted-dashed line in the schematic. This result is consistent with the experimental observation that 180° stripe domains form along the surface step edges in $PbTiO_3$ thin films epitaxially grown on $SrTiO_3$(001) substrates [37]. This comparison is reasonable because $PbTiO_3$ thin films should have the same geometry as the substrate and the lateral lattice parameters of the simulation model under the strain, 3.86 and 3.84 Å in the x and y directions, respectively, are nearly the same as the lattice parameter of the substrate $SrTiO_3$ of a = 3.85 Å. The local polarization tends to lie along the surface (the x direction) in both the upper and lower terraces, whereas the polarization is aligned in the z direction within the film. As a result, the polarization forms a flux along the surface step and the domain

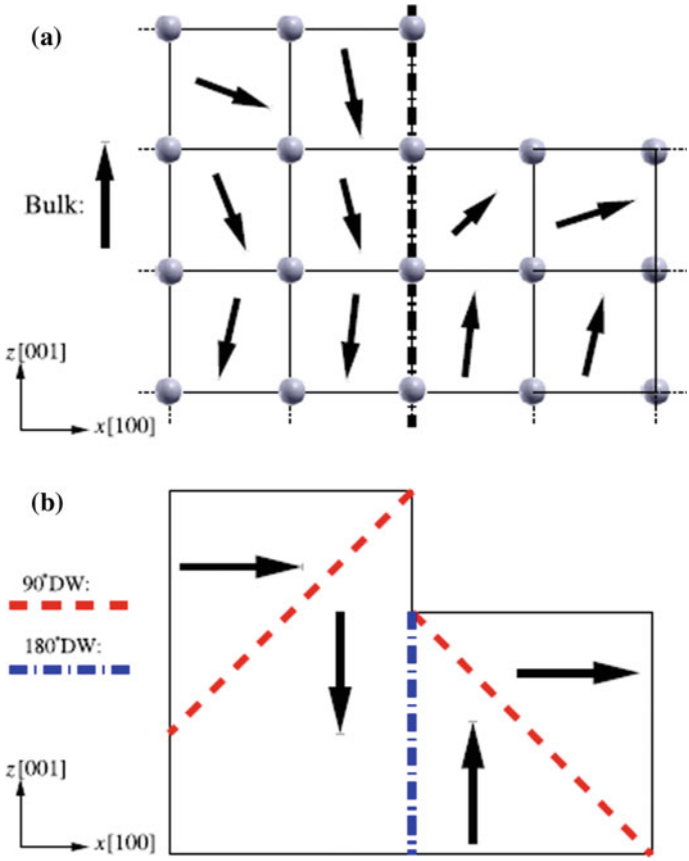

Fig. 5.15 Vector field of the local polarization distribution of the PbO-terminated surface step at a strain of $\varepsilon_{yy} = -0.05$. Reprinted with permission from Ref. [31]. Copyright 2010 by Institute of Physics

wall. In terms of the polarization distribution pattern, the domain configuration at the surface step seems to consist of not only the 180°, but also 90° domain walls (see the schematic in Fig. 5.15). This ferroelectric polydomain structure is not a special case for PbO-terminations but is formed for the TiO_2-terminated surface steps under compression.

5.2.3 Ferroelectric Thin-Film Capacitor

As was mentioned in the previous section, a metal electrode layer attached on the surface of ferroelectric substrate compensates for surface charges due to the screening effect of the electrode, which substantially reduces the depolarizing field

and facilitates the ferroelectric polarization. A ferroelectric film sandwiched by electrode layers has been of much interest not only because the screening effect allowing ferroelectric instability in an ultrathin film but also because the multilayer component works as a thin-film capacitor that has potential application to microelectric devices such as the nonvolatile ferroelectric random access memory (FeRAM). As the smallness (small thickness) gives tremendous advantage to the microelectric devices, it has been a central issue how thin the ferroelectric capacitor can be. To estimate the critical thickness of the ferroelectric film between electrodes (i.e., a thickness below which the spontaneous ferroelectric polarization is lost), theoretical investigations using first-principles calculations have been performed.

A pioneering effort of this matter can be found in a study by Junquera and Ghosez [38]. The work was stimulated by earlier experimental observations that identified ferroelectric ground states in perovskite oxide films and crystalline copolymer films with the thickness down to 10–40 Å [39, 40]. To address the issue of ferroelectric instability in the sandwiched electrode-perovskite-electrode structure, the influence of a real metal–perovskite interface, including the finite screening length of the electrode, the interface chemistry, and the strain conditions imposed by the substrate, must be accurately evaluated, which can be done by the first-principles approach. Junquera and Ghosez performed first-principles DFT calculations based on the local density approximation for a $SrRuO_3/BaTiO_3/SrRuO_3$ capacitor as shown in Fig. 5.16. The capacitor was assumed to be grown on a (001)-oriented thick $SrTiO_3$ substrate, meaning that the capacitor component possesses the same in-plane lattice constant as the substrate. The electrical boundary conditions were fixed by putting the electrodes in short circuit. A SrO/TiO_2 interface with $BaTiO_3$ was expected experimentally and assumed in the calculations. The basic unit, which is periodically repeated in space, as shown in Fig. 5.16 corresponds to the generic formula $[SrO\text{-}(RuO_2\text{-}SrO)_n/TiO_2\text{-}(BaO\text{-}TiO_2)_m]$, where n was fixed to five (large enough to avoid interaction between the two interfaces) and m ranged from two to ten. The periodic boundary conditions naturally impose the required short-circuit condition between the electrodes.

The structures were first relaxed imposing a mirror symmetry plane on the central TiO_2 layer, and then, the existence of a ferroelectric instability was searched for by moving $BaTiO_3$ atoms continuously following the displacement pattern of the bulk tetragonal soft mode, ξ, determined for the same tetragonal cell geometry as in the supercell. Figure 5.17 presents the evolution of the energy as a function of the soft-mode distortion for different thickness of the $BaTiO_3$ layer. The figure points out a change of behavior at a critical thickness around $m = 6$ (26 Å). For larger thicknesses, the energy decreases when the atoms follow the bulk soft-mode path, clearly demonstrating that the thin film has a ferroelectric ground state in agreement with experimental results for the same thickness [39]. In contrast, below the critical value the energy is minimized for the paraelectric configuration. This observation alone was not sufficient to make conclusions about the absence of ferroelectricity, but the absence in the thinnest model was confirmed by performing a full atomic relaxation for the $m = 2$ supercell starting from a ferroelectric configuration, which resulted in the atoms moving back to their paraelectric positions.

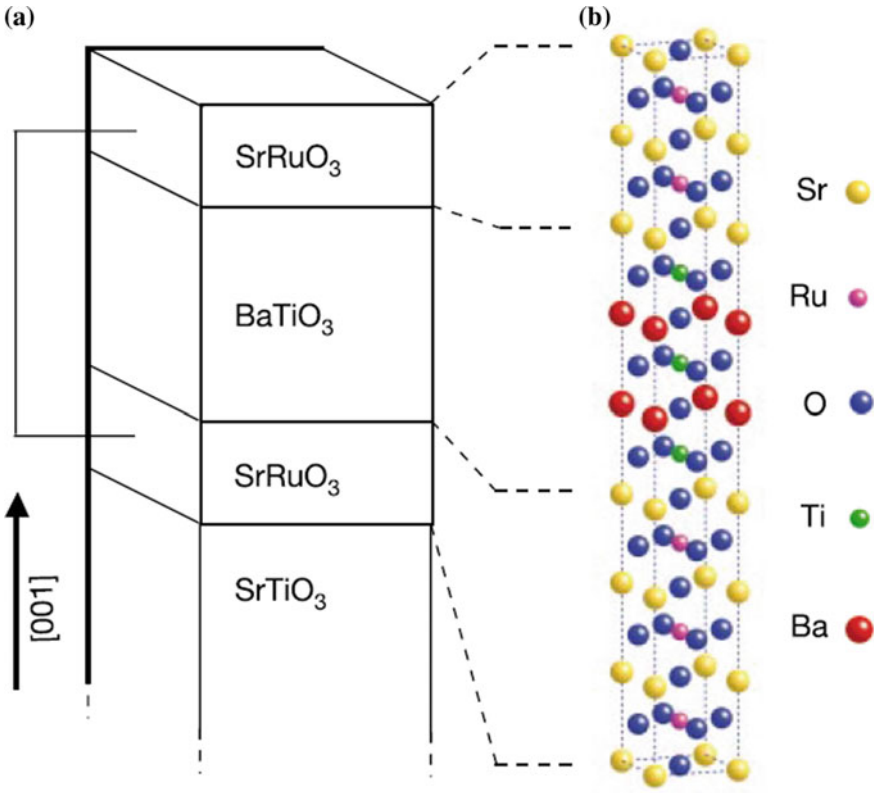

Fig. 5.16 Structure of a typical ferroelectric capacitor. Reprinted by permission from Macmillan Publishers Ltd: Ref. [38], copyright 2003

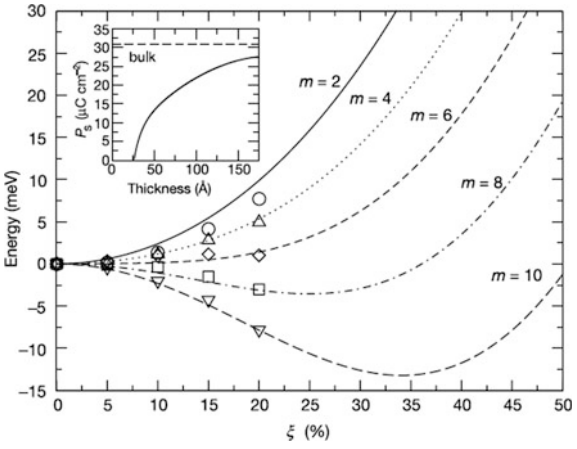

Fig. 5.17 Evolution of the energy as a function of the soft-mode distortion ξ. Reprinted by permission from Macmillan Publishers Ltd: Ref. [38], copyright 2003

The results suggest the existence of a lower limit for the thickness of useful ferroelectric layers in electronic devices.

First-principles DFT calculations to investigate the critical thickness for ferroelectricity in other perovskite thin-film capacitors have also been done. Umeno et al. performed DFT calculations with the local density approximation for $Pt/PbTiO_3/Pt$ films and estimated the critical thickness of the $PbTiO_3$ layer for ferroelectric instability [41]. The electrode-perovskite-electrode multilayers with TiO_2- and PbO-terminations of the perovskite layer are modeled in the form $Pt_3/TiO_2\text{-}(PbO\text{-}TiO_2)_m/Pt_3$ and $Pt_3/PbO\text{-}(TiO_2\text{-}PbO)_m/Pt_3$, respectively. This study took into account the effect of the in-plane lattice constant a to be used for the thin-film geometries, because an artificial compression of the in-plane lattice constant will cause an increase of the c/a ratio, which will enhance the preference of the ferroelectric state. Under a predetermined in-plane lattice constant, the cubic paraelectric (PE) state was obtained by a full atomic relaxation with imposed centrosymmetry. To determine the ferroelectric (FE) state, a soft-mode distortion of the atoms was superimposed to the PE configuration followed by a full relaxation of all atomic positions normal to the interfaces. From these calculations, the energy difference between the FE and PE states was obtained as a function of the in-plane lattice constant as shown in Fig. 5.18. At the theoretical equilibrium lattice constant of $SrTiO_3$ ($a = 3.845\,\text{Å}$), which is often used as a substrate to grow stress-free FE $PbTiO_3$ films, it was found that the $m = 4$ PbO-terminated films (~ 16 Å) become ferroelectric. The $m = 4$ TiO_2-terminated films stay paraelectric unless the lattice parameter is compressed to be below 3.815 Å. Using the extrapolation formula, $\Delta E_{\text{est}}(m = 4) = [2\Delta E(m = 2) + 2\Delta E(m = \infty)]/4$, an estimate of ΔE for the $m = 4$ geometry. The result, plotted by short dashed lines in Fig. 5.18, is nearly identical to the full $m = 4$ DFT calculations. This extrapolation scheme indicates that the critical thickness of the TiO_2-terminated film is around $m = 6$ or 24 Å.

Importantly, in Ref. [41], the effect of the exchange–correlation energy functional on the result was investigated to address the discrepancy with the DFT study by Sai et al. [42], in which the persistence of ferroelectricity for $Pt/PbTiO_3/Pt$ films down to one unit cell (4 Å) was reported. Careful tests with different types of pseudopotentials and density functionals revealed that this discrepancy was due to insufficiencies of the widely used GGA approximations (PW91 and PBE functionals). The lattice parameter of bulk $PbTiO_3$ is fairly well reproduced by both LDA and GGA within 1% deviation from the experimental value. In contrast, the tetragonal distortion (c/a) obtained by the GGA functionals is considerably larger, while the LDA result agrees very well with experiment. The FE displacements of the atoms again are much larger in the GGA than in the LDA, as is the energy difference between the tetragonal FE and cubic PE states. Overall, these results demonstrated that LDA clearly gives more realistic results for the structural parameters of tetragonal bulk phase of $PbTiO_3$ compared to experiment, whereas the GGAs substantially overestimate the FE distortions and favor by a great amount the stability of the FE state.

Umeno et al. extended the simulation scheme for investigating the critical thickness of ferroelectric instability at varying lattice parameters to $SrRuO_3/PbTiO_3/$

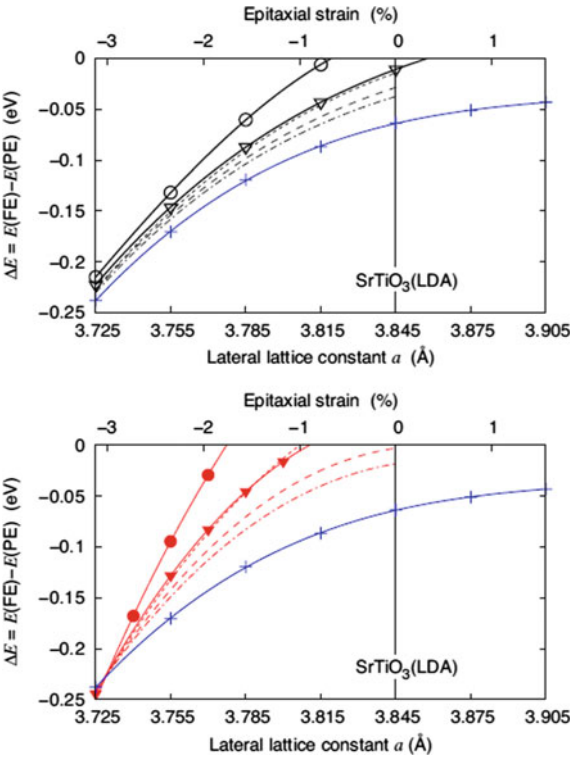

Fig. 5.18 Ferroelectric stability ΔE per perovskite unit cell of $PbTiO_3$ layers for different film thicknesses m as a function of the in-plane lattice constant a. Reprinted with permission from Ref. [41]. Copyright 2006 by American Physical Society

$SrRuO_3$ and $Pt/PbTiO_3/SrRuO_3$ models [43]. The asymmetric capacitor like the latter one has a particular importance for applications because dissimilar materials are usually used for the bottom and top electrodes because different properties are required for the two electrodes. Metals having resistance against oxidation, e.g., Pt, are suitable for the top electrode as it protects the capacitor in gaseous environments. In contrast, the bottom layers must have good properties for epitaxial growth, and the perovskite $SrRuO_3$ is best suited because of its close structural relationship and lattice match to the ferroelectric perovskites $PbTiO_3$ and $BaTiO_3$ and to typically used substrates like $SrTiO_3$, enabling epitaxy on a perovskite substrate with well-ordered atomic arrangement.

Figure 5.19 shows the ferroelectric stability of the asymmetric $Pt/PbTiO_3/SrRuO_3$ capacitor. Because of the asymmetry in the structure, the FE states with opposite polarization directions exhibit different energies relative to the PE reference state. The structure with the oxygen atoms being shifted toward the Pt layer and metal atoms (Ti and Pb) toward $SrRuO_3$ (i.e., negatively charged Pt side and positive $SrRuO_3$) exhibits the lower FE stability. The FE stability decreases with increasing lateral lattice parameter, and the FE state with the lower stability is no longer stable at $a \geq 3.755$ Å in the $m = 2$ capacitors. While $m = 4$ is sufficient for both the PbO-terminated $Pt/PbTiO_3/Pt$ and $SrRuO_3/PbTiO_3/SrRuO_3$ capacitors to have stable ferroelectric states at $a = 3.845$ Å, the ab initio calculated cubic lattice

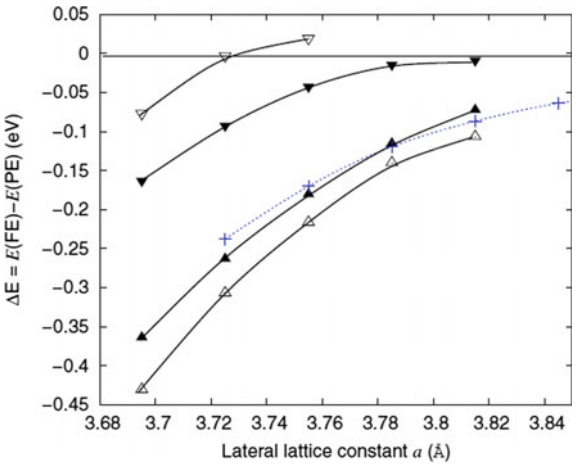

Fig. 5.19 Ferroelectric stability of $Pt/PbTiO_3/SrRuO_3$ capacitors with PbO-terminated perovskite layers. Reprinted with permission from Ref. [43]. Copyright 2009 by American Physical Society

parameter of $SrTiO_3$, thicker layers are expected to be necessary for the $Pt/PbTiO_3/SrRuO_3$ capacitor. This is because the ferroelectric state with the polarization pointing to the $SrRuO_3$ side is suppressed by the combination of the dissimilar electrodes, while the polarization in the opposite direction is stabilized.

In the study, the electronic Schottky barriers of the capacitors were evaluated by measuring the potential level located in the most central part of the $PbTiO_3$ slab. It was found that the epitaxial lattice strain does not strongly affect the Schottky barriers of the symmetric capacitors. The contribution of the electrical field originating from the asymmetric electrodes influences the Schottky barriers significantly. The switching of the polarization in the asymmetric capacitor changes the Schottky barrier height from $PbTiO_3$ to the $SrRuO_3$ electrode by about 1.0 eV, while the Schottky barrier between Pt and $PbTiO_3$ is rather insensitive to the polarization state of $PbTiO_3$ (see Fig. 5.20).

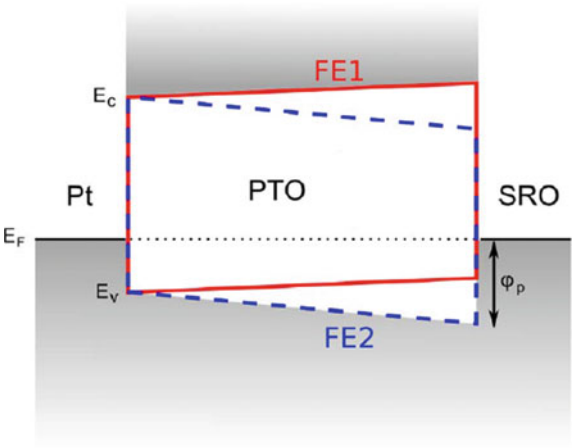

Fig. 5.20 Variation in the Schottky barrier height by switching of the FE polarization in $PbTiO_3$. Reprinted with permission from Ref. [43]. Copyright 2009 by American Physical Society

5.3 Nanowire and Nanotube: One-Dimensional Structure

5.3.1 FE Structure in Perovskite Nanowire

Ferroelectric perovskites with one-dimensional structures such as nanorods, nanowires, and nanotubes have recently been drawing a lot of attention due to their potential use in various devices including nonvolatile random access memory, energy-harvesting devices, and advanced sensors. As is nicely summarized in a review by Rørvik et al. [44], a number of studies have been done for the synthesis of one-dimensional nanostructures of ferroelectric perovskites.

Ferroelectric properties of $BaTiO_3$ and $PbTiO_3$ nanowires were studied with first-principles DFT calculations by Cai et al. [45, 46]. In their study, nanowires in a cylindrical shape with varying diameter were investigated. When the diameter is above the critical value, the nanowires possess FE polarization along the axis of the wire (z axis), while the total ferroelectricity in the perpendicular direction (x and y axes) was zero because displacement of atoms in the $x - y$ plane is axisymmetric. Figure 5.21 shows the longitudinal z-axis spontaneous polarization as a function of the diameter of the $BaTiO_3$ nanowire. The critical diameter was found to be 14 Å, below which the nanowire lost the spontaneous polarization. In contrast, at a diameter of 20 Å the FE polarization was even enhanced (by 18%) compared to the bulk counterpart. The finding that the nanowire can retain the FE state with the diameter down to 20 Å suggests the possibility of the application of FE nanostructures. Their DFT calculations for the $PbTiO_3$ nanowires revealed that the critical diameter was nearly the same as that of the $BaTiO_3$ nanowire and that the change of the FE polarization with varying diameter was very similar. At a diameter of 18 Å, the FE polarization exhibited the maximum value, which is 26% larger than that of bulk $PbTiO_3$.

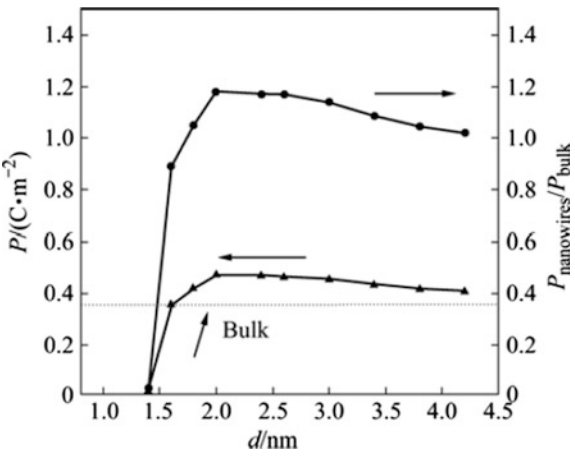

Fig. 5.21 Longitudinal z-axis spontaneous polarization as a function of diameter in $BaTiO_3$ nanowire. Reprinted from Ref. [45], Copyright 2009, with permission from Elsevier

Ferroelectricity in $BaTiO_3$ and $PbTiO_3$ nanowires along [001] (z axis) with (100) and (010) sidewalls (normal to x and y axes, respectively) was investigated by first-principles DFT calculations by Pilania et al. [47, 48]. In Ref. [47], nonstoichiometric (BaO- or TiO_2-terminations for all four sidewalls) and stoichiometric (BaO-terminations for two adjoining sidewalls and TiO_2-termination for the rest) $BaTiO_3$ nanowires of varying cross-sectional sizes and sidewall termination were examined. Stoichiometric nanowires contained an integer number of $BaTiO_3$ unit cells, which were arranged in a $n \times n$ square grid along the plane normal to the nanowire axis (referred to as S-$n \times n$). BaO- and TiO_2-terminated nonstoichiometric nanowires were referred to as BaO- $n \times n$ and TiO_2-$n \times n$, respectively. BaO-$n \times n$ (TiO_2-$n \times n$) nanowires were created by adding BaO (TiO_2) layers to the TiO_2 (BaO) facets of S-$n \times n$ nanowires. Nanowires of $n = 1 \sim 4$ were examined in the study. Figure 5.22 presents the ferroelectric well depth, or the energy difference per $BaTiO_3$ unit between the reference paraelectric structure and the distorted ferroelectric structure, for the stoichiometric nanowires. A positive value in the ferroelectric well depth means that the ferroelectric structure is more favored than the paraelectric one. While the S-1×1 and S-2×2 nanowires exhibited paraelectric structures, the ferroelectric states with polarization along the nanowire axis were stable in S-3×3 and S-4×4 nanowires with the ferroelectric well depths of 12.8 and 13.8 meV, respectively. The magnitude of the ferroelectric distortions was in general smaller than in bulk $BaTiO_3$, while the distortions were much larger at and close to TiO_2-terminated facets than at and close to BaO-terminated ones. From the results, it was concluded that the ferroelectric distortion is stabilized even in nanowires as small as 12 Å thick. It was also found that the band gap is narrowed in $BaTiO_3$ units containing both TiO_2- and BaO-terminated facets.

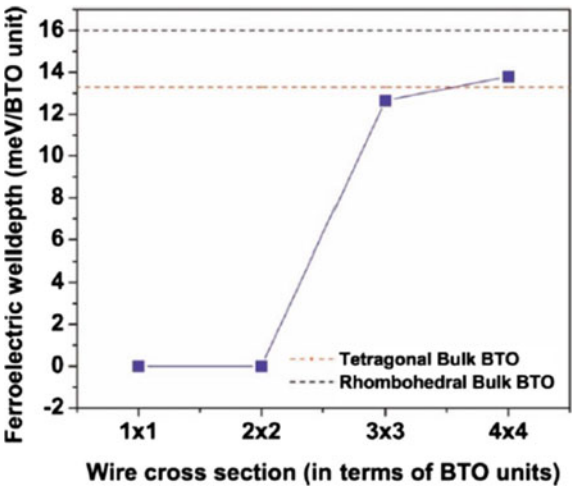

Fig. 5.22 Ferroelectric well depth per $BaTiO_3$ unit for stoichiometric nanowires as a function of wire cross section. Reprinted with permission from Ref. [47]. Copyright 2009 by American Physical Society

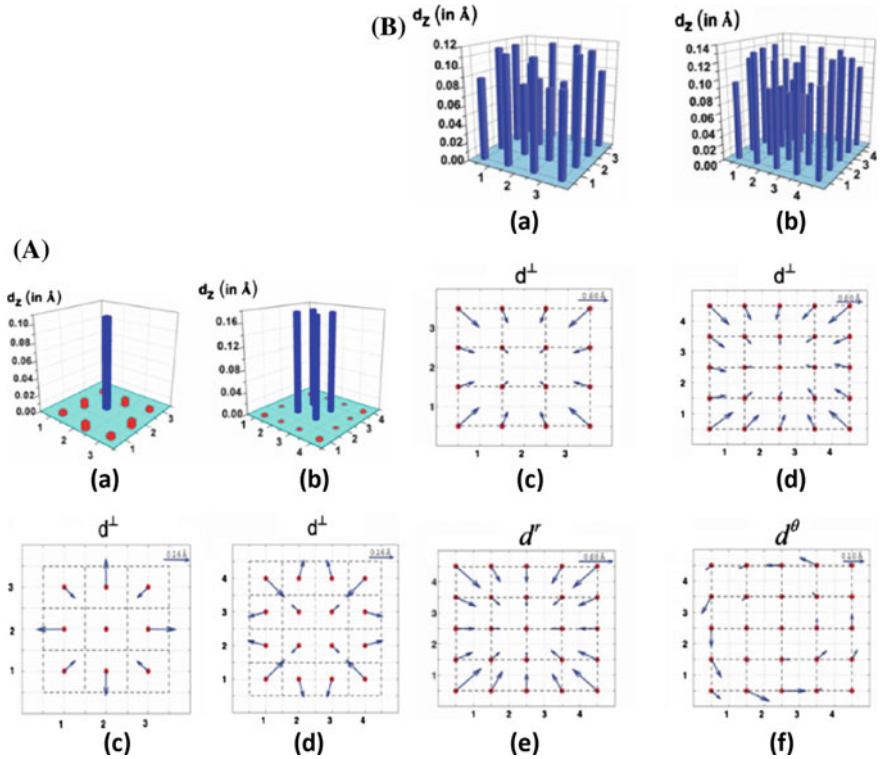

Fig. 5.23 Calculation results of $BaTiO_3$ nanowires. **A** Axial polarization distortion (*a* and *b*) transverse polarization patterns (*c* and *d*) for BaO-3 × 3 and BaO-4 × 4 nanowires, respectively. **B** Axial polarization distortion (*a* and *b*), transverse polarization patterns (*c* and *d*) of TiO_2-3 × 3 and TiO_2-4 × 4 nanowires, respectively. Radial and azimuthal decompositions of the transverse polarization pattern of the TiO_2-4 × 4 nanowire are shown in (*d*). Reprinted with permission from Ref. [47]. Copyright 2009 by American Physical Society

In the case of nonstoichiometric BaO-$n \times n$ and TiO_2-$n \times n$, not only the axial ferroelectric distortion ($d^z \equiv z_{O_1} + z_{O_2} - 2z_{Ti}$) but also the transverse component of the distortion ($d^{\perp} \equiv d^x + d^y; d^x \equiv x_{O_1} + x_{O_2} - 2x_{Ti}, d^y \equiv y_{O_1} + y_{O_2} - 2y_{Ti}$) was considered. Among the BaO-$n \times n$ nanowires, axial polarization was not stable in the BaO-1 × 1 and BaO-2 × 2 models, similar to the stoichiometric case. The axial and transverse distortions in the BaO-3 × 3 and BaO-4 × 4 nanowires are presented in Fig. 5.23A. Clearly, the axial polarization was realized only in the core of the nanowires. It is worth noting that in the outer "shell" region, the sign of the d^z value is opposite to that in the core region. It was thus concluded that the nanowires having BaO-terminated sidewalls display a core–shell polarization structure. In the core region, the magnitude of the distortion is closer to the corresponding bulk tetragonal value of 0.20 Å than was displayed by the stoichiometric nanowires.

The TiO_2-1×1 and TiO_2-2×2 nanowires were found to exhibit no spontaneous polarization. It thus appears that the critical size for ferroelectricity in $BaTiO_3$ nanowires with polarization along the nanowire axis is about 12 Å regardless of stoichiometry or surface terminations. Figure 5.23B shows the axial and transverse components of the ferroelectric distortions for the TiO_2-3×3 and TiO_2-4×4 nanowires. Qualitative difference in the distortions from the case of the BaO-terminated nanowires was found. The axial component was significant in both the TiO_2-3×3 and TiO_2-4×4 nanowires regardless of the location (center or near-surface) of the Ti atom, indicating that the core–shell structure seen in the BaO-terminated nanowires was not present. As for the transverse component of the distortion, it was seen that all Ti atoms tended to move radially inward significantly rather than the "mixed" trend shown in the BaO-$n \times n$ nanowires. The most remarkable polarization behavior was displayed by the TiO_2-4×4 nanowires as shown in Fig. 5.23B(e) and (d), where the transverse component of the ferroelectric distortion was decomposed into radial and azimuthal (vortex) components. As opposed to the 3×3 nanowire, the TiO_2-4×4 nanowire exhibited the vortex component of the transverse ferroelectric distortion.

A similar analysis now for $PbTiO_3$ nanowires was made in Ref. [48]. In this study, only nonstoichiometric nanowires, PbO-$n \times n$ and TiO_2-$n \times n$ with n being $1 \sim 4$, were investigated by first-principles DFT calculations. In the stress-free ground state, $n \times n$ TiO_2-terminated nanowires for $n = 1 \sim 3$ remained paraelectric and did not show any spontaneous polarization (axial or vortex). In contrast, the ground state of the 4×4 TiO_2-terminated nanowire turned out to be a pure vortex state. The local polarization vector forms a closed loop in the plane normal to the nanowire axis, showing a zero net polarization along the x, y and z directions and a nonzero moment of polarization along the axial direction. On the other hand, in relaxed state all four $n \times n$ PbO-terminated $PbTiO_3$ nanowires exhibited a significant axial polarization as listed in Table 5.1. The PbO-terminated nanowires showed zero toroidal moment. Table 5.1 contains the preferred polarization state and the optimal c lattice parameter along the axis at equilibrium, where a clear

Table 5.1 Polarization states and c lattice parameters at equilibrium for $PbTiO_3$ nanowires. "P" and "F" represent the paraelectric and ferroelectric ground states, respectively. Reprinted with permission from Ref. [48]

	Nanowire size				Bulk
	1×1	2×2	3×3	4×4	
TiO_2-terminated					
Ground state	P	P	P	F (vortex)	F
c (Å)	3.39	3.57	3.68	3.76	4.03
PbO-terminated					
Ground state	F (axial)	F (axial)	F (axial)	F (axial)	F
Axial polarization ($\mu C/cm^2$)	103.05	100.34	92.93	90.53	84.50
c (Å)	3.92	3.97	3.99	4.01	4.03

correlation between the polarization state and c can be seen. The effect of strain on axial polarization and toroidal moment was also investigated in Ref. [47], which will be explained in the next subsection.

Pilania and Ramprasad also evaluated dielectric permittivity of ultrathin $PbTiO_3$ nanowires [49]. In the study, they proposed an efficient method to compute the dielectric permittivity of nanostructures by combining first-principles density functional perturbation theory with effective medium theory. The dielectric permittivity along the off-axis directions was predicted to reduce in the nanowires from that in bulk. On the other hand, along the axial direction, the dielectric permittivity showed a significant enhancement over the corresponding bulk value. Their calculations also suggested that the nanowires with unconventional vortex-type polarization states were expected to have an increased dielectric response as compared to those with conventional uniform axial polarization.

5.3.2 *FE Perovskite Nanotube*

Ferroelectric perovskite nanotubes have been drawing tremendous attention due to their unique properties and potential technological applications. Various fabrication methods have been therefore intensively studied [50], and peculiar properties that arise in perovskite nanotubes have been investigated experimentally [51, 52]. In such a context, theoretical studies based on first-principles methods have also been performed to reveal the complicated mechanism behind the unique properties of ferroelectric perovskite nanotubes.

First-principles DFT calculations of structural and ferroelectric properties of ultrathin $PbTiO_3$ nanotubes were carried out by Shimada et al. [53]. The lower size limit of the thickness for ferroelectric instability in a thin film does not apply for nanotubes with the same materials and thickness due to internal strain caused by the nanotube geometry, i.e., compression (tension) on the side of the inner (outer) wall that occurs during rolling up the thin-film layer into a tube shape. In fact, the DFT calculations in Ref. [53] revealed the absence of the intrinsic critical size for ferroelectricity in the $PbTiO_3$ nanotube. In the study, the nanotube was modeled by rolling up a (100) film with a single-unit-cell thickness about the [001] axis (a single-wall nanotube). The axial, circumferential, and radial directions were therefore [001], [010], and [100], respectively. Two possible configurations of the outside–inside wall stacking sequence are considered as the PbO/TiO_2 (PbO-outside) and the TiO_2/PbO (TiO_2-outside) layers. Nanotubes with different radii (the number of perovskite unit cells around the circumference $N_{NT} = 10 \sim 22$) were simulated. Figure 5.24 shows total energy and cohesive energy of the nanotubes as a function of tube curvature $1/R_{NT}$. Here, positive and negative values represent the curvature of the PbO-outside and TiO_2-outside nanotubes, respectively. The cohesive energy was positive for all nanotubes, meaning that the nanotubes possess stable structures. The PbO-outside nanotubes were energetically more favored than the thin film, while the TiO_2-outside nanotubes were

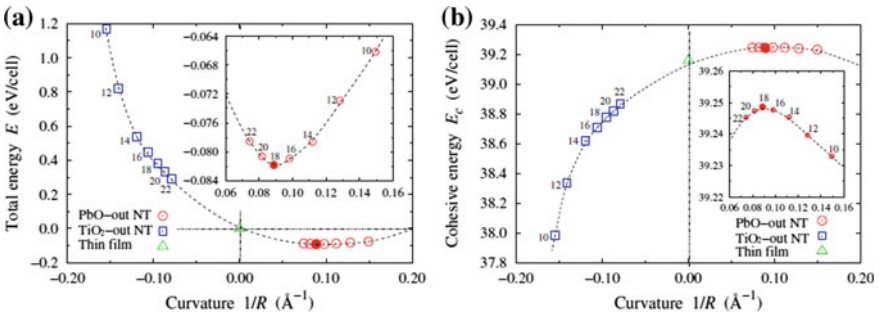

Fig. 5.24 Calculation results of (100) $PbTiO_3$ nanotubes. **a** Total energy per perovskite unit cell as a function of nanotube curvature. Positive and negative values represent the curvature of PbO-outside and TiO_2-outside nanotubes, respectively. **b** Cohesive energy as a function of nanotube curvature. The numbers next to the points denote the number of perovskite unit cells. Reprinted with permission from Ref. [53]. Copyright 2012 by American Physical Society

unfavorable. The highest cohesive energy of 39.25 eV/cell was found at $N_{NT} = 18$ (PbO-outside, $1/R_{NT} = 0.089 Å^{-1}$), while that for bulk $PbTiO_3$ is 40.61 eV/cell, meaning that the nanotube was metastable.

The atomic structure of the $N_{NT} = 18$ PbO-outside nanotube is shown in Fig. 5.25. Buckling of the tube sidewall is observed, especially in the inside TiO_2 layer, which is clearly induced by periodic clockwise and counterclockwise tilting of oxygen atoms in the circumferential direction, namely AFD instability. The AFD instability in $PbTiO_3$ bulk is known to occur under high pressure, which explains the occurrence of the AFD instability in the nanotubes; i.e., curvature in the nanotubes with respect to the flat thin-film geometry results in the inside TiO_2 layer being subject to relatively high compression, leading to the emergence of the AFD rotations. In fact, The AFD rotations were found in all the PbO-outside nanotubes, whereas they were not found in the TiO_2-outside nanotubes. The emergence of the

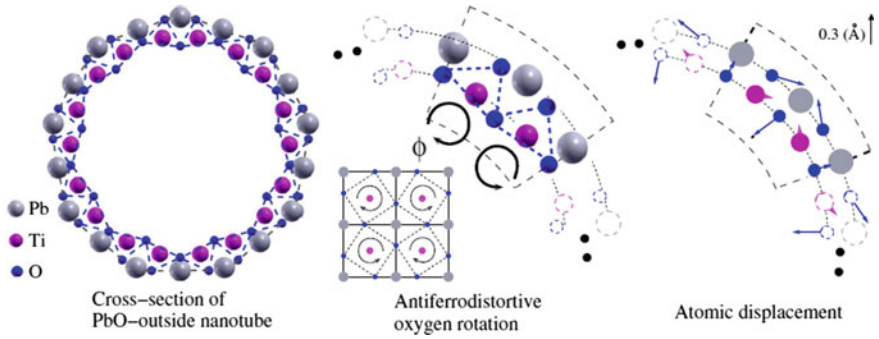

Fig. 5.25 Cross section of the stable PbO-outside $N = 18$ (100) $PbTiO_3$ nanotube. Reprinted with permission from Ref. [53]. Copyright 2012 by American Physical Society

AFD rotation plays a central role in stabilizing the nanotubular structure, which was confirmed by calculations excluding the AFD displacement in the PbO-outside $N_{NT} = 18$ nanotube, which resulted in the dramatic increase in the total energy.

A spontaneous polarization of 5.12 $\mu C/cm^2$ along the axial direction is found in the PbO-outside $N_{NT} = 18$ nanotube, which is one order of magnitude smaller than that of bulk, 78.6 $\mu C/cm^2$. Figure 5.26 shows the axial FE polarization (P_z) and the AFD rotation angle (ϕ) as functions of the tube curvature. Obviously, the FE

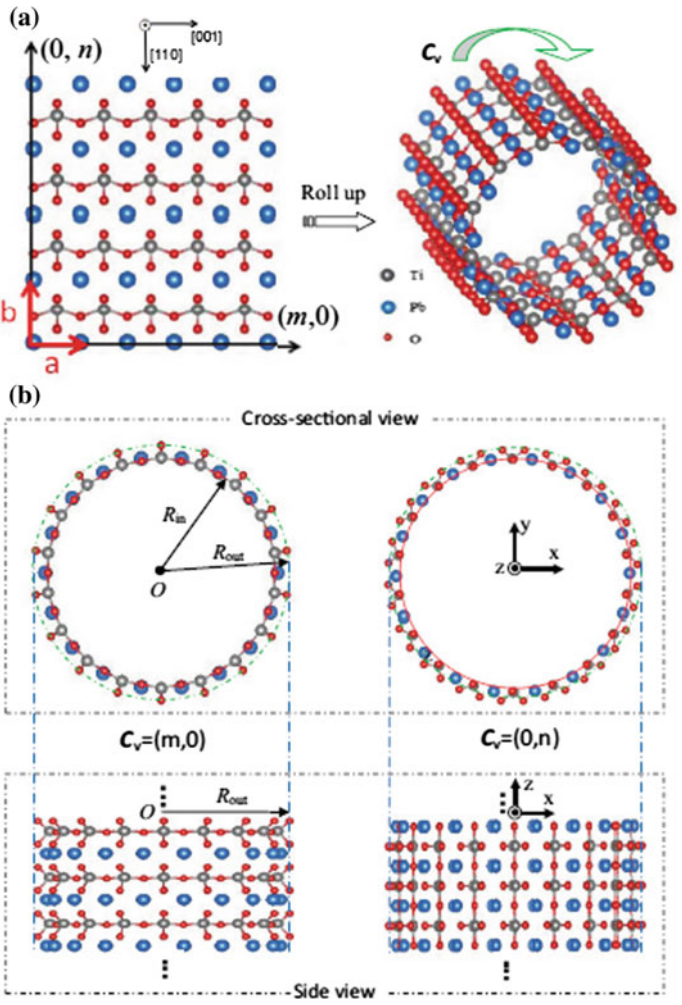

Fig. 5.26 Configurations of (110) $PbTiO_3$ nanotubes. **a** Arrangement of atoms in nanosheets, which is to be rolled up to form (110) $PbTiO_3$ nanotubes. **b** Cross-sectional and side views of nanotubes with different chiralities. Reprinted with permission from Ref. [54]. Copyright 2012 by American Physical Society

polarization and AFD rotation are directly coupled to each other. That is, the polarization increases smoothly as the AFD rotation is enhanced with increasing curvature (compressing the inner TiO_2 layer) for the PbO-outside nanotubes, whereas the TiO_2-outside nanotubes, which does not involve any AFD rotation, are all paraelectric. These results suggest that ferroelectricity in the ultrathin $PbTiO_3$ nanotubes is realized owing to the presence of the AFD rotation.

The ferroelectric properties of $PbTiO_3$ nanotubes with different rolling directions were investigated by Wang et al. [54]. In their study, single-wall nanotubes folded from (110) nanosheets with a specific chirality of (m_{NT}, n_{NT}) were examined by DFT calculations (see Fig. 5.26). The study covered only two rolling directions, namely $(m_{NT}, 0)$ and $(0, n_{NT})$, as nanotubes with other chiral vectors require simulation cells with much larger dimensions along the tube axis. Figure 5.27a presents the relationship between total energy and curvature of the nanotubes with a rolling direction of $(m_{NT}, 0)$. Positive and negative curvatures correspond to PbTiO-outside and O-outside nanotubes, respectively. The total energies of PbTiO-outside nanotubes are higher than that of the flat thin film, showing that the PbTiO-outside nanotubes are unstable. In contrast, O-outside nanotubes were found to be more stable than the flat film. The lowest energy is located at $1/R_{NT} = -0.097\,\text{Å}^{-1} (m_{NT} = 18)$. It is interesting to note that the curvature of the O-outside (110) nanotube with the lowest energy is the same as that of the PbO-outside (100) $PbTiO_3$ nanotube. However, the minimum total energy (-0.343 eV/cell) for the (110) nanotube is much lower than that (-0.082 eV/cell) for the (100) nanotube, indicating that the $PbTiO_3$ nanotube rolled from (110) nanosheets is more stable than that from (100). Figure 5.27b shows total energy as a function of curvature for the $(0, n_{NT})$ nanotubes, where the trend opposite to the case of the $(m_{NT}, 0)$ nanotubes was found; the total energies of the PbTiO-outside nanotubes were lower than that of the flat film.

Ferroelectricity of the (110) $PbTiO_3$ nanotubes is shown in Fig. 5.27c and d. It was found that spontaneous polarizations existed only in nanotubes with a rolling direction of $(m_{NT}, 0)$. A remarkable axial polarization existed for the O-out nanotubes even though the tube wall was much thinner than the critical thickness for ferroelectricity of the flat thin film. This indicates the absence of the FE critical thickness in the (110) $PbTiO_3$ nanotubes, which was also the case in the (100) $PbTiO_3$ nanotubes. The axial polarization in the (110) $PbTiO_3$ nanotube with a curvature of $-0.097\,\text{Å}^{-1} (m = 18)$ was calculated to be 42.42 $\mu C/cm^2$, which is more than half of the bulk value of 78.6 $\mu C/cm^2$ and one order of magnitude higher than that in the (100) $PbTiO_3$ nanotubes. The polarization in the (110) nanotubes stems from breaking the symmetry of oxygen octahedra along the axial direction due to the asymmetric cutting of the unit cell. This symmetry breaking results in a large tilt of oxygen octahedra in the longitudinal section of the nanotubes, as is indicated by the angle θ in Fig. 5.28. The curvature dependence of the polarization and tilting angle in Fig. 5.27c and d clearly demonstrates the strong coupling between FE distortion and AFD distortion. The FE-AFD coupling is responsible for the stability of the FE distortion in the nanotubes, which gives rise to anomalous ferroelectricity in the (110) $PbTiO_3$ nanotubes.

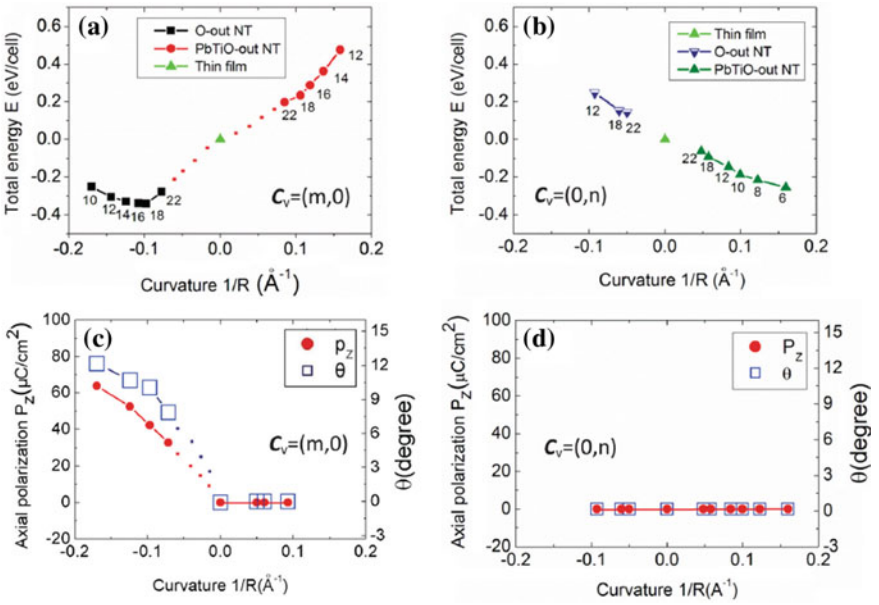

Fig. 5.27 Curvature dependence of total energy in (110) $PbTiO_3$ nanotubes (**a** and **b**), and curvature dependence of spontaneous polarization and tilting angle of oxygen octahedra (**c** and **d**). Reprinted with permission from Ref. [54]. Copyright 2012 by American Physical Society

5.3.3 *Strain Effect in Nanowire and Nanotube*

The crucial role of axial tensile strain on ferroelectricity in $PbTiO_3$ nanowires with cross sections of $1 \times 1, 2 \times 2$ and 3×3 unit cells (sidewalls of (100) and (010) surfaces) was investigated by Shimada et al. using DFT calculations [55]. Figure 5.29 shows the axial polarization averaged over the cross section $(\bar{P})$ in the PbO-terminated and TiO_2-terminated nanowires as a function of the tensile strain (ε_{zz}). The averaged polarization increases almost linearly with respect to the tensile strain for all the PbO-terminated nanowires, suggesting that the axial tensile strain enhances the ferroelectricity. As for the TiO_2-terminated nanowires, which are initially paraelectric, ferroelectricity appears under axial tension. Critical tensile strains for ferroelectricity in the TiO_2-terminated nanowires were found to be 0.15, 0.08, and 0.04 for the $1 \times 1, 2 \times 2$ and 3×3 models, respectively. Figure 5.30 shows the change in atomistic and electronic configurations on the PbO (010) planes in the PbO- and TiO_2-terminated 3×3 nanowires under axial tension. At equilibrium in the PbO-terminated nanowire, O atoms are displaced in the $-z$ direction from their ideal lattice sites and the covalent Pb–O bonds are formed, indicating ferroelectric distortion. During axial tension, all the Pb–O covalent bonds are sustained and the charge density at these sites increases. On the other hand, in the TiO_2-terminated nanowires, which present no ferroelectric distortion at

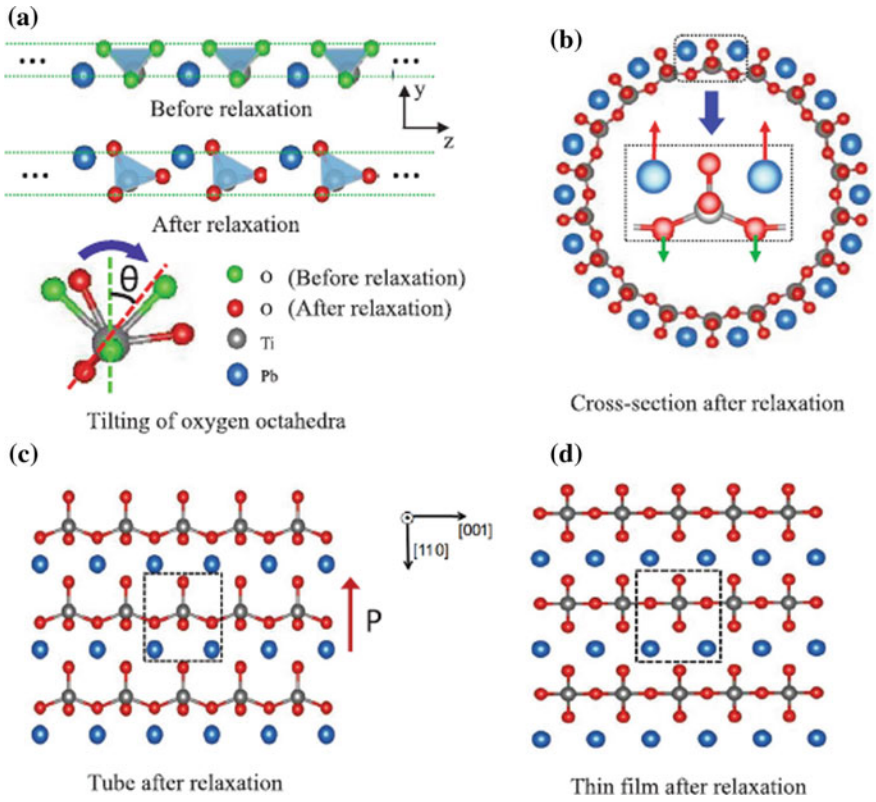

Fig. 5.28 Atomic configuration after relaxation for the O-outside (110) $PbTiO_3$ nanotube with the lowest energy $(m = 18)$. **a** Tilting of oxygen octahedra in the longitudinal section. **b**, **c** Cross-sectional and expanded side views of the nanotube, respectively. **d** Side view of thin film. Reprinted with permission from Ref. [54]. Copyright 2012 by American Physical Society

equilibrium, the upper (lower) Pb–O bond is strengthened (weakened) when the axial strain exceeds the critical strain for ferroelectricity, which is associated with the upward displacement of the Pb atom.

In Refs. [47, 48], Pilania et al. examined the effect of axial strain on ferroelectricity in $BaTiO_3$ and $PbTiO_3$ nanowires having a 4×4 cross section. As was already described in Sect. 5.3.2, toroidal moment (vortex structure) in ferroelectric distortion can emerge in the 4×4 nanowires. For the TiO_2-4×4 $BaTiO_3$ nanowire, the toroidal moment was found to reduce with increasing axial strain, showing that the axial tension destabilizes the vortex state of ferroelectricity in the nanotube. Similarly, the destabilizing effect of axial tension on the toroidal moment was found in the $PbTiO_3$ nanotubes. Figure 5.31 shows the ferroelectric behavior of the 4×4 TiO_2- and 4×4 PbO-terminated $PbTiO_3$ nanowires as a function of axial strain. In both cases, axial strain can cause a phase transition between the pure vortex and pure axial polarization states. The 4×4 TiO_2-terminated and PbO-terminated

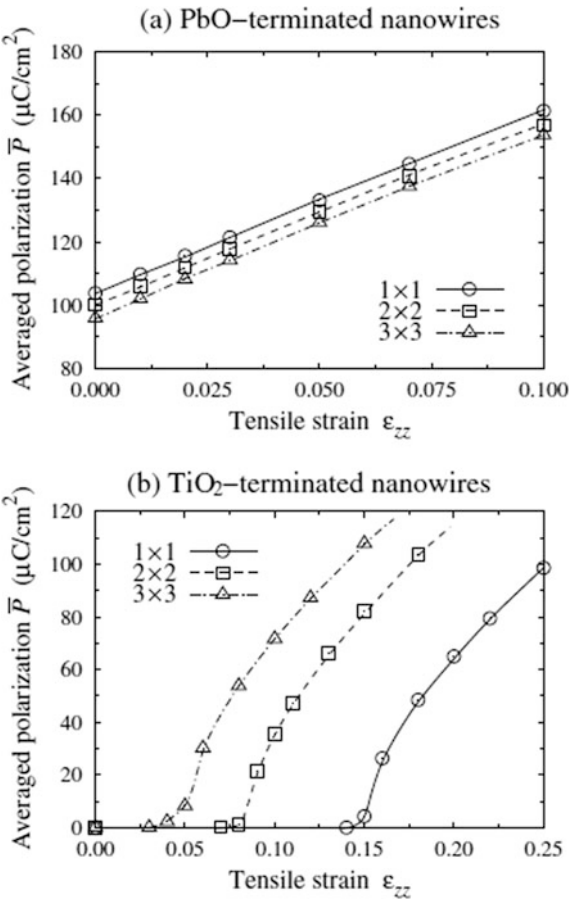

Fig. 5.29 Averaged polarization in $PbTiO_3$ nanowires as a function of tensile strain. **a** PbO-terminated nanowires. **b** TiO_2-terminated nanowires. Reprinted with permission from Ref. [55]. Copyright 2009 by American Physical Society

nanowires take nonrectilinear vortex and rectilinear axial polarizations state at equilibrium, respectively. In the TiO_2-terminated nanowire, an axial tensile strain of about 3% results in the onset of an abrupt phase transition to a pure axial polarization. In contrast, the opposite transition is found in the PbO-terminated nanowire under axial compression (negative strain) and an abrupt phase transition to a pure vortex polarization occurs at an axial strain of about -3.5%. That is, axial tension enhances the axial polarization and axial compression promotes the vortex one in both the TiO_2- and PbO-terminated nanowires, and only the transition strains are different between the two nanowires. Figure 5.32 depicts the local dipoles for each unit cell and displacements of atoms with respect to the corresponding positions in the reference paraelectric state. In the figure, the configurations of the TiO_2-terminated nanowire at 0% strain and the PbO-terminated nanowire at -3.7% strain both present the vortex structure of ferroelectric distortion. It is worth noting that the transition between vortex and axial polarization states occurs at exactly the same

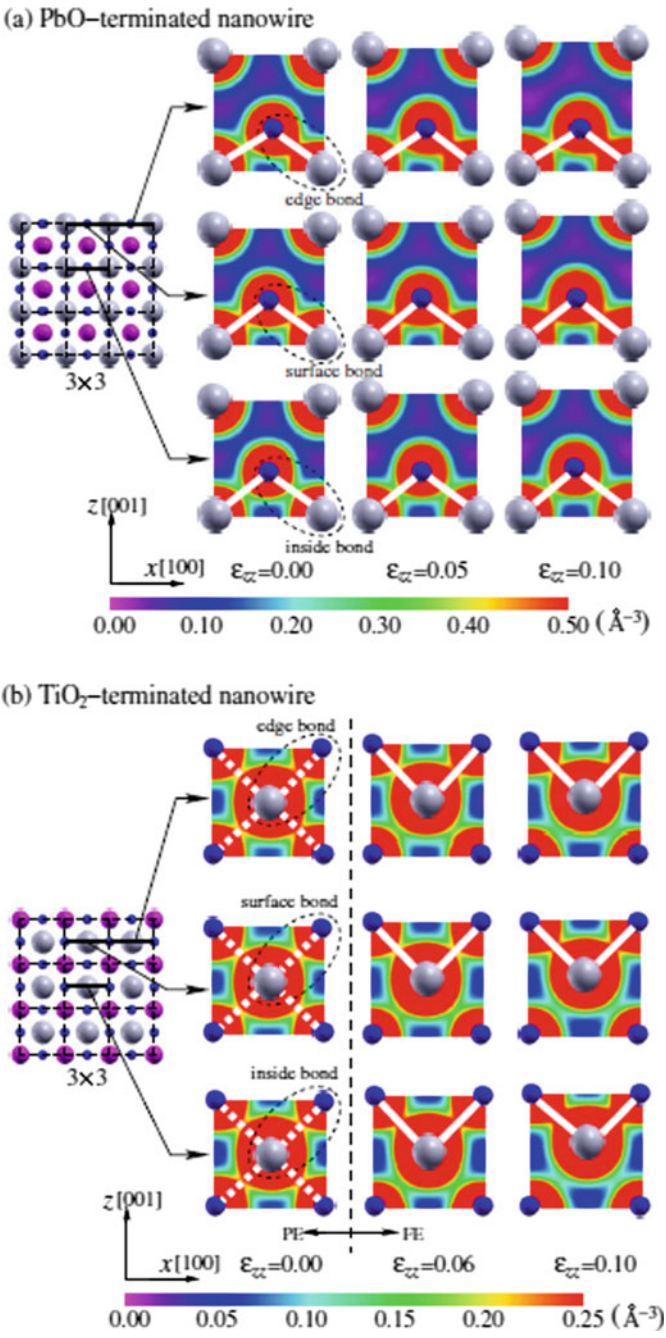

Fig. 5.30 Change in atomic structures and charge density distributions on the PbO (010) planes in **a** PbO-terminated and **b** TiO_2-terminated 3×3 $PbTiO_3$ nanowires under axial tensile strain. The larger and smaller spheres are for Pb and O atoms, respectively. The covalent Pb-O bonds are shown by white lines. Reprinted with permission from Ref. [55]. Copyright 2009 by American Physical Society

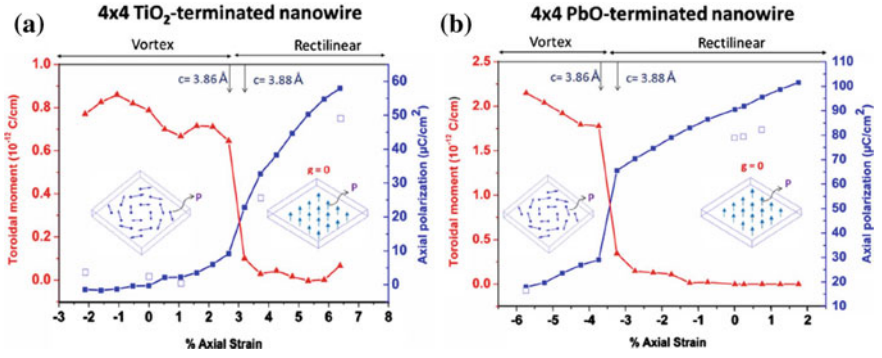

Fig. 5.31 **a** Toroidal moment (*triangle*) and axial polarization (*square*) as a function of axial strain for the 4×4 TiO_2-terminated $PbTiO_3$ nanowire. Hollow squares represent axial polarization calculated with actual Born effective charges instead of bulk Born effective charges. **b** Same as (**a**) but for the PbO-terminated nanowire. Reprinted with permission from Ref. [48]. Copyright 2010 by American Physical Society

value for the longitudinal lattice parameter c(3.87 Å) for both the types of $PbTiO_3$ nanowires, implying that the c value determines the nature of polarization state.

Shimada et al. assessed in Ref. [53] the effect of axial strain on ferroelectricity in the $PbTiO_3$ nanotube rolled up from a (100) nanosheet. Figure 5.33 shows spontaneous polarization and FE orderings in the PbO-outside $N_{NT} = 18$ nanotube as a function of axial strain. Under axial tension, FE distortions appear in the axial direction (FE_{Axial}) and the axial polarization P_z increases smoothly due to the increase in the tetragonal distortion of the perovskite lattice, while no polarization vortex (FE_{Vortex}) exists. The AFD oxygen rotation (ϕ) is enhanced by the application of tensile strain (see inset), which also supports the increase in the axial polarization due to direct AFD-FE coupling. On the other hand, under axial compression, the spontaneous axial polarization disappears at $\varepsilon_{zz} = -0.02$, and the system becomes paraelectric (PE). Further application of compressive strain causes the system to regain polarization at $\varepsilon_{zz} = -0.05$, forming FE_{Vortex} with spontaneous polarization being aligned in the circumferential direction. Their analysis of the band gap energy of the nanotube found that the PbO-outside nanotube exhibits a wider band gap energy of 2.21 eV to be compared with the bulk value of 1.65 eV. The band gap energy of the nanotube was found to be sensitive to axial strain, varying from 2.90 to 1.91 eV with the axial strain ranging from 0.10 to −0.10. In contrast, the change in band gap energy in bulk is relatively small, being less than 0.3 eV in the same strain range.

As an approach other than first-principles analysis to investigate the response of ferroelectric structure to strain, Zhang et al. performed molecular dynamics simulations using an empirical interatomic model [56, 57]. Focusing on the axial polarization, TiO_2-terminated $BaTiO_3$ nanowires with $n \times n$ cross sections $(n = 2 \sim 10)$ were studied with the atomic potentials obtained by Tinte et al. [58]. It

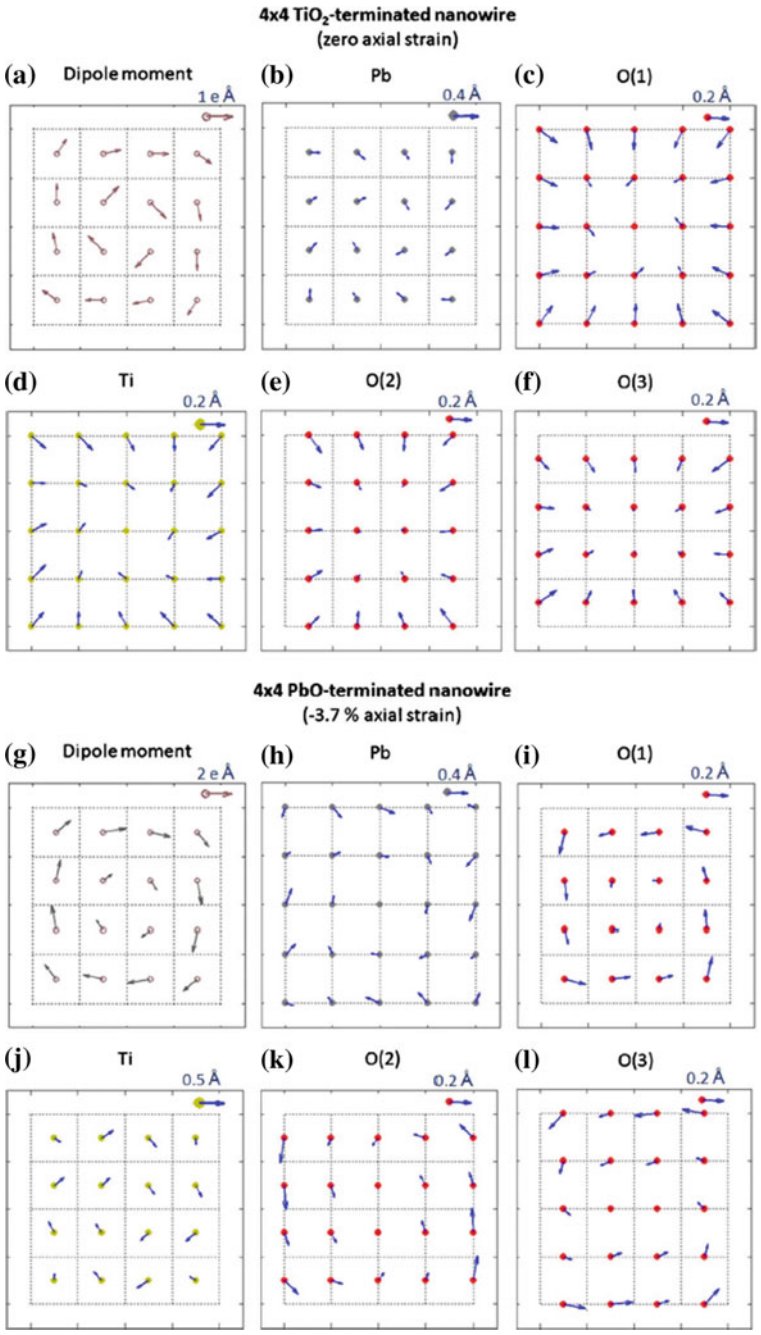

Fig. 5.32 Unit-cell-decomposed dipole moments (**a**) and in-plane displacements of individual atoms with respect to a paraelectric reference state in PbO (**b**, **c**) and TiO_2 (**d**–**f**) transverse planes for the 4×4 TiO_2-terminated $PbTiO_3$ nanowire at zero strain. The counterparts in the PbO-terminated nanowire at a strain of -3.7% are shown in (**g**–**l**). Reprinted with permission from Ref. [48]. Copyright 2010 by American Physical Society

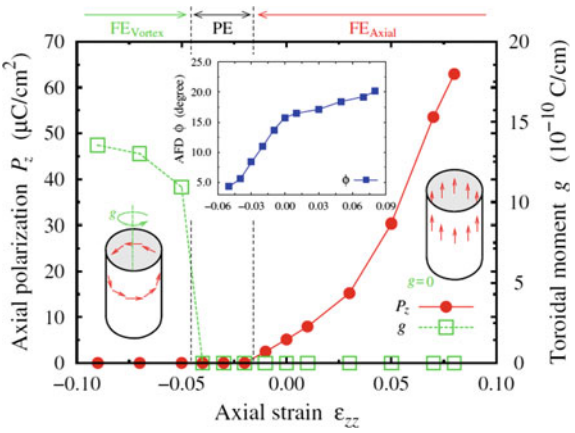

Fig. 5.33 Axial polarization and toroidal moment of the stable PbO-outside (100) $PbTiO_3$ nanotube ($N_{NT} = 18$) as a function of axial strain. The variation in the AFD rotation angle with strain is also shown. Reprinted with permission from Ref. [53]. Copyright 2012 by American Physical Society

was found that the polarization states in response to axial stress consist of three stages: pure FE, pure PE, and mixed FE-PE stages in between. In fact, the $n = 8$ nanotube, which is at the pure FE state at equilibrium, shows transition at a certain (critical) value of compressive stress to the mixed FE-PE stage where PE phase arises from the inner cells. When the compressive axial stress is further increased and exceeds a critical value, the nanowire exhibits the pure PE state. The variations of polarization and strain in response to axial electrical loading were also obtained under the stress-free mechanical boundary condition. The hysteresis loop in the polarization–electrical loading relation and the butterfly loops in the strain–electrical loading relation were calculated, where strong size effects were found. Due to the inhomogeneous polarization distribution in the nanowire, the domain switches layer-by-layer and thus both loops exhibit stepwise shapes.

5.4 Nanodot: Zero-Dimensional Structure

The emergence of toroidal FE structures in ferroelectric nanowires we saw in Sect. 5.3 suggests that such characteristic FE structures stem from the structural low dimensionality and smallness and implies that peculiar FE structures may occur in zero-dimensional nanostructures. It is therefore interesting to FE structures in ferroelectric nanodots, which are representative zero-dimensional nanostructures. From both scientific and industrial points of view, it is urged to reveal the effects of various factors such as the size, boundary conditions, and temperature.

Little investigation of FE structures in nanodots has been done by first-principles calculations or related approaches. Instead, a number of studies have been devoted for this issue based on phenomenological approaches. The effective Hamiltonian approach based on the Landau–Ginzburg–Devonshire theory [13, 14], which is also called a phase-field model in engineering communities, has been a most widely

used model to be applied for FE nanodots thus far. In such a scheme, the effective Hamiltonian is usually constructed based on first-principles calculations, but obviously, it is not a first-principles simulation because the model neither involves electron structure calculations nor contains the degrees of freedom of all atoms. Although such phenomenological approaches are basically not central in the scope of this book, here we give an overview to these efforts.

Prosandeev and Bellaiche investigated vortex domain structures (VDSs) in FE nanodots with various shapes. Their calculations showed that the vortex structure can exist in PZT60 ($PbZr_{0.4}Ti_{0.6}O_3$) nanodots in various shapes as shown in Fig. 5.34. Here, nanodots with the size of $n \times n \times n$ unit cells were examined. The rectangular and pyramidal dots have vortices directed along the shortest c-axis, while the vortex in the spherical dot can be oriented along any $\langle 001 \rangle$ direction. The bigger $(12 \times 12 \times 24)$ rectangular dot has two vortices with the opposite directions in contrast to the case of the shorter $(16 \times 16 \times 8)$ rectangular dot showing only one vortex.

The electrocaloric effect, which is a change in temperature of ferroelectric materials under adiabatic conditions, in the FE nanodots was investigated by Li et al. [60]. Their effective Hamiltonian approach successfully simulated the electrocaloric effect in $PbTiO_3$ nanodots due to the toroidal moment change in the FE structure. The simulation predicted a relatively large adiabatic temperature change (around 6 K) in a $PbTiO_3$ nanodot. It was also suggested that the magnitude of the adiabatic temperature change should decrease with increasing surface tension.

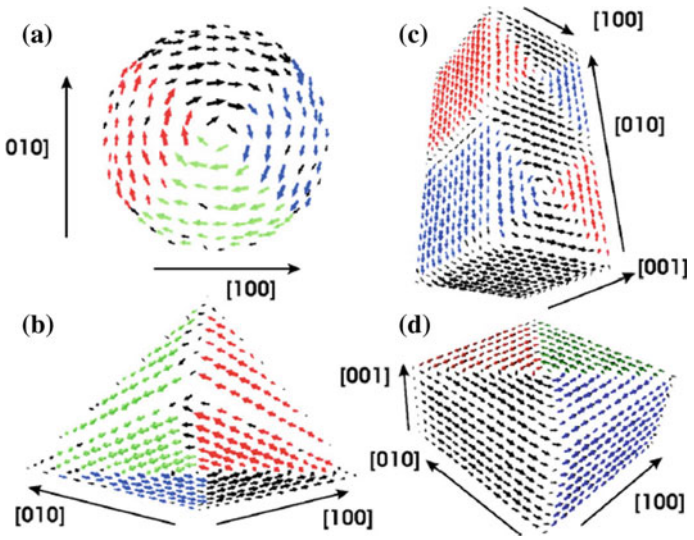

Fig. 5.34 Polarization patterns at 1 K for stress-free PZT60 nanodots. Spherical $16 \times 16 \times 16$ (**a**), pyramidal $16 \times 16 \times 8$ (**b**), rectangular $12 \times 12 \times 24$ (**c**), and rectangular $16 \times 16 \times 8$ (**d**) nanodots. Reprinted with permission from Ref. [59]. Copyright 2007 by American Physical Society

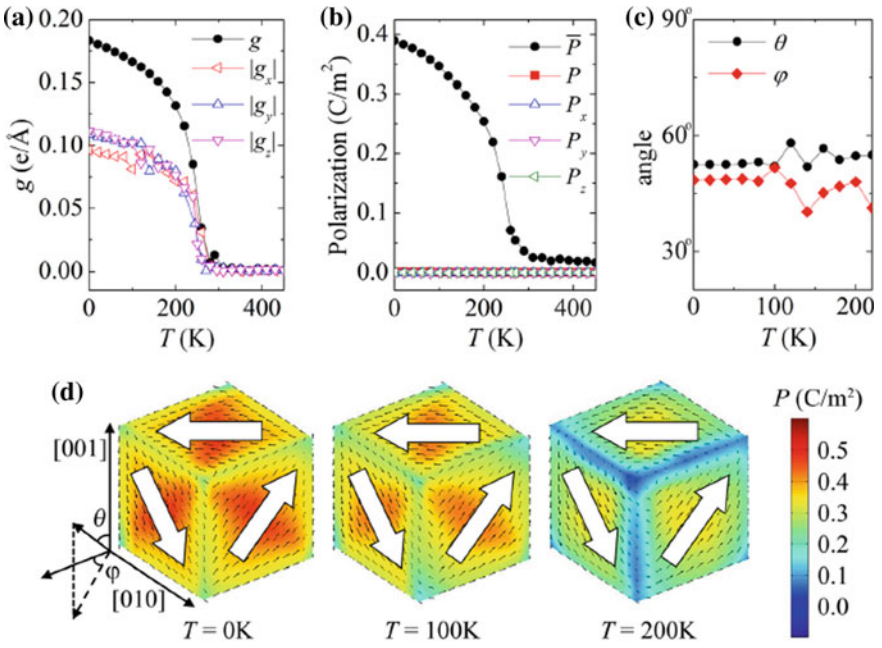

Fig. 5.35 Results of the $12 \times 12 \times 12$ $BaTiO_3$ nanodot under a pressure of -4.8 GPa. The variation of **a** magnitude of toroidal moment, **b** polarization, **c** direction of toroidal moment, and **d** VDSs with temperature. Reprinted with permission from Ref. [61]. Copyright 2013, American Institute of Physics

Various factors affecting VDSs in FE nanodots, such as size, pressure, and temperature, were then extensively investigated using the effective Hamiltonian approach in Refs. [61–63]. Figure 5.35 shows the simulation results of a $12 \times 12 \times 12$ $BaTiO_3$ nanodot. The toroidal FE moment was defined as $\boldsymbol{g} = 1/(2Nv)\sum_i \boldsymbol{r}_i \times \boldsymbol{p}_i$. The magnitude of the toroidal moment is close to zero at high temperature, but when the nanodot is cooled down below 270 K, the magnitude $(\boldsymbol{g})$ becomes nonzero and increases with decreasing temperature, implying a paraelectric–ferrotoroidic phase transition. The average of polarization magnitude of all sites $\bar{P}$ is nonzero when the temperature is lower than about 270 K, whereas the total polarization remains nearly vanishing at all temperatures. As for the size effect in the nanodot, it was found that at a given temperature the toroidal moment increases with increasing size. The paraelectric–ferrotoroidic phase transition temperature (T_{FD}) also increases with increasing size. As the size decreases, the phase transition becomes less abrupt. In a relatively large nanodot model $(16 \times 16 \times 16)$, small additional vortex patterns arose around the midpoint of the edges of the cubic nanodot. The simulation also indicated that both T_{FD} and g decrease significantly with increasing hydrostatic pressure. Overall, the series of their simulations indicated promising controllability of VDS in the nanodots.

5.5 Conclusion

Recent theoretical approaches for ferroelectric structures and properties in low-dimensional nanostructures were reviewed in this chapter, introducing mainly first-principles DFT calculations while briefly covering calculations based on related methods. For two-dimensional nanostructures, first-principles calculations are being used for industrial-oriented issues and direct comparison with experiments, as exemplified by the estimation of critical thickness in ferroelectric thin-film capacitors and closure domain structures in thin films. First-principles calculations are also intensively used for one-dimensional FE nanostructures to find interesting and characteristic FE behavior, suggesting promising application of such nanostructures for novel devices. However, in some cases where consideration of temperature effects is essential, first-principles approaches alone cannot address all problems and a combined approach with empirical or phenomenological models may be necessary. For zero-dimensional nanostructures, first-principles calculations are not yet widely used and phenomenological approaches, though they are often "first-principles based," are still dominant in this field. Application of first-principles calculation methods to the problem of ferroelectric properties in zero-dimensional nanostructures is much expected with fast and ongoing development of computer resources.

References

1. W. Cochran, Adv. Phys. **9**, 387 (1960)
2. Ph Gosez, X. Gonze, J.-P. Michenaud, Europhys. Lett. **33**, 713 (1996)
3. S. Tinte, K.M. Rabe, D. Vanderbilt, Phys. Rev. B **68**, 144105 (2003)
4. M. Dawber, K.M. Rabe, J.F. Scott, Rev. Mod. Phys. **77**, 1083 (2005)
5. K.M. Rabe, Curr. Opin. Solid State Mater. Sci. **9**, 122 (2005)
6. J. Junquera, Ph Gosez, J. Comput. Theor. Nanosci. **5**, 2071 (2008)
7. N.A. Pertsev, A.G. Zembilgotov, A.K. Tagantsev, Phys. Rev. Lett. **80**, 1988 (1998)
8. N.A. Pertsev, A.K. Tagantsev, N. Setter, Phys. Rev. B **61**, R825 (2000)
9. O. Diéguez, K.M. Rabe, D. Vanderbilt, Phys. Rev. B **72**, 144101 (2005)
10. K.J. Choi, M. Biegalski, Y.L. Li, A. Sharan, J. Schubert, R. Uecker, P. Reiche, Y.B. Chen, X.Q. Pan, V. Gopalan, L.-Q. Chen, D.G. Schlom, C.B. Eom, Science **306**, 1005 (2004)
11. A. Vasudevarao, A. Kumar, L. Tian, J.H. Haeni, Y.L. Li, C.-J. Eklund, Q.X. Jia, R. Uecker, P. Reiche, K.M. Rabe, L.Q. Chen, D.G. Schrom, V. Gopalan, Phys. Rev. Lett. **97**, 257602 (2006)
12. J.H. Haeni, P. Irvin, W. Chang, R. Uecker, P. Reiche, Y.L. Li, S. Choudhury, W. Tian, M.E. Hawley, B. Craigo, A.K. Tagantsev, X.Q. Pan, S.K. Streiffer, L.Q. Chen, S.W. Kirchoefer, J. Levy, D.G. Schlom, Nature **430**, 758 (2004)
13. A.F. Devonshire, Phil. Mag. **40**, 1040 (1949)
14. A.F. Devonshire, Phil. Mag. **42**, 1065 (1951)
15. Q.Y. Qiu, R. Mahjoub, S.P. Alpay, V. Nagarajan, Acta Mater. **58**, 823 (2010)
16. H. Sharma, J. Kreisel, Ph Ghosez, Phys. Rev. B **90**, 214102 (2014)
17. J. Padilla, W. Zhong, D. Vanderbilt, Phys. Rev. B **53**, R5969 (1996)

18. S. Pöykkö, D.J. Chadi, Appl. Phys. Lett. **75**, 2830 (1999)
19. B. Meyer, D. Vanderbilt, Phys. Rev. B **65**, 104111 (2002)
20. T. Shimada, Y. Umeno, T. Kitamura, Phys. Rev. B **77**, 094105 (2008)
21. A. Kubo, Y. Umeno, J. Solid Mech. Mater. Eng. **6**, 90 (2012)
22. A. Kubo, J.-M. Albina, Y. Umeno, Modelling Simul. Mater. Sci. Eng. **21**, 065019 (2013)
23. S. Pöykkö, D.J. Chadi, J. Phys. Chem. Solids **61**, 291 (2000)
24. X. Du, I.-W. Chen, J. Appl. Phys. **83**, 7789 (1998)
25. Z. Zhang, P. Wu, L. Lu, C. Shu, J. Alloys Compd. **449**, 362 (2008)
26. B. Meyer, D. Vanderbilt, Phys. Rev. B **63**, 205426 (2001)
27. Y. Umeno, T. Shimada, T. Kitamura, C. Elsässer, Phys. Rev. B **74**, 174111 (2006)
28. D.D. Fong, G.B. Stephenson, S.K. Streiffer, J.A. Eastman, O. Auciello, P.H. Fuoss, C. Thompson, Science **304**, 1650 (2004)
29. T. Shimada, S. Tomoda, T. Kitamura, Phys. Rev. B **81**, 144116 (2010)
30. C. Kittel, Phys. Rev. **70**, 965 (1946)
31. T. Shimada, S. Tomoda, T. Kitamura, J. Phys. Condens. Matter **22**, 355901 (2010)
32. X. Wang, S. Tomoda, T. Shimada, T. Kitamura, J. Phys. Condens. Matter **24**, 045903 (2012)
33. Y. Yamashita, K. Mukai, J. Yoshinobu, M. Lippmaa, T. Kinoshita, M. Kawasaki, Surf. Sci. **514**, 54 (2002)
34. E.A. Kotomin, E. Heifets, S. Dorfman, D. Fuks, A. Gordon, J. Maier, Surf. Sci. **566–568**, 231 (2004)
35. M.-W. Chu, I. Szafraniak, R. Scholz, C. Harnagea, D. Hesse, M. Alexe, U. Gösele, Nat. Mater. **3**, 87 (2004)
36. J.H. Jeon, S.K. Choi, Appl. Phys. Lett. **91**, 091916 (2007)
37. C. Thompson, D.D. Fong, R.V. Wang, F. Jiang, S.K. Streiffer, K. Latifi, J.A. Eastman, P.H. Fuoss, G.B. Stephenson, Appl. Phys. Lett. **93**, 182901 (2008)
38. J. Junquera, P. Ghosez, Nature **422**, 506 (2003)
39. Th Tybell, C.H. Ahn, J.-M. Triscone, Appl. Phys. Lett. **75**, 856 (1999)
40. A.V. Bune, V.M. Fridkin, S. Ducharme, L.M. Blinov, S.P. Palto, A.V. Sorokin, S.G. Yudin, A. Zlatkin, Nature **391**, 874 (1998)
41. Y. Umeno, B. Meyer, C. Elsässer, P. Gumbsch, Phys. Rev. B **74**, 060101(R) (2006)
42. N. Sai, A.M. Kolpak, A.M. Rappe, Phys. Rev. B **72**, 020101(R) (2005)
43. Y. Umeno, J.-M. Albina, B. Meyer, C. Elsässer, Phys. Rev. B **80**, 205122 (2009)
44. P.M. Rørvik, T. Grande, M.-A. Einarsrud, Adv. Mater. **23**, 4007 (2011)
45. M.-Q. Cai, D.U. Yong, B.-Y. Huang, Trans. Nonferrous Met. Soc. China **19**, 1634 (2009)
46. M.Q. Cai, Y. Zheng, B. Wang, G.W. Yang, Appl. Phys. Lett. **95**, 232901 (2009)
47. G. Pilania, S.P. Alpay, R. Ramprasad, Phys. Rev. B **80**, 014113 (2009)
48. G. Pilania, R. Ramprasad, Phys. Rev. B **82**, 155442 (2010)
49. G. Pilania, R. Ramprasad, J. Mater. Sci. **47**, 7580 (2012)
50. X. Zhu, Z. Liu, N. Ming, J. Mater. Chem. **20**, 4015 (2010)
51. Y. Yang, X. Wang, C. Sun, L. Li, J. Appl. Phys. **104**, 124108 (2008)
52. Y. Yang, X. Wang, C. Sun, L. Li, J. Appl. Phys. **109**, 014109 (2011)
53. T. Shimada, X. Wang, Y. Kondo, T. Kitamura, Phys. Rev. Lett. **108**, 067601 (2012)
54. J. Wang, T. Xu, T. Shimada, X. Wang, T.-Y. Zhang, T. Kitamura, Phys. Rev. B **89**, 144102 (2014)
55. T. Shimada, S. Tomoda, T. Kitamura, Phys. Rev. B **79**, 024102 (2009)
56. Y. Zhang, J. Hong, B. Liu, D. Fang, Nanotechnology **21**, 015701 (2010)
57. Y. Zhang, B. Liu, D. Fang, J. Appl. Phys. **110**, 054109 (2011)
58. S. Tinte, M.G. Stachiotti, M. Sepliarsky, R.L. Migone, C.O. Rodriguez, J. Phys. Condens. Matter **11**, 9679 (1999)
59. S. Prosandeev, L. Bellaiche, Phys. Rev. B **75**, 094102 (2007)
60. B. Li, J.B. Wang, X.L. Zhong, F. Wang, L.J. Wang, Y.C. Zhou, J. Appl. Phys. **114**, 044301 (2013)

61. J.Y. Liu, W.J. Chen, B. Wang, Y. Zheng, J. Appl. Phys. **114**, 044101 (2013)
62. W.J. Chen, Y. Zheng, B. Wang, D.C. Ma, F.R. Ling, Phys. Chem. Chem. Phys. **15**, 7277 (2013)
63. C.M. Wu, W.J. Chen, Y. Zheng, D.C. Ma, B. Wang, J.Y. Liu, C.H. Woo, Sci. Rep. **4**, 3946 (2014)

Chapter 6
Magnetism in Nanostructures

Abstract As magnetism is susceptible to structures and strains as well as crystal defects, extensive investigations by means of first-principles calculations have been devoted to this matter. We review a series of studies on magnetism in bulk and nanostructures. After giving an overview of studies of internal defects as understructures, we discuss magnetism in low-dimensional nanostructures, including thin films, monolayers, nanowires, nanotubes, atomic chains, and atomic clusters.

Keywords Magnetism · Crystal defect · Non-collinear

6.1 Magnetism in Bulk

6.1.1 Magnetism and Its Response to Strain

Magnetism is a material property that the material exhibits spontaneous or permanent magnetic moment in the absence of external magnetic field. The magnetic moment aligns in a certain order in a crystal (e.g., ferromagnetism and antiferromagnetism), and the direction can be manipulated by applying external magnetic field. In general, transition metals (e.g., Fe, Ni, Co) and their oxides exhibit magnetism and have been studied as a prototype of magnetic materials [1].

Magnetism is, in general, sensitive to the crystal lattice structure and applied mechanical strains. The interplay between magnetism and mechanical strain (lattice deformation) was first investigated for ferromagnetic iron bulk materials [2]. The most typical and famous research is probably the magnetic phase transitions along the body-centered-cubic (bcc) to face-centered-cubic (fcc) transformation path (namely a Bain's path) [3]. This deformation path can be seen, for example, in an iron thin film on a (001) fcc-metal substrate, due to the pseudo-epitaxy on a fcc(001) surface. Figure 6.1a shows the total energy of the iron bulk with the ferromagnetic (FM), nonmagnetic (NM), single-layered antiferromagnetic (AFM1), and double-layered antiferromagnetic (AFMD) phases as a function of lattice deformation c/a, where a and c denote the in-plane and out-of-plane lattice parameters of epitaxially

Y. Umeno et al., *Multiphysics in Nanostructures*, Nanostructure Science and Technology, DOI 10.1007/978-4-431-56573-4_6

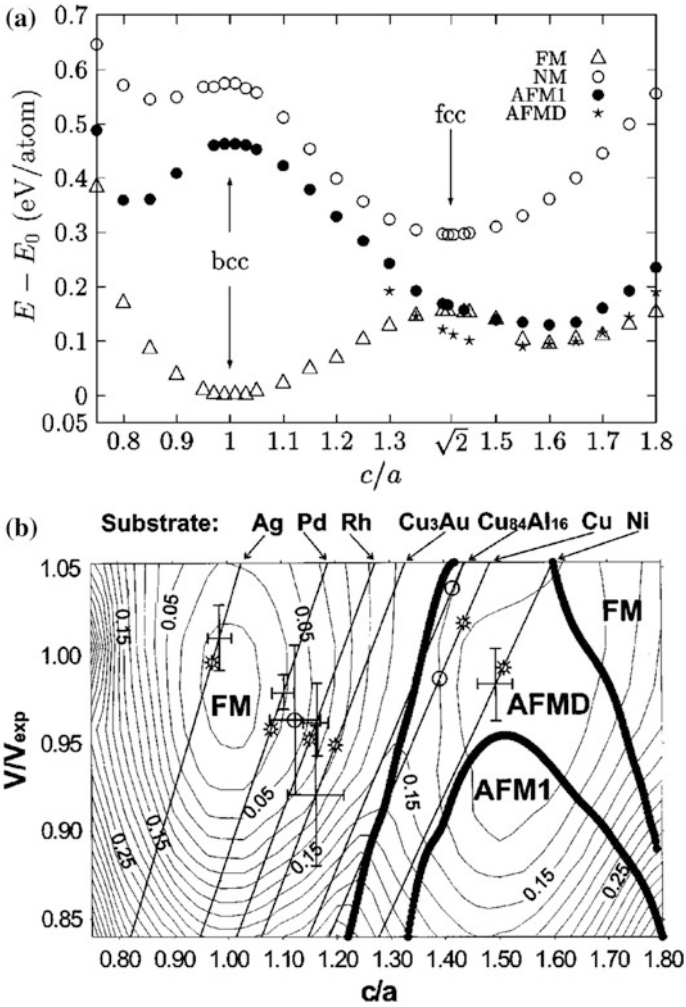

Fig. 6.1 **a** Total energies of iron films along the constant-volume bcc–fcc transformation path for the experimental unit-cell volume calculated within the generalized gradient approximation (GGA) relative to the equilibrium energy E_0 of FM bcc iron. **b** Total energy of iron as a function of tetragonality of lattice c/a and volume relative to the FM bcc ground-state energy calculated within the generalized gradient approximation. Only the states with the minimum energy are shown in the figure. The interval of isolines is 20 meV/atom. The *thick lines* represent the phase boundaries of FM/AFMD and AFMD/AFM1. The straight lines correspond to constant lateral lattice parameters of different (001) substrates. Reprinted with permission from Ref. [2]. Copyright 2001 by American Physical Society

strained iron [2]. Here, the spin configuration of AFM1 is …↑ ↓ ↑ ↓…, while the spin configuration of …↑ ↑ ↓ ↓… is for the AFMD. These calculations were performed using full-potential linearized augmented plane waves (FLAPW) method within the generalized gradient approximation (GGA), which is essential for a correct

description of structural and magnetic ground state of transition metals [4, 5]. The NM and FM states exhibit energy extrema at $c/a = 1$ and $c/a = \sqrt{2}$ corresponding to the higher-symmetry lattices (i.e., bcc and fcc, respectively). On the other hand, the AFM1 iron keeps its cubic symmetry only for $c/a = 1$ (i.e., bcc), while at $c/a = \sqrt{2}$, the atoms occupy the fcc lattice positions, but as the atoms with spins up and down are not equivalent by the magnetic symmetry breaking, resulting in tetragonal symmetry; i.e., the structural and magnetic degrees of freedom are coupled with each other. The same behavior is also seen in the AFMD iron. Obviously, the most stable structure is FM bcc at $c/a = 1$, while the lattice deformation leads to FM-to-AFMD phase transition around $c/a = \sqrt{2}$. Further deformation stabilizes the AFM1 state with respect to the AFMD state. As a result, iron undergoes a FM–AFMD–AMD1 phase transition along the Bain's deformation path.

Figure 6.1b plots the magnetic phase diagram along the Bain's path with various unit-cell volumes. It can be easily understood that the magnetic phases of iron are mainly tailored among the FM, AFM1, and AFMD states by applying in-plane lattice strains through the choice of substrates. For example, the choice of substrates with a larger lattice parameter, e.g., Ag, Pd, Rh, and Cu_3Au, sustains ferromagnetism as the unstrained bcc iron does. On the other hand, the choice of $Cu_{84}Al_{16}$, Cu, and Ni substrates with a smaller lattice parameter results in the FM-to-AFM1 or FM-to-AMFD phase transitions via the coincident bcc–fcc structural transition. In fact, the experiments showed that these substrates change the magnetic phase of iron films [6–8]. It is also noted that the calculations based on the local density approximation (LDA) are less effective for prediction than the GGA, which is an opposite trend to ferroelectricity, which can be successfully described by the LDA instead of the GGA, as shown in the previous section. Therefore, the following results related to magnetism reviewed here are mostly within the GGA.

Such strain-induced magnetic phase transitions are intensively studied for various deformation paths, e.g., bcc–hcp (hexagonal-closed packed), uniaxial, biaxial, and trigonal deformations) in various transition metals (e.g., iron, nickel, and cobalt) [9–11]. We should remark that these are all for perfect crystals and do not include any influence from the defects (understructures) and nanostructures (outer shapes), which modify the magnetic properties locally or even globally. These effects are systematically reviewed in the following sections.

6.1.1.1 Effects of Defects

Grain boundaries

Grain boundaries are the most typical defects in materials and regarded as a two-dimensional (2D) understructure in crystals. Since the lattice misfit and misorientation at the grain boundaries induce local lattice strain and rapid change in coordination number, the grain boundaries are considered to alter the magnetic properties in the materials. Here, we take a $\Sigma 5$ tilt grain boundary in ferromagnetic iron as a representative example of planar defects in the ferromagnetic materials.

Figure 6.2a shows the atomic configuration of a Σ5 tilt boundary in ferromagnetic bcc iron [12]. The stable atomic configuration of the grain boundary structure can be determined by calculating and mapping the total energy with shifting one grain with respect to another grain (i.e., γ-surface calculations). The local (muffin-tin) magnetic moment is evaluated around each of the relaxed potentials from the spin-resolved muffin-tin charge densities, and it is plotted in Fig. 6.2b together with the magnetic moment of the bulk counterpart [13]. The remarkable enhancement of the local magnetic moment is found near the grain boundary. Such enhancement is, however, limited to the two layers adjacent to the grain boundary interface although a small fluctuation in the magnetic moment exists further away from the defect which rapidly converges to the bulk value. Such behavior is common among grain boundaries [14]. The enhancement of the magnetic moment might be due to the increased volume around the grain boundary layer (i.e., free-volume effect), which is caused by the lower coordination of nearest and second-nearest neighbors at the boundary. These local enhancement and distribution of the relaxation away from the grain boundary are qualitatively similar to those calculated for iron (100) surfaces, as simultaneously shown in Fig. 6.2b. The effect of surfaces will be discussed in the later section.

Dislocations

Following the grain boundaries as the 2D understructure (or defect), here we review the dislocations in ferromagnetic iron as the one-dimensional (1D) defects.

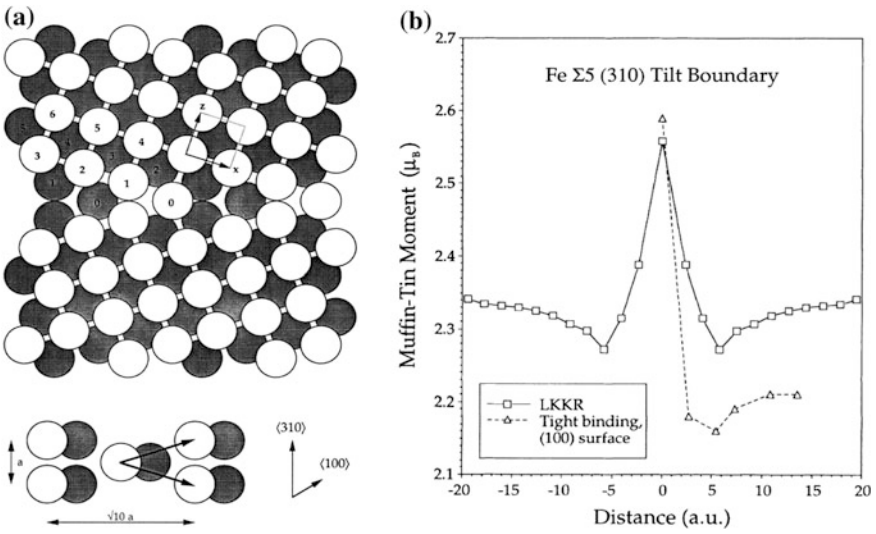

Fig. 6.2 **a** The atomic configuration of Fe Σ5(310) tilt grain boundary for density-functional theory calculations. **b** A layer-by-layer magnetic moment near the Fe Σ5 tilt grain boundary. Reprinted with permission from Ref. [12]. Copyright 1993 by American Physical Society

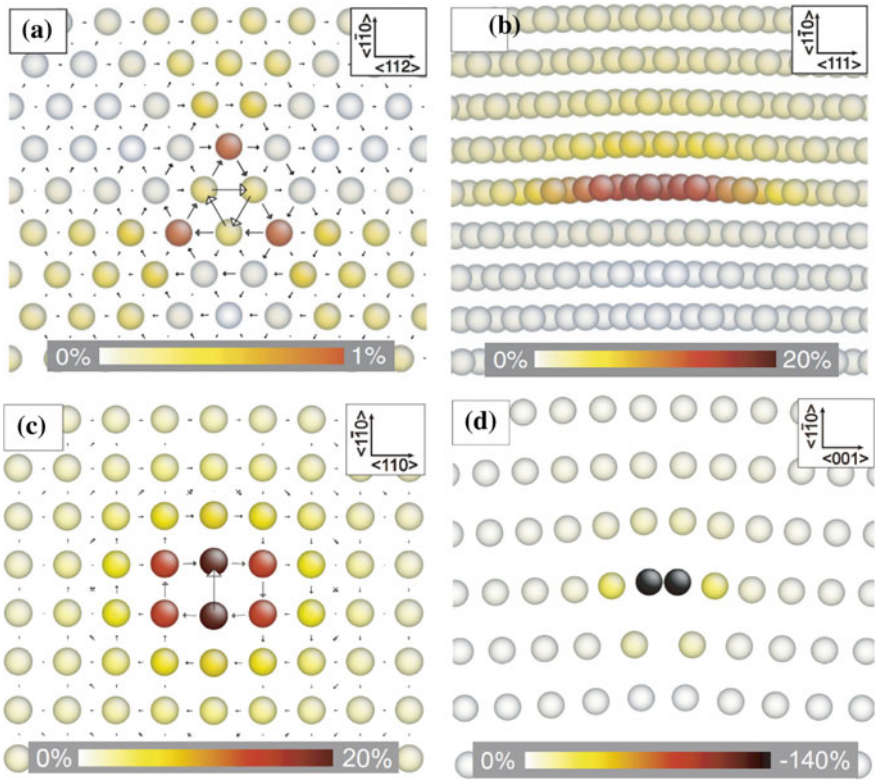

Fig. 6.3 Atomic configuration of the 1/2 ⟨111⟩ screw **a** and edge **b** dislocation cores. Atomic configuration of the ⟨100⟩ screw **c** and edge **d** dislocation cores. The contour color represents the suppression of local magnetic moments from the bulk value. Reprinted with permission from Ref. [15]. Copyright 2011 by American Physical Society

Figure 6.3a–d shows the structural and magnetic properties of both screw and edge dislocation cores with the 1/2 ⟨111⟩ and ⟨100⟩ Burgers vectors, respectively, in ferromagnetic iron bulk [15]. For the 1/2 ⟨111⟩ screw dislocation, both the changes of atomic configuration and magnetic moments are less remarkable, and the dislocation core is the nondegenerate structure [16]. The small changes in local magnetic moments around the core arise from a slight change in interatomic bond lengths. In contrast, a significant decrease in the magnetic moments (about 20% of the bulk value) can be seen in the 1/2 ⟨111⟩ edge dislocation core, which subject to quite high compressive stress due to the lattice misfit. Due to such higher strains around the cores, even larger changes occur at the ⟨100⟩ dislocation cores (see Fig. 6.3c–d). Thus, such local strains at the dislocation cores suppress or enhance the local magnetic moments slightly or dramatically, depending on the core configuration.

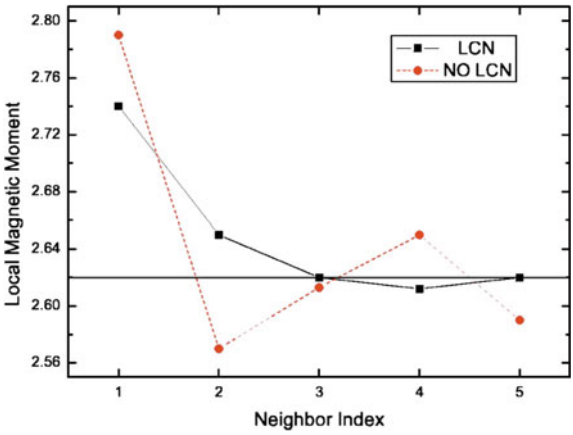

Fig. 6.4 Local magnetic moments on the first–fifth neighboring iron atoms from a vacancy site in ferromagnetic iron with and without locally charge neutral (LCN) conditions. Reprinted with permission from Ref. [17]. Copyright 2005 by American Physical Society

Vacancies

Atomic vacancies are the typical point (0D) defects (or understructure) in magnetic materials. Figure 6.4 shows the distributions of the local magnetic moments around a vacancy in ferromagnetic iron bulk [17]. Since the iron is the metallic system with ideally perfect screening, each atom is assumed to remain locally charge neutral (LCN). To include this charge self-consistency around defects, which is important for a realistic treatment of point defects, two kinds of simulations were performed; with and without the LCN conditions. The local magnetic moment is strongly enhanced from the bulk value at the nearest neighbors to the vacancy, while the magnetic moment tends to decrease as the distance from the vacancy increase, and it becomes close to the bulk value at the third neighbors with the LCN condition. On the other hand, without the LCN condition, the local magnetic moment shows the oscillation behavior around the vacancy with respect to that with the LCN condition, leading to a long-range influence on magnetic properties. This indicates that the extent of screening in transition metallic systems plays an important role in determining the magnetic properties of point defects in ferromagnetic materials.

6.2 Thin Film and Monolayer: Two-Dimensional Structure

6.2.1 Thin Films and Surface Properties

The ferromagnetic properties of thin films are generally dominated by the surface effects due to their significantly large surface-to-volume ratio. Here, we focus on the ferromagnetic iron thin films as a representative of 2D nanostructures. We review the ferromagnetic properties of a stable (001) surface of the iron thin films [18]. Figure 6.5a shows the layer-resolved local magnetic moment in the ferromagnetic

iron film obtained from the density-functional theory calculations with the Perdew–Burke–Ernzerhof (PBE) functional within the GGA, which can reproduce the both structural and magnetic properties of iron. Within the plane wave basis set and the projector-augmented wave formalism [19, 20], the local magnetic moments can be evaluated by projecting the wave functions onto the spherical harmonics around each atom within the Wigner-Seitz radii. The magnetic moment increases near the surface layer compared with the bulk value. In particular, the surface layer (see layer number 1) exhibits the highest magnetic moment of 2.94 μ_B. It should be noted that, comparing with the unrelaxed (fix all the atoms at the bcc lattice site) and relaxed (stable) surface structures, the surface relaxation leads to little influence on the increase in magnetic moment. This implies that the low coordination number intrinsic to the surface layer plays an important role in determining the enhanced magnetic moment.

Such enhancement of ferromagnetism at the surface layer can be understood from the electronic point of view. One can consider d orbitals in iron that predominate ferromagnetism of transition metals. Based on the crystal-field theory [21, 22], the five d orbitals, d_{xy}, d_{yz}, d_{zx}, d_{x2-y2}, and d_{z2}, in the bcc crystal reduce to the hybridized t_{2g} (d_{xy}, d_{yz}, and d_{zx}) and e_g (d_{x2-y2} and d_{z2}) states. Here, the distribution of t_{2g} orbital is toward the nearest neighbors in the bcc lattice via the formation of ddσ bonds, whereas the e_g orbital is distributed toward the second-nearest neighbors. Figure 6.5b, c represents the spin-polarized local electronic density of states (DOS) for the total d- and t_{2g}-e_g-decomposed states inside the film and at the (001) surface, respectively. Here, the local DOS can be evaluated by projecting the wave functions onto the spherical harmonics around each atom, as we can do for local magnetic moment. Inside the iron thin films, the majority-spin t_{2g} and e_g states are almost fully occupied. This shows that both the t_{2g} and e_g states are almost equitably occupied, which leads to no specific directionality in the majority-spin density distribution [18]. On the other hand, the t_{2g} state is majorly occupied in the minority-spin DOS, while the e_g states are located above the Fermi level. For the (001) surface, in contrast, a large difference can be seen in the t_{2g} state: The majority-spin t_{2g} states are localized around the lower energy level of −3.0 eV, which thus results in a fully occupied majority-spin state. On the other hand, the minority-spin t_{2g} bands are almost localized around an energy level slightly higher than the Fermi level. Therefore, the t_{2g} state has much narrower bandwidth with respect to that inside the film (or bulk). Considering the fact that the bandwidth is reduced as the coordination number decreases [23], having 50% fewer nearest neighbors at the surface (i.e., low coordination number) leads to a significant suppression in the bandwidth of the t_{2g} state, while the slight change in the e_g DOS arises from the more limited loss of second-nearest neighbors. As a result of such energy-level shift of the band structure between the surface and the inside, the number of occupied majority-spin and minority-spin electrons changes. This results in the enhancement of the magnetic moment at the (001) surface.

In general, thin films are prepared by an epitaxial growth on a certain substrate, which apply in-plane (epitaxial) tension or compression to the film due to lattice mismatch at the interface between the film and the substrate, as shown in the bulk

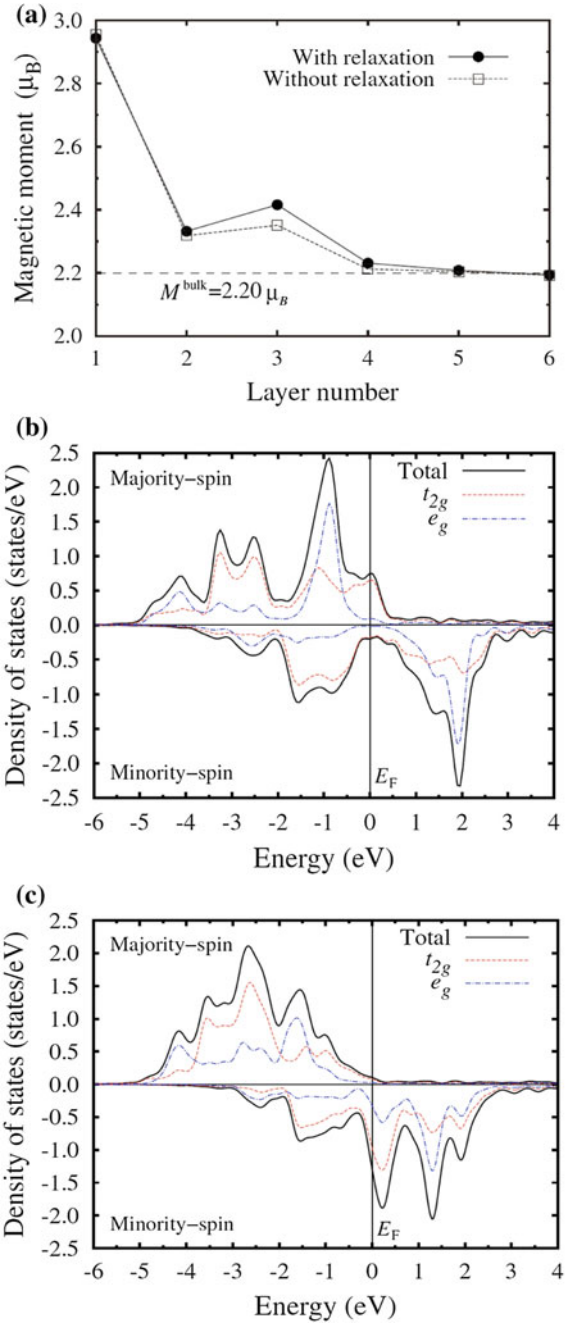

Fig. 6.5 **a** Layer-by-layer magnetic moment in the FM iron film with and without structural relaxation. The layer numbers 1–5 correspond to the first–fifth atomic layers from the Fe (001) surface. The dashed horizontal line denotes the bulk value 2.20 μ_B. Total *d* and t_{2g}-e_g decomposed electronic local density of states (DOS) **b** in the Fe bulk and **c** at the (001) surface. Reprinted with permission from Ref. [18]. Copyright 2010 by American Physical Society

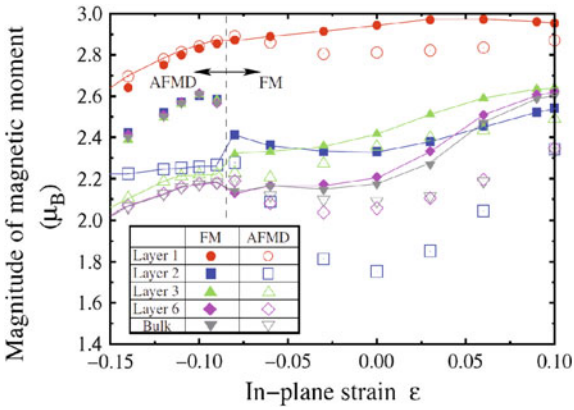

Fig. 6.6 Magnetic moment as a function of epitaxial strain in the iron thin film and bulk. Reprinted with permission from Ref. [18]. Copyright 2010 by American Physical Society

section. The FM phase is energetically preferable under tension, while the FM-to-AFMD phase transition occurs at a compressive strain of −0.09 [18]. Figure 6.6 plots the magnitude of the magnetic moment as a function of epitaxial strain, in the iron nanofilm [18]. The tensile epitaxial strain smoothly enhances the magnetic moments for all layers, while the magnetic moments are suppressed by compression in both the film and bulk. Under compression, on the other hand, a discontinuous change in magnetic moment can be seen at the critical strain of $\varepsilon = -0.09$ due to the magnetic and structural phase transition from FM to AFMD. At the surface layer, in contrast, the magnetic moment has no discontinuities and is thus insensitive to the phase transition, although the crystal lattice undergoes an abrupt change from bcc to fcc. The thin film, especially the surface layer, always exhibits a high magnetic moment despite any applied strains.

6.2.2 *Monolayer*

Due to the strong surface effect shown above, the ferromagnetic properties of thin films strongly depend on the thickness. The ultimately thinner geometry is obviously a monolayer that consists of only one atomic layer. Recent advances have enabled to fabricate the Fe monolayer that consists of a single (110) atomic layer of a bcc lattice by pseudomorphic growth on substrates [24]. Here, we present the magnetic properties and unusual phase transitions in such an ultimately thin geometry.

Before going to the details, a brief description for the calculations of incommensurate SS (spin-spiral) noncollinear magnetism should be provided here. Noncollinear helimagnetism with incommensurate SS waves is described by the fully unconstrained formalism [25] in which density-functional theory is expressed in terms of a 2×2 density matrix, in conjunction with the generalized Bloch

conditions. According to generalized Bloch theorem, the wave function is described as a two component Bloch spinors instead of the usual one-electron wave function [26, 27]. The Bloch spinors are still characterized by a wave vector **k** vector in the first Brillouin zone as:

$$\Psi_k(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}} \begin{pmatrix} e^{-i\mathbf{q}\cdot\mathbf{r}/2}\alpha_{\mathbf{k}}(\mathbf{r}) \\ e^{-i\mathbf{q}\cdot\mathbf{r}/2}\beta_{\mathbf{k}}(\mathbf{r}) \end{pmatrix} \tag{6.1}$$

where **r** is position vector. Here, $\alpha_{\mathbf{k}}(\mathbf{r})$ and $\beta_{\mathbf{k}}(\mathbf{r})$ correspond to spin-up and spin-down components and are "invariant" with respect to lattice translations. Therefore, the incommensurate spin-spiral configurations can be analyzed by applying this generalized Bloch theorem (condition).

Figure 6.7a represents the SS energy, $\Delta E = E_{\mathrm{SS,min}} - E_{\mathrm{FM}}$, with respect to applied strains [28]. Here, the minimum energy at the SS phase within q in the Brillouin zone is denoted by $E_{\mathrm{SS,min}}$. The negative ΔE means a stable noncollinear SS state. Under the strain-free and tensile condition, ΔE is always zero and iron monolayer keeps FM. Under compression, on the other hand, the noncollinear SS phase appears via the FM-to-SS phase transition (see the white line). Since this noncollinear SS state cannot be seen in the iron bulk, the SS phase is characteristic to the monolayer system, i.e., the ultimate single-layer geometry, and the effect of strains stabilizes such unusual SS magnetic phase. Note that the Fe monolayer grown on a W(110) substrate, which corresponds to the strain of $(\varepsilon_{xx}, \varepsilon_{yy}) = (0.10, 0.10)$, exhibits the FM ground state. This is consistent with the experimental observation in the Fe monolayer and suggests that the quite large strain can be actually applied to the monolayer [24, 29]. Figure 6.7b plots the SS wave vector q as functions of strains [28]. The minimum SS excitation energy is located along the Γ-N path (see Fig. 6.7d) in the Brillouin zone $[q = 2\pi/\mathrm{L}(q_x, 0)]$, where L is the lattice parameter of the monolayer. So, here, we show the q_x value in the plot, which smoothly increases with compressive strain. We can find anisotropy of strain effects: A larger q_x (=0.4–0.5) is found around $\varepsilon_{xx} = -0.20$, whereas q_x is smaller (0.1–0.25) under $\varepsilon_{yy} = -0.20$. This suggests that ε_{xx} has large influence on the SS wave vector q_x. Figure 6.7c shows the magnitude of the magnetic moment m as a function of the in-plane strains. The magnetic moment in the Fe monolayer seems to vary smoothly even across the FM–SS transition line, whereas a discontinuous change in the magnetic moment is observed in bulk iron that undergoes a FM-to-AFMD phase transition (see the previous section). This characteristic corresponds to the fact that the magnitude of m in the monolayer is less sensitive to the SS wave vector q, as seen in the surface layer of the thin film [18]. This suggests that ultimate low coordination number in monolayer (or surface layer) leads to high magnetic moment and unusual magnetic phases.

To see the energetics details of this unusual SS phase and its transition in the monolayer, here we provide the interatomic exchange interaction based on the effective Heisenberg Hamiltonian:

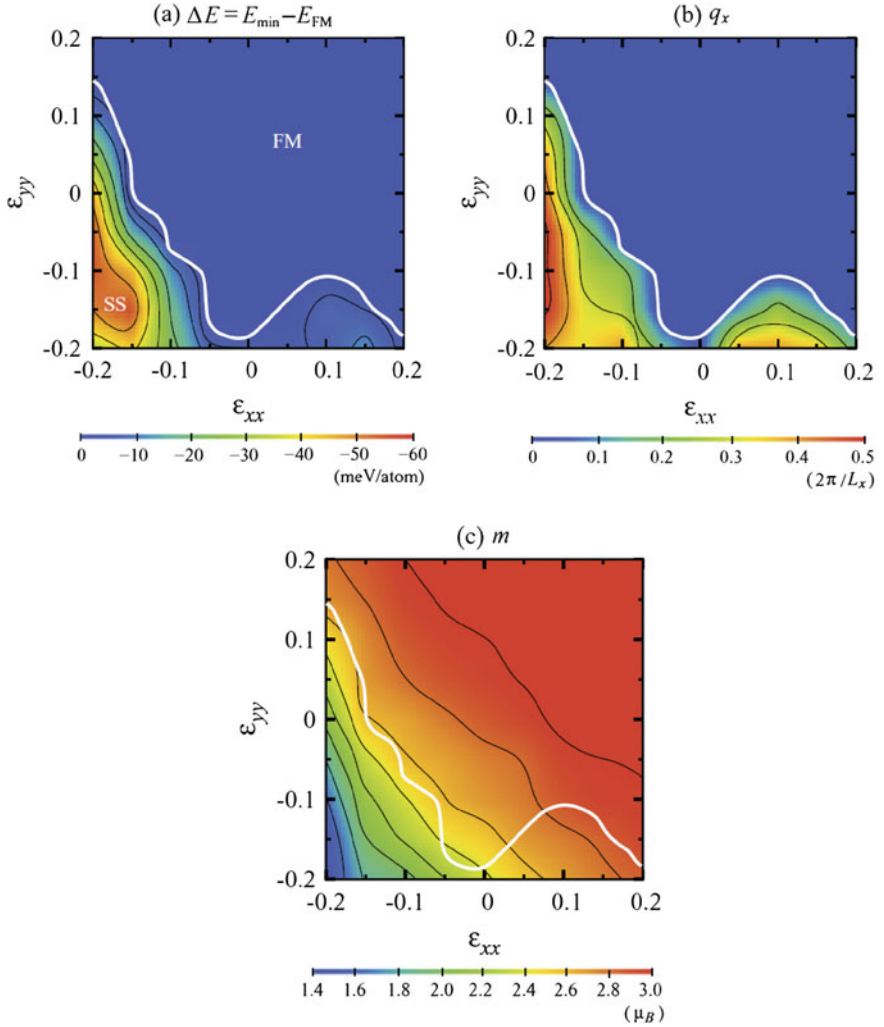

Fig. 6.7 **a** SS energy of $\Delta E = E_{SS,min} - E_{FM}$, **b** wave vector q_x, and **c** magnitude of magnetic moment m of iron monolayer as functions of strains (ε_{xx}, ε_{yy}). The *white lines* denote the magnetic phase transition of FM-SS. Reprinted with permission from Ref. [28]. Copyright 2012 by American Physical Society

$$H_{\text{eff}} = -\frac{1}{2}\sum_{i,j} J_{ij}\mathbf{e}_i \cdot \mathbf{e}_j, \tag{6.2}$$

where J_{ij} is the interatomic exchange interaction between the ith and jth neighbors and $\mathbf{e}_i$ is the unit vector representing the direction of the local magnetic moment on

the *i*th neighbor. From Eq. (6.2), the total energy of the noncollinear SS state with $\mathbf{q}$, $E_{SS}(\mathbf{q})$, is given by

$$E_{SS}(\mathbf{q}) = -\frac{1}{2} J_{0j} \exp(-i\mathbf{q} \cdot \mathbf{R}_{0j}), \tag{6.3}$$

where $\mathbf{R}_{0j}$ denotes the coordinate of the *j*th neighbor. In the classical frozen magnon approach [30], J_{0j} can be related to the SS wave excitation energy (magnon dispersion relation), $\Delta E(\mathbf{q}) = E_{SS}(\mathbf{q}) - E_{FM}$, by Fourier backtransformation:

$$J_{0j} = \frac{1}{N_{\mathbf{q}}} \sum_{\mathbf{q}} \Delta E(\mathbf{q}) \exp(-i\mathbf{q} \cdot \mathbf{R}_{0j}), \tag{6.4}$$

where $N\mathbf{q}$ denotes the number of q points in the Brillouin zone. The exchange parameter of the monolayer can be evaluated by summing the SS wave excitation energies for various $\mathbf{q}$ vectors using Eq. (6.4).

Figure 6.8 represents the interatomic exchange interaction of the *j*th neighbor, J_{0j}, in the Fe(110) monolayer as a function of strain (ε_{xx}, ε_{yy}) [28]. Under strain-free conditions, $\varepsilon_{xx} = \varepsilon_{yy} = 0.0$, J_{01} has a large positive value, which means that the positive exchange interaction tends to align the magnetic moment in the same direction (i.e., ferromagnetic coupling). In contrast, J_{02} shows a small negative value, meaning that the interaction from the second neighbor aligns the magnetic moment in the opposite direction (i.e., antiferromagnetic coupling). J_{03} is positive (FM coupling), but its magnitude is one order of magnitude smaller than those of J_{01} and J_{02}, indicating that the exchange interaction from the third neighbors has a smaller influence than those from the first and second neighbors. Therefore, the FM ground state can be stabilized by the competing exchange interaction between the larger ferromagnetically coupled nearest neighbors and the smaller antiferromagnetically coupled second neighbors. This trend is more remarkable under tensile strains. Under compression, on the other hand, J_{01} and J_{02} decrease rapidly and the compression therefore weakens the FM interaction and strengthens the AFM interaction. This change destroys the original balance of J_{01} and J_{02} and leads to the noncollinear SS ground state.

It is also possible to predict the critical condition for the FM-to-SS phase transition from the exchange interaction parameters: As the transition occurs when $E_{SS}(\mathbf{q})$ in Eq. (6.3) becomes a minimum at $q \neq 0$, the critical condition for the transition can be given by $J_{02}/J_{01} = -1/4$ [31]. Figure 6.8c shows the FM-to-SS magnetic phase transition predicted by the critical condition based on J_{01} and J_{02} (see the red line). The $J_{01} - J_{02}$-dominated transition line coincides well with the actual FM–SS transition around the isotropic strain ($\varepsilon_{xx} = \varepsilon_{yy} = -0.1$). On the other hand, in the transition in the anisotropic strain region ($\varepsilon_{xx} \neq \varepsilon_{yy}$), they deviate from each other. The deviation can be corrected by including three exchange parameters, J_{01}, J_{02}, and J_{03}, represented by the blue line, which is in excellent agreement with the actual FM–SS transition for both isotropic and anisotropic strains. This indicates

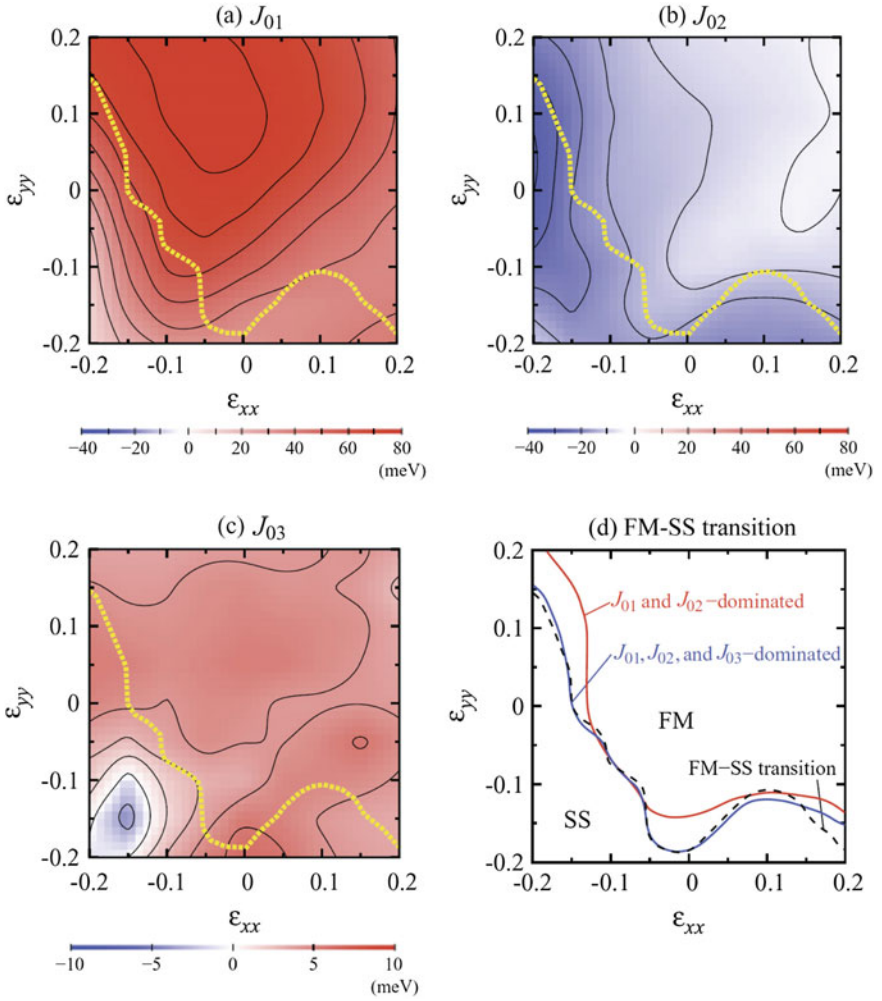

Fig. 6.8 Interatomic exchange interactions **a** J_{01}, **b** J_{02}, and **c** J_{03} as functions of strains (ε_{xx}, ε_{yy}). The FM–SS transition is given by the *yellow line*. **d** FM–SS transition line predicted using the exchange parameters J_{01} and J_{02} (*red line*) and J_{01}, J_{02}, and J_{03} (*blue line*). The *black dashed line* shows the actual FM–SS transition directly calculated from ab initio calculations. Reprinted with permission from Ref. [28]. Copyright 2012 by American Physical Society

that the FM-to-SS phase transition under the isotropic strain is dominated by the exchange interaction of the first and second neighbors, while the long-range interaction from the third neighbors becomes critical for the transition under anisotropic strain. As a whole, even the ultimate monolayer system can be understood and modeled by considering the exchange interactions, which will help to predict the unconventional phase transition in low-dimensional system.

6.3 Nanowire, Nanotube, and Atomic Chain: One-Dimensional Structure

6.3.1 Nanowires

Nanowires are one of the typical samples of one-dimensional ferromagnetic nanostructures. The nanowires are surrounded by surfaces, and in particular, the edge of nanowires is expected to lead to a much less coordination number and resulting strong effect of low dimensionality than the thin-film geometry. Here, we review the magnetic properties of nanowires by focusing on the influence of edges. Figure 6.9 shows the distribution of local magnetic moments and bond length in an iron nanowire with a 3 × 3 unit-cell cross section [32]. The edge of the nanowire exhibits the highest magnetic moment of 2.83 μ_B (about 30% larger with respect to the bulk value), i.e., strong enhancement of magnetic moment. In addition, the atom located at the edge also has the shortest bond length of 2.402 Å (e.g., others are from 2.43 to 2.46 Å). Similarly, the magnetic moment at the surface area exhibits relatively large values (2.53–2.57 μ_B), while the magnetic moment inside the nanowire (2.18 μ_B) is almost the same as the bulk value. The effects of the edge

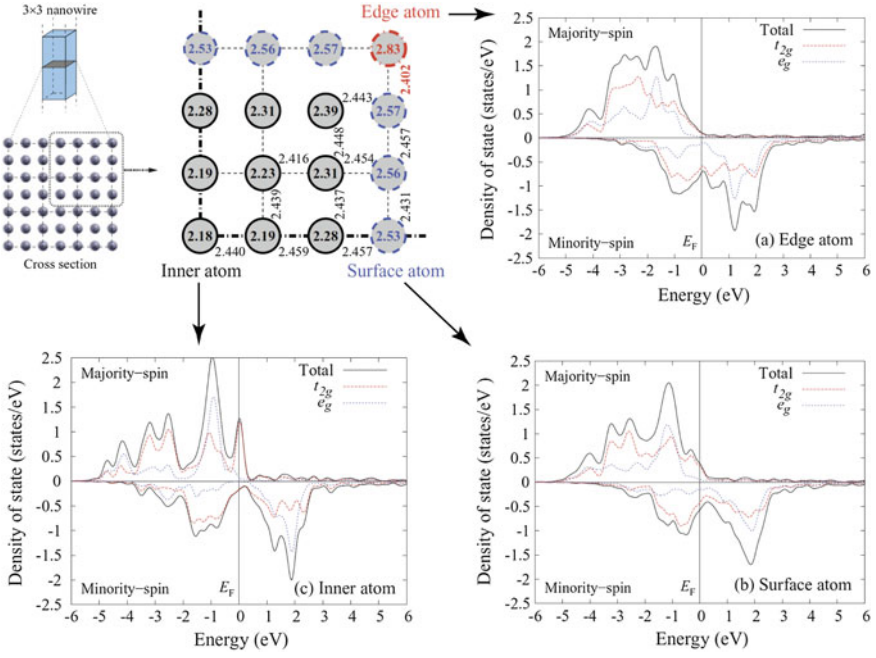

Fig. 6.9 Magnetic moment distribution (in μ_B) and bond length (in Å) in 3 × 3 iron nanowire. The top right quarter of the nanowire is presented. Total d and t_{2g}-e_g decomposed electronic local densities of states of the edge, surface, and inner atoms in the 3 × 3 nanowire. Reprinted with permission from Ref. [32]. Copyright 2011 by American Physical Society

thus range approximately 4 Å, suggesting that the edge predominates the magnetism in such ultrathin nanowires.

Electronic structure can provide a detail of such enhancement at the edge. Figure 6.9 shows the spin-polarized local electronic density of states (DOS) of the total d and t_{2g}-e_g-decomposed states of the edge, surface, and inner atoms in the 3 × 3 nanowire [32]. For the majority spins, both the t_{2g} and e_g states are highly occupied; i.e., both the t_{2g} and e_g states are almost equally occupied and results in the isotropic distribution of the majority-spin density. On the other hand, for minority spin, t_{2g} is almost occupied, while e_g is little occupied. This corresponds to the directional minority-spin density distribution toward the nearest neighbors [33]. For the minority spin, the t_{2g} states of the inner atom distributed across E_F and exhibits a broader bandwidth, while the t_{2g} state of the edge atom is highly localized around E_F, resulting in a relatively narrow bandwidth. Because the edge atom has fewer coordination numbers, the bandwidth of the corresponding t_{2g} states decreases. As a result, a remarkable increase (reduction) in the number of electrons occurs in the majority-spin (minority-spin) t_{2g} state; i.e., the electrons of the edge atom change from minority-spin t_{2g} to majority-spin t_{2g} states. Therefore, this charge transfer enhances the magnetic moment at the edge.

The magnetic ground state of bulk iron is FM, but changes to AFM when the tensile strain is applied around 0.2 due to the bcc–fcc transition along the Bain's path. On the other hand, the nanowire keeps FM under high tensile strain without such FM-to-AFM magnetic phase transition. To see structure–magnetic relationship, Fig. 6.10a, b shows the tetragonality of the lattice *c/a* and magnetic moments, respectively, around the edge, surface, and inner atoms, and in the FM 3 × 3 nanowire as a function of axial strain ε_z [32]. The lattice tetragonality is almost *c/a* = 1 (i.e., bcc) when no strain is applied, while an abrupt increase in the lattice tetragonality can be seen at $\varepsilon_{zz} = 0.30$ in which *c/a* is almost $\sqrt{2}$ (i.e., fcc). At this straining point, the magnetic moments decrease dramatically. This suggests that the sharp reduction in the magnetic moment is due to the structural transition. On the other hand, even at the critical strain, the magnetic moment at the edge remains constant. Therefore, the edge of nanowires always keeps the FM phase and the high magnetic moment even under a high strain.

6.3.2 Nanotubes

Another important and exotic 1D nanostructure of transition metal is helical nanotubes. Since the surprising fabrication of a single-walled helical gold nanotube with a smaller diameter of 0.4 nm [34], such single-walled helical nanotubes have been studied both experimentally and theoretically for various metals as well as Au [35–38]. Such ultrathin nanotubes exhibit unique properties arising from the helical shell geometry, including quantum ballistic conductance, superplasticity, and helical conduction channels [38–40]. Due to the novelty, here we review the magnetic properties of single-walled helical nanotubes.

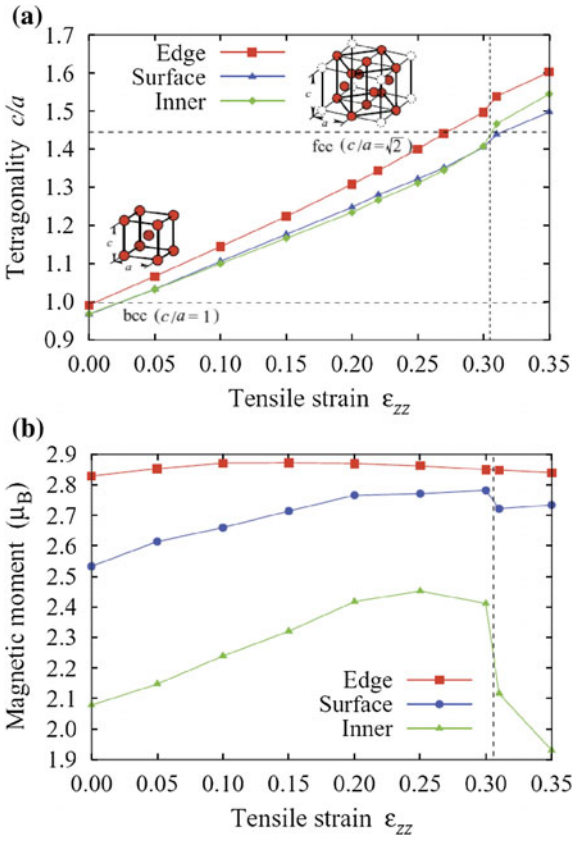

Fig. 6.10 **a** Lattice tetragonality c/a around the edge, surface, and inner atoms in the 3 × 3 nanowire as a function of axial strain ε_{zz}. **b** Magnetic moment of the edge, surface, and inner atoms in the 3 × 3 nanowire as a function of axial strain ε_{zz}. Reprinted with permission from Ref. [32]. Copyright 2011 by American Physical Society

Geometrical characteristics of nanotubes can be uniquely described by a chiral vector: Fig. 6.11a shows the chiral structure of iron single-walled nanotubes (SWNTs) [41]. The FeSWNTs can be considered by rolling up an iron monolayer. The circumferential configuration of a SWNT can be described by the chiral vector, $\mathbf{C}\mathrm{v} = m_{\mathbf{a}1} + n_{\mathbf{a}2} \equiv (m, n)$, where $\mathbf{a}_1$ and $\mathbf{a}_2$ are the primitive translation lattice vectors of the two-dimensional lattice. All independent configurations are represented by the chiral vectors within the irreducible range surrounded by the lines of $m = 2n$ and $m = n$.

Structural and energetics characteristics of SWNTs are shown in Fig. 6.11b [41]. One can find local minima in the total energy curve at the (4, 2), (5, 3), and (6, 3) SWNTs, which correspond to long-lived stable configurations in a freestanding condition [38]. On the other hand, the stability under a tip-suspended condition can be evaluated by a string tension concept f, which can be defined as the work needed to extract a nanotube out of the counterpart bulk metal, $f = (F - N\mu)/L_z$ [36], where μ is the chemical potential of the bulk iron and F is the free energy of a nanotube. The string tension f entirely increases with respect to the radius (see Fig. 6.11c), but a small drop can be found at the (4, 2), (5, 3), and (6, 3) configurations. The minima

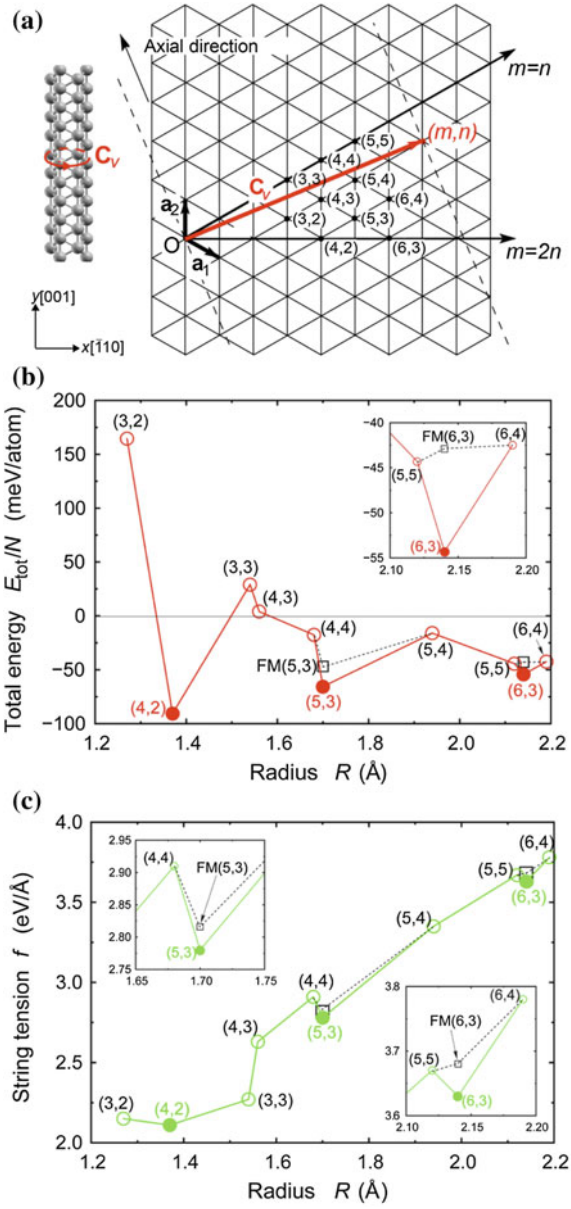

Fig. 6.11 **a** Chiral vector Cv on the two-dimensional lattice of the iron monolayer and atomic configurations of (*m*, *n*) FeSWNTs. **b** Total energy and **c** string tension *f* as a function of radius *R*. Reprinted with the permission from Ref. [41]. Copyright 2013 American Chemical Society

signify a long-lived "magic" structure under the tip-suspended condition, as has been shown for Au(5, 3) single-walled nanotubes [38]. As a whole, the (4, 2), (5, 3), and (6, 3) FeSWNTs are expected to be experimentally observed in a freestanding and a tip-suspended conditions.

These stable "magic" FeSWNTs also exhibit unique magnetic properties. Figure 6.12a shows the spin-spiral (SS) energy, $\Delta E_{SS-FM} = E_{SS} - E_{FM}$, as a

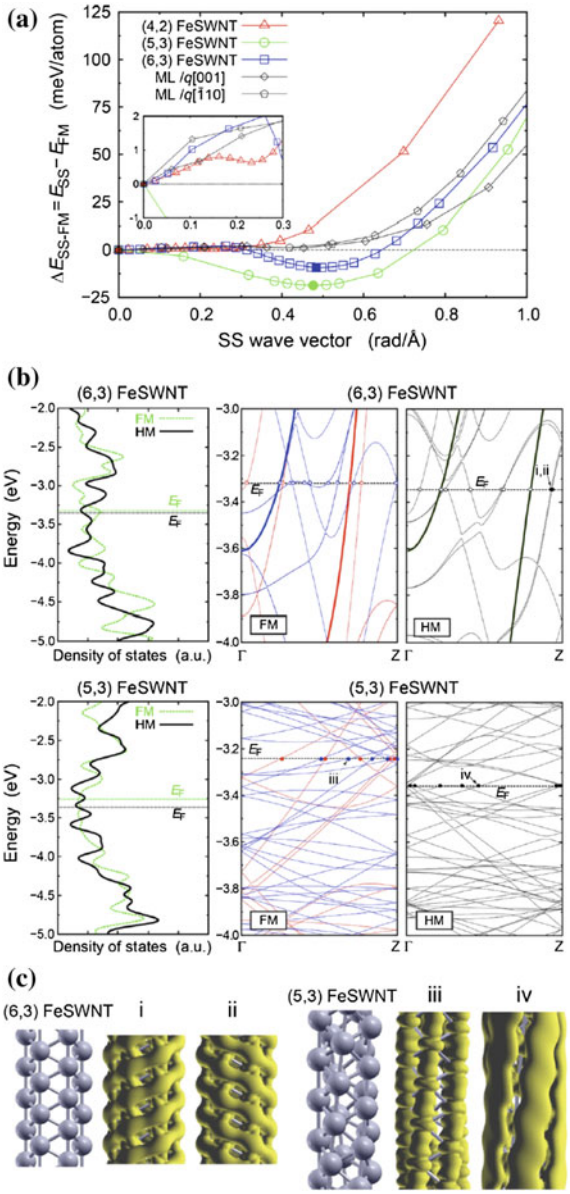

Fig. 6.12 **a** Spin-spiral (SS) excitation energy, $\Delta E_{SS-FM} = E_{SS} - E_{FM}$, as a function of wave vector q_z. **b** Electronic density of states (DOS) and band structures of the "magic" (6, 3) and (5, 3) FeSWNTs. The *red* and *blue lines* indicate the up (majority) and down (minority) spins, respectively. **c** Squared wave function of the states at the Fermi level, the states (i) and (ii) for (6, 3), and the states (iii) and (iv) for (5, 3), indicating the helical current channels. Reprinted with the permission from Ref. [41]. Copyright 2013 American Chemical Society

function of the SS wave vector q_z [41], where E_{SS} and E_{FM} are the total energies of the SWNT with a nonzero SS wave vector ($q_z \neq 0$) and that of the FM phase ($q_z = 0$), respectively. Although the flat ML (and energetically unfavorable FeSWNTs) is simply ferromagnetic, the "magic" (6, 3) and (5, 3) FeSWNTs show a minimum energy at the nonzero SS wave vectors, indicating the excitation of SS

wave, i.e., the noncollinear SS transition. This chirality dependence of formation of FeSWNT quantum helimagnets indicates the "chirally selectivity" of this phenomenon. Note that the SS wavelengths of the (6, 3) and (5, 3) SWNTs are $2\pi/q_z$ = 12.83 and 13.09 Å, respectively; i.e., the SS waves are incommensurate.

Such chiral-selective SS transition can be understood from the electronic point of view: Fig. 6.12b shows the electronic density of states (DOS) and band structures of the (6, 3) and (5, 3) FeSWNTs [41]. In the (6, 3) FeSWNT, the electronic bands below the Fermi level are located down to lower levels through the SS transition, while the bands above the Fermi level shift up, decreasing the Fermi energy and the total energy of the SS state with respect to the original ferromagnetic phase. This arises from the hybridization of the majority- and minority-spin bands via the SS wave excitation, so that a pair of bands are repulsively pushed away [42]. As a result, such interband repulsion via the hybridization stabilizes the noncollinear HM phase in the FeSWNTs. It should be noted that the flat ML has so narrow bandwidth that such hybridization of spins is less effective, which retains the ML with the FM phase.

In addition to the chirality-induced SS transition, chiral conductivity emerges: The current transporting channels of the SS FeSWNTs are visualized in Fig. 6.12c [41]. Some of channels form helical paths, i.e., chiral conductivity in the SS FeSWNTs. On the other hand, no such helical channel is present in the FM (6, 3) FeSWNT, meaning that the chiral conductivity is driven by the SS transition. This structure–magnetic–electronic relation suggests the existence of intriguing coupling between the chiral conductivity and helimagnetism in chiral FeSWNTs.

6.3.3 Atomic Chains

The ultimately small one-dimensional nanostructures should be atomic chains. Due to extremely small low coordination number and high structural degrees of freedom, the magnetic properties of atomic chains are strongly dependent on their exotic atomic structures changing with axial strains.

To include structural degrees of freedom of atomic chains into account, the Fe atomic chains are modeled by an ensemble of six atoms in a periodically repeated tetragonal cell with the axial dimension c along the z-axis [43]. The tension or compression on the atomic chains was simulated by changing the height of the cell. Figure 6.13a, b shows the atomic configuration of the chains and the binding energies of Fe atomic chains and the average magnetic moments, respectively [43]. The atomic configurations along the height of simulation cell are also shown. Under tensile strain (c > 14 Å), a dimerization of the chain, introducing two short and long distances, can be seen as an energetically favorable configuration. Such dimerization with short and long bonds results in a slight reduction in the energy and stabilization. By increasing tensile strain, such dimerization further stabilized, and the difference between short and long bonds becomes remarkable. A slight reduction in the magnetic moments is found due to the dimerization. Under compression, on the other hand, a strong reduction in magnetic moment is found: Under

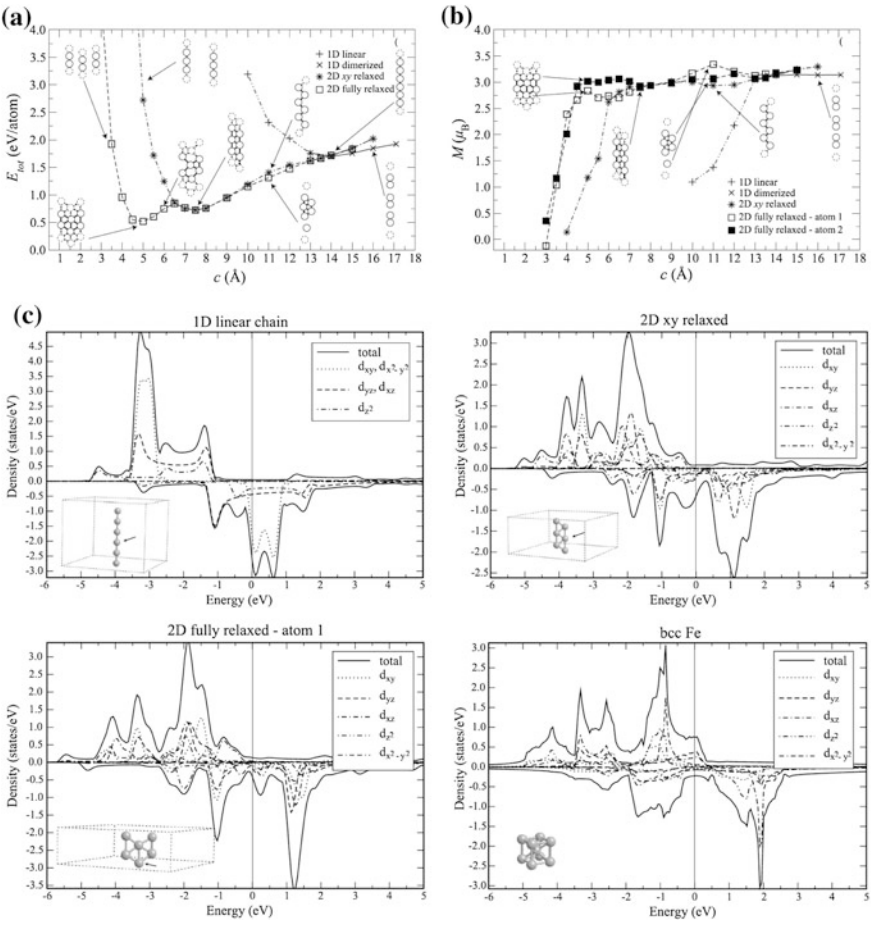

Fig. 6.13 **a** Total energy and **b** magnetic moment in iron atomic chains with respect to axial cell (strain). **c** Total and partial spin-polarized DOS of iron atomic chains, in comparison with that of bcc bulk counterpart. Reprinted with permission from Ref. [43]. Copyright 2009 by American Physical Society

the cell size within 7.0–13.5 Å, the zigzag chain configuration is formed via the gradual transformation to triangular stripes with reducing magnetic moments. This is because the interaction among not only the nearest neighbors (one-dimensional) but also the next nearest neighbors begins owing to the planar (two-dimensional) geometry. With further high compression, the triangular stripe chain transforms to a hexagonal one. The magnetic moment is higher for the atoms at the outer edge of the stripe 2.92 μ_B than for the central atoms 2.67 μ_B. This trend is similar as found in the nanowires (see the previous section). Finally, the atomic chain becomes nonmagnetic s at c = 4.5–6.0 Å with the stripe structures. In this way, the structural (dimensional) and magnetic degrees of freedom are strongly coupled with each other in the atomic chains.

The dimensionality of the atomic chain configuration corresponds to electronic properties. The total and partial spin-polarized electronic densities of states (DOSs) in atomic chains are shown in Fig. 6.13c [43]. In the linear atomic chain, the DOS consists of the largest ddσ bands by a d_{z2} orbital extended in the axial direction, the lower ddπ bands by d_{xz} and d_{yz} orbitals, and the lowest ddδ band by d_{xy} and d_{x2-y2} orbitals. On the other hand, in the triangular (or hexagonal) stripe chains with a planar geometry, the largest contribution to the bonds along the chain axis arises from ddσ bonds by a d_{z2} orbital and ddπ bonds by a d_{yz} orbital, whereas ddσ bonds by a d_{xz} orbital contributes to the transverse bonds. As a result, the Fermi level goes down to the minimum of the minority states by the bonding–antibonding splitting. This broader bandwidth due to the planar geometry results in the high magnetic moment of 2.92 μ_B.

It should be remarked that the most significant change of magnetism in atomic chains was demonstrated for Mn, which exhibits various collinear and noncollinear magnetic states even in the bulk. Figure 6.14 shows the noncollinear magnetic structures in various local equilibrium atomic chain configurations [44]. One can

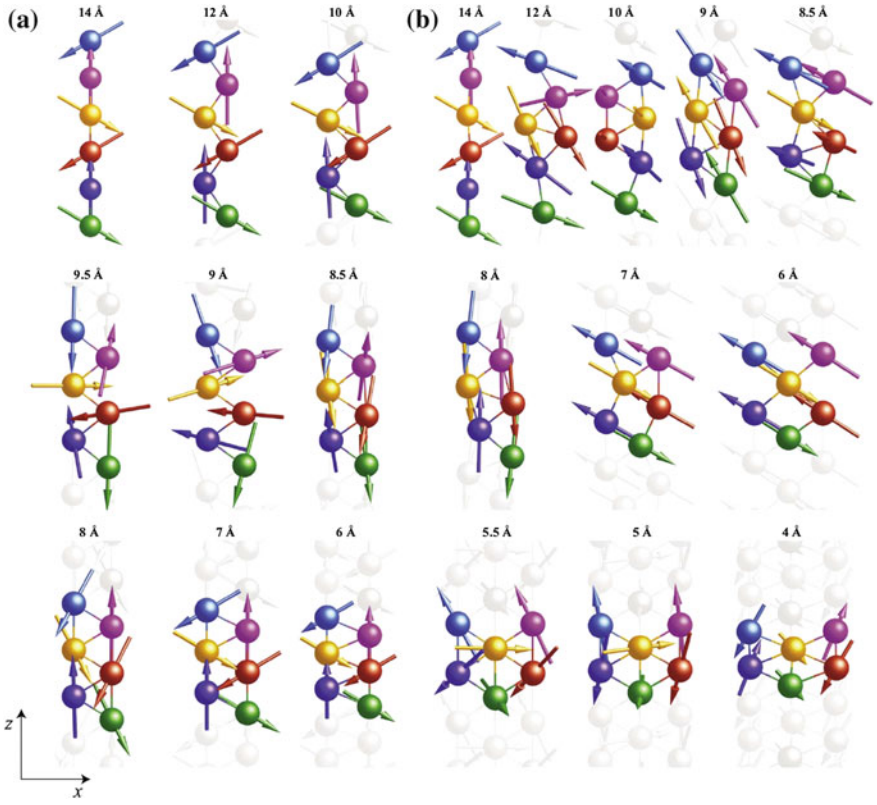

Fig. 6.14 Atomic and magnetic structures of Mn atomic chains as a function of axial length (*strains*). Reprinted with permission from Ref. [43]. Copyright 2009 by American Physical Society

find not only collinear FM and AFM, but also noncollinear SS-, Néel-, and 3Q-type states depending on the Mn atomic chain configurations. This suggests that the magnetism in Mn atomic chains is much sensitive to strain. It is thus interesting to study magnetic transitions in other transition metal atomic chains.

6.4 Atomic Cluster: Zero-Dimensional Structure

Atomic clusters are the representative of 0D nanostructures. Here, magnetic properties in atomic clusters consisting of a small number of atoms are reviewed. Figure 6.15 shows the atomic configuration and three-dimensional magnetization vectors in atomic clusters of Cr_n ($n \leq 5$) [45]. The fully unconstrained noncollinear magnetism calculations [46] are used to evaluate the noncollinear ground states of atomic clusters. Conventional FM and AFM configurations are found in a simple atom Cr_1 and dimer Cr_2, respectively, due to no or very few structural and magnetic degrees of freedom. On the other hand, noncollinear magnetism appears in the atomic clusters of $Cr_3 - Cr_5$.

The Cr_3 atomic cluster forms the distorted triangular atomic configuration and exhibits the noncollinear "spiral" arrangement, as shown in the left panel of Fig. 6.15 [45]. In general, the interatomic exchange interaction of Cr atoms is antiferromagnetic, and this magnetic configuration can be regarded as the conventional frustration of magnetic moments in the antiferromagnetic trimer. Since Cr atoms possess a localized magnetization, its change from one direction of magnetization to another goes through zero magnetization like "Bloch wall." In a similar way, tetrahedral Cr_4 is also showing noncollinear configuration of magnetic moments. The spin-spiral arrangement is found with the magnetization directions

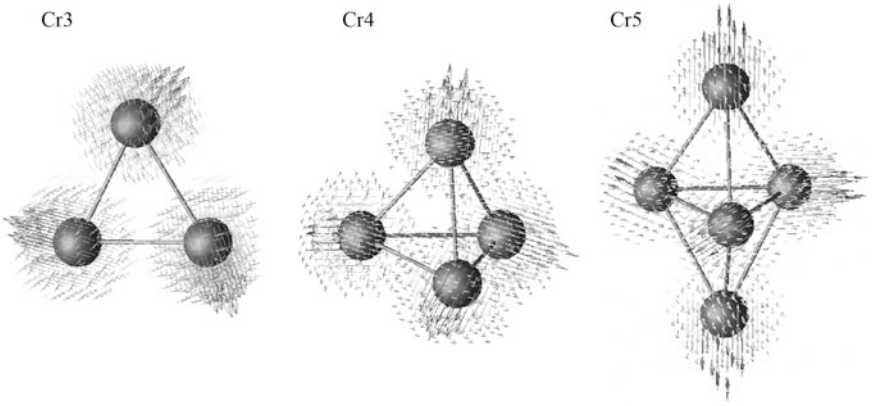

Fig. 6.15 (Meta)stable atomic and magnetic configurations of atomic clusters Cr_n ($n \leq 5$). The magnetization density is visualized for $Cr_3 - Cr_5$ atomic clusters. Reprinted with permission from Ref. [45]. Copyright 2000 by American Physical Society

on the neighboring atoms keeping an angle of 109.5° (see the middle panel of Fig. 6.15 [45]). The spin-spiral configuration in Cr_4 originates from the frustration of neighboring magnetic moments. In addition, this noncollinear configuration results in large magnetic moment of 4.02 μ_B/atom, suggesting a strong enhancement of magnetism. On the other hand, a Cr_5 cluster forms a bipyramidal pentamer with a ferrimagnetic ground state, but the apical atoms are spin-up while the central triangular atoms are spin-down (see the right panel of Fig. 6.15 [45]). In this way, the atomic clusters exhibits unique magnetic phases and orders strongly coupled with their geometrical characteristics.

6.5 Conclusion

In this chapter, we reviewed a series of studies on magnetism from bulk to nanostructures. In Sect. 6.1, we first introduced ferromagnetic iron being quite sensitive to strain through the ferromagnetic to antiferromagnetic phase transition along the deformation paths. We also briefly reviewed first-principles studies of internal defects as a representative of understructures. All defects including grain boundaries (2D), dislocations (1D), and vacancies (0D) locally enhance the magnetic properties due to its low coordination number and local strains.

After that, in Sects. 6.2–6.4, we discussed the magnetism in low-dimensional (2D-0D) nanostructures. By the reduction of dimensionality, from bulk to nanofilm and nanowires, the magnetic moment increases and stabilizes the ferromagnetic states over the other magnetic states. On the other hand, in an extremely thin geometries, such as monolayers, single-walled nanotubes, atomic chains, and clusters, the unexpected noncollinear magnetic phases including the spin-spiral and frustrated magnetism emerge due to their exotic atomic arrangements distinct from the bulk counterpart. A series of studies demonstrates the great potential of designing and controlling the unusual magnetic phase and properties through the low-dimensional nanostructures and mechanical strains.

References

1. C. Heck, Magnetic materials and their applications (Butterworth, London, 1974)
2. M. Friak, M. Sob, V. Vitek, Phys. Rev. B **63**, 052405 (2001)
3. P.J. Craievich, M. Weinert, J.M. Sanchez, R.E. Watson, Phys. Rev. Lett. **72**, 3076 (1994)
4. D.J. Singh, W.E. Pickett, H. Krakauer, Phys. Rev. B **43**, 11628 (1991)
5. P. Bagno, O. Jepsen, O. Gunnarsson, Phys. Rev. B **40**, 1997 (1989)
6. H. Li, Y.S. Li, J. Quinn, D. Tian, J. Sokolov, F. Jona, P.M. Marcus, Phys. Rev. B **42**, 9195 (1990)
7. M. Stampanoni, A. Vaterlaus, M. Aesschliman, F. Meier, Phys. Rev. Lett. **59**, 2483 (1987)
8. J. Quinn, Y.S. Li, H. Li, D. Tian, F. Jona, P.M. Marcus, Phys. Rev. B **43**, 3959 (1991)
9. E.G. Moroni, G. Kresse, J. Hafner, Phys. Rev. B **56**, 15629 (1997)

10. C. Elsässer, J. Zhu, S.G. Louie, M. Fa¨hnle, and C. T. Chan. J. Phys Condens. Matter **10**, 5081 (1998)
11. H.C. Herper, E. Hoffmann, P. Entel, Phys. Rev. B **60**, 3839 (1999)
12. K. Hampel, D.D. Vvedensky, S. Crampin, Phys. Rev. B **47**, 4810 (1993)
13. A. Mokrani, C. Demangeat, H. Dreysse, Phys. Rev. B **42**, 8670 (1990)
14. A. Gonis, X.-G. Zhang, J.M. MacLaren, S. Crampin, Phys. Rev. B **42**, 3798 (1990)
15. M. Mrovec, D. Nguyen-Manh, C. Elsösser, P. Gumbsch, Phys. Rev. Lett. **106**, 246402 (2011)
16. L. Ventelon, F. Willaime, J. Comput. Aided Mater. Des. **14**, 85 (2007)
17. G. Liu, D. Nguyen-Manh, B.G. Liu, D.G. Pettifor, Phys. Rev. B **71**, 174115 (2005)
18. T. Shimada, Y. Ishii, T. Kitamura, Phys. Rev. B **81**, 134420 (2010)
19. P.E. Blöchl, Phys. Rev. B **50**, 17953 (1994)
20. G. Kresse, D. Joubert, Phys. Rev. B **59**, 1758 (1999)
21. J.B. Goodenough, Phys. Rev. **120**, 67 (1960)
22. Y.Q. Xie, Acta Metall. Mater. **42**, 3705 (1994)
23. F. Liu, S.N. Khanna, P. Jena, Phys. Rev. B **43**, 8179 (1991)
24. M. Pratzer, H.J. Elmers, M. Bode, O. Pietzsch, A. Kubetzka, R. Wiesendanger, Phys. Rev. Lett. **87**, 127201 (2001)
25. D. Hobbs, G. Kresse, J. Hafner, Phys. Rev. B **62**, 11556 (2000)
26. C. Herring, *Magnetism* (Academic, New York, 1966)
27. L.M. Sandratskii, J. Phys Condens. Matter **3**, 8565–8585 (1993)
28. T. Shimada, J. Okuno, T. Kitamura, Phys. Rev. B **85**, 134440 (2012)
29. M. Pratzer, H.J. Elmers, Phys. Rev. B **67**, 094416 (2003)
30. J.C. Tung, G.Y. Guo, Phys. Rev. B **83**, 144403 (2011)
31. M. Zelený, M. Sob, J. Hafner, Phys. Rev. B **80**, 144414 (2009)
32. T. Shimada, Y. Ishii, T. Kitamura, Phys. Rev. B **84**, 174405 (2011)
33. T.E. Jones, M.E. Eberhart, D.P. Clougherty, Phys. Rev. Lett. **100**, 017208 (2008)
34. Y. Oshima, A. Onga, K. Takayanagi, Phys. Rev. Lett. **91**, 205503 (2003)
35. Y. Kondo, K. Takayanagi, Science **289**, 606–608 (2000)
36. E. Tosatti, S. Prestipino, S. Kostlmeier, A. Dal Corso, F.D. Di Tolla, Science **291**, 288–290 (2001)
37. M.J. Lagos, F. Sato, J. Bettini, V. Rodrigues, D.S. Galvão, D. Ugarte, Nat. Nanotechnol. **4**, 52–149 (2009)
38. S.L. Elizondo, J.W. Mintmire, Phys. Rev. B **73**, 045431 (2006)
39. R.T. Senger, S. Dag, S. Ciraci, Phys. Rev. Lett. **93**, 196807 (2004)
40. B. Wang, D. Shi, J. Jia, G. Wang, X. Chen, J. Zhao, Physica E **30**, 45–50 (2005)
41. T. Shimada, J. Okuno, T. Kitmura, Nano Lett. **13**, 2792 (2013)
42. R. Lizárraga, L. Nordström, L. Bergqvist, A. Bergman, E. Sjöstedt, P. Mohn, O. Eriksson, Phys. Rev. Lett. **93**, 107205 (2004)
43. M. Zelený, M. Šob, J. Hafner, Phys. Rev. B **79**, 134421 (2009)
44. M. Zelený, M. Šob, J. Hafner, Phys. Rev. B **80**, 144414 (2009)
45. D. Hobbs, G. Kresse, J. Hafner, Phys. Rev. B **62**, 11556 (2000)
46. T. Oda, A. Pasquarello, R. Car, Phys. Rev. Lett. **80**, 3622 (1998)

Chapter 7
Multiferroic Nanostructures

Abstract Multiferroics, two or more ferroic orders such as ferroelectricity and antiferromagnetism coexist and coupled with each other, is a tremendously fascinating topic in terms of scientific interests as well as potential industrial applications. Investigations of $BiFeO_3$, a prototypical intrinsic multiferroic material, are introduced. The material's sensitive response to strain, i.e., interaction between ferroelectrics and magnetics, is exemplified. After briefly reviewing first-principles studies on the effect of understructures, multiferroic and magnetoelastic properties in low-dimensional nanostructures are discussed. We also present an extrinsic (defect-induced) multiferroics with atomic-scale structures.

Keywords Multiferroicity · Magnetoelectric coupling · Multiferroic transition · Strain effect

7.1 Multiferroicity in Bulk

7.1.1 Multiferroic Properties and Response to Strain

In previous chapters, we have already reviewed ferroelectricity and magnetism each. These properties are, however, mutually exclusive because a formal d^0 (empty) electron configuration of the transition-metal cations (e.g., Ti^{4+} in $PbTiO_3$) drives their positions off-center in classical ferroelectrics, while a partially filled d state is required for magnetism [1]. This competition has prevented the realization of intrinsic multiferroic materials for a long time.

In recent years, materials that exhibit both ferroelectricity and antiferromagnetism have been discovered [2, 3]. These materials are called as "multiferroics." These two ferroic orderings not only coexist in the material, but they also strongly couple with each other. The coupling effect is known as magnetoelectric (ME) coupling, which enables spontaneous magnetization (polarization) to be controlled by applying an external electric (magnetic) field [2–6]. Owing to the intriguing properties,

Y. Umeno et al., *Multiphysics in Nanostructures*, Nanostructure Science and Technology, DOI 10.1007/978-4-431-56573-4_7

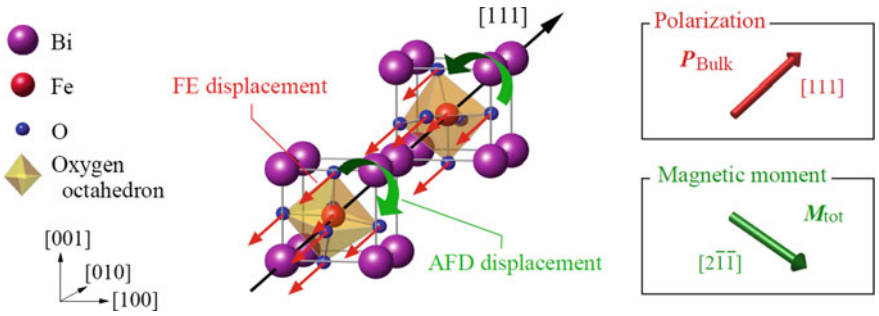

Fig. 7.1 Crystal structure of rhombohedral $BiFeO_3$. There exist FE displacement and AFD displacement in $BiFeO_3$. P_{Bulk} and M_{tot} indicate the polarization and the magnetic moment in $BiFeO_3$, respectively. Reprinted with permission from Ref. [12]. Copyright 2014 by American Physical Society

multiferroics have the potential to be used in next-generation devices, such as multistate memory elements, transducers, sensors, and spintronics devices [7–11].

$BiFeO_3$ is a representative and prototype of multiferroics. Figure 7.1 shows the unit cell of bulk rhombohedral (R3c) $BiFeO_3$ [12]. Mainly two types of structural distortions are found in the $BiFeO_3$ lattice: (i) ferroelectric (FE) distortions associated with an irreducible representation of Γ^{-4} and (ii) antiferrodistortive (AFD) rotations of O_6 octahedra corresponding to the R^{+4} mode. The FE distortions of bulk $BiFeO_3$ are in the [111] direction in the figure, resulting in a spontaneous polarization, P, in the same direction. The magnetic moments are essentially in the rock salt antiferromagnetic (G-type AFM) order (see M_{Fe1} and M_{Fe2} Fig. 7.2a) but with a small spin-canting due to the R^{+4} mode (see Fig. 7.2b) [13]. The canting of the magnetic moments results in a macroscopic magnetization, M, in the [2-1-1] direction.

As reviewed in previous chapters, ferroelectricity and magnetism is quite sensitive to strains, which is also true in multiferroics: Fig. 7.3 shows the spontaneous

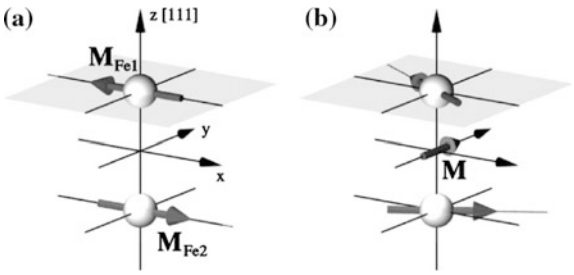

Fig. 7.2 **a** Rock salt antiferromagnetic (G-type AFM) order in $BiFeO_3$. M_{Fe1} and M_{Fe2} show the magnetic moment of iron in $BiFeO_3$. **b** The spin-canting and resulting total magnetization M in $BiFeO_3$. Reprinted with permission from Ref. [13]. Copyright 2005 by American Physical Society

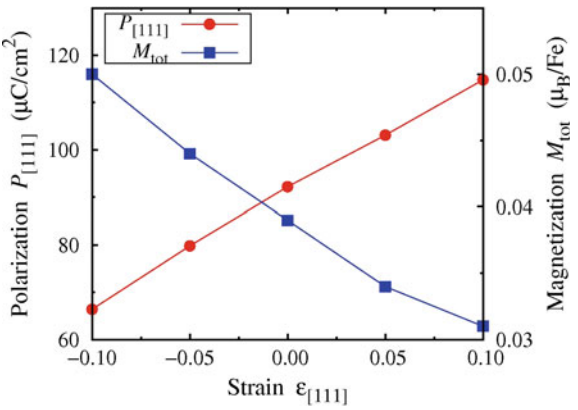

Fig. 7.3 Spontaneous polarization $P_{[111]}$ and total magnetic moment M_{tot} as a function of strain $\varepsilon_{[111]}$ in $BiFeO_3$. Reprinted with permission from Ref. [14]. Copyright 2014 by the Society of Materials Science, Japan

polarization and magnetic moment as a function of strain in $BiFeO_3$ [14]. The tensile (or compressive) strain tends to increase (or decrease) the spontaneous polarization but decrease (or increase) magnetic moment. As a result, ferroelectricity and magnetism exhibits quite opposite strain-responses. This suggests that ferroelectricity and magnetism coexists but compete with each other in multiferroic $BiFeO_3$. This interaction (or competition) between two different ferroic properties leads to distinct behaviors in nanostructures from pure ferroelectrics or magnets.

7.1.2 Multiferroic Domains and Domain Walls

Although we reviewed multiferroic properties in single-crystal and single-domain $BiFeO_3$ above, multiferroics, in general, consists of ferroelectric/ferromagnetic domains, one of characteristic and representative understructures of multiferroics. The multiferroic domains are characterized by domain walls. Here, we review the multiferroic properties of domain walls in $BiFeO_3$.

Figure 7.4 shows the polarization distribution along the domain walls in $BiFeO_3$ [15]. Since the $BiFeO_3$ is rhombohedral and its polar axis is [111], a possible domain wall (DW) configurations are limited to 71°, 109°, and 180° walls. In all cases, the polarization smoothly rotates from its direction across the wall with a DW width of two or three lattice constants (approximately 1 nm). As simultaneously shown in Fig. 7.4, macroscopic electrostatic potentials have discontinuous step due to the polarization change at the DW, especially in 109° DW. This is quite similar as 90° DWs in $PbTiO_3$ shown in previous chapters. Such DW characteristics are related to structural distortions: Fig. 7.5 shows the atomic and polarization configurations around DWs [15]. As clearly seen in the 71° DW, the polarization rotation across the DW is associated with rhombohedral to pseudocubic lattice deformations. Such lattice deformation may change the polarization rotation smoothly at the DW and contribute to an atomically thin wall thickness.

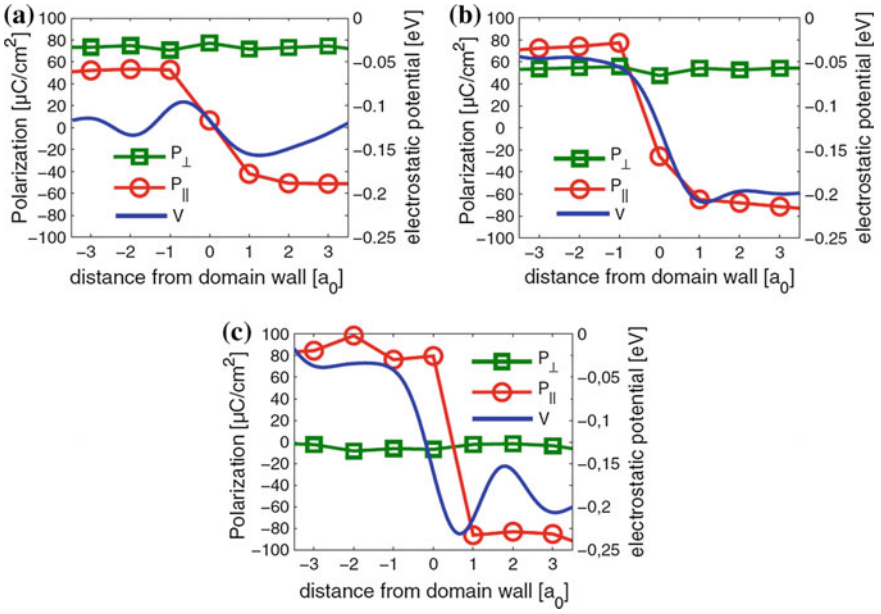

Fig. 7.4 Polarization distribution across the **a** 71°, **b** 109°, and **c** 180° domain walls in $BiFeO_3$. $P\perp$ and $P\parallel$ denote normal and parallel components of polarization, respectively. Figure includes the planar-averaged electrostatic potential V. Reprinted with permission from Ref. [15]. Copyright 2009 by American Physical Society

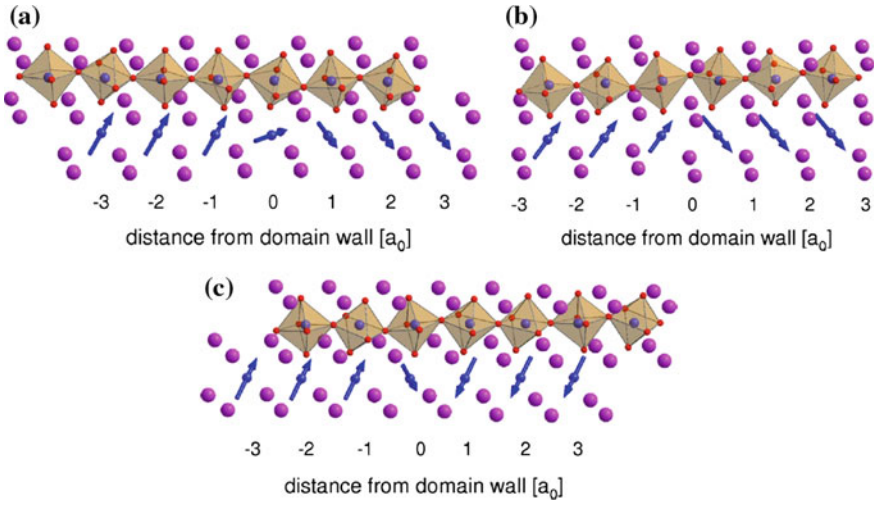

Fig. 7.5 Structural and polarization changes across **a** 71°, **b** 109°, and **c** 180° domain walls in $BiFeO_3$. *Blue arrows* represent the magnitude and orientation of the local polarization. Reprinted with permission from Ref. [15]. Copyright 2009 by American Physical Society

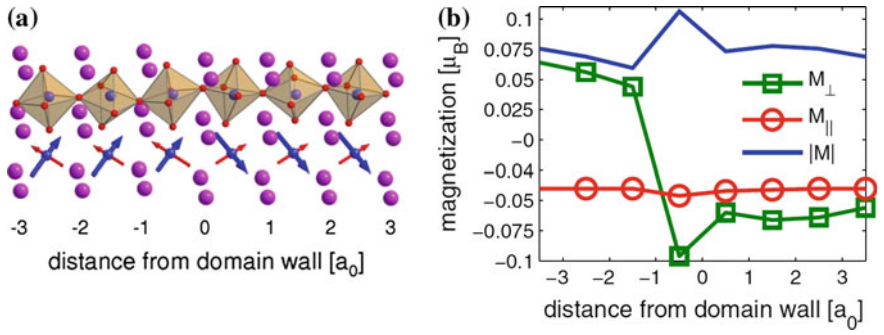

Fig. 7.6 Magnetic moment distribution across the 109° DW. *Red* and *blue arrows* represent the local magnetic moment and polarizations, respectively, in the panel (**a**). $M\perp$ and $M_{\parallel}$ denote normal and parallel components of magnetic moment, respectively, in the panel (**b**). Reprinted with permission from Ref. [15]. Copyright 2009 by American Physical Society

Coupled with the rotation of polarizations, magnetization also changes its directions smoothly at the wall. Figure 7.6 shows the magnetization distribution across the 109° DW [15]. In the domains, the magnetic moment is [11-2] on one side and [112] on the other side of the wall as same as the bulk value. Importantly, the spin-canting increases by 33% at the DW, corresponding to a larger Fe–O–Fe bond angles. Again, the magnetization direction changes smoothly across the wall, and the magnetic domain wall thickness is estimated to be about 1 nm, as same as the ferroelectric domain wall thickness [16]. This consistency means that ferroelectric and magnetic rotations at the DW region are coupled with each other, leading to the formation of simultaneously "multiferroic" DWs where electric and magnetic order changes at the same time.

7.1.3 Atomic Defects

Another important understructure in multiferroics is lattice defects such as grain boundaries [17], dislocations [18], and atomic defects [19–21]. Here, we review the effect of atomic vacancies on multiferroic properties.

Figure 7.7 shows the ferroelectric polarization field near an atomic vacancy V_i (i = O, Bi, and Fe) in multiferroics $BiFeO_3$ [22]. To show the intrinsic effect of a vacancy itself, the difference of polarization field between the system including a vacancy and perfect (defect-free) $BiFeO_3$, $\Delta \boldsymbol{P} = \boldsymbol{P}_{\text{vacancy}} - \boldsymbol{P}_{\text{perfect}}$, is presented. The effect of the vacancy ranges a two-unit-cell length near the vacancy site. The neutral V_O changes polarization field very slightly and the effect is trivial, while it becomes remarkable when the oxygen vacancy is charged. This suggests that the effect of charge state is quite strong. This extrinsic effect of charged V_O directs the polarization outward to the vacancy site. On the other hand, the Bi vacancy drives

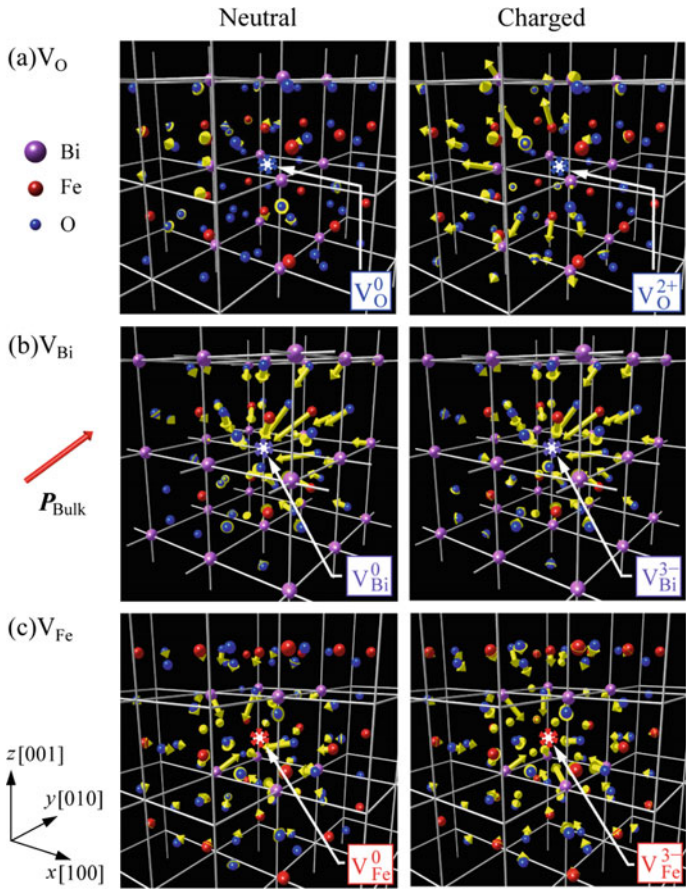

Fig. 7.7 Ferroelectric polarization field **a** V_O, **b** V_{Bi}, and **c** V_{Fe} in multiferroic $BiFeO_3$. The spontaneous polarization of 89.4 μC/cm^2 in perfect $BiFeO_3$ is shown by the *red arrow* for comparison. Reprinted with permission from Ref. [22]. Copyright 2016 by American Physical Society

an isotropic and inward polarization, which is remarkable at the [111] side of the vacancy (see Fig. 7.7b) [22]. This is opposite to the spontaneous polarization of the $BiFeO_3$ matrix, and thereby, V_{Bi} suppresses ferroelectricity. This extrinsic polarization induced by V_{Bi} is, however, quite insensitive to the charge state. Such trend is also found in V_{Fe}, as shown in Fig. 7.7c [22]. Thus, the ferroelectric nature of vacancies exhibits strong dependence on the cationic/anionic character and/or charge state.

Atomic vacancies strongly affect the magnetic properties, as well. Figure 7.8 shows the magnetic properties of each vacancy. In the figure, the difference in the

magnetic moments from those of perfect $BiFeO_3$, $\Delta \boldsymbol{M} = \boldsymbol{M}_{\text{vacancy}} - \boldsymbol{M}_{\text{perfect}}$, is visualized to clarify the vacancy-driven magnetic moments [22]. From Fig. 7.8a, neutral V_O locally suppresses magnetic moments of neighboring Fe atoms. However, such effect disappears when oxygen vacancies are charged. In contrast, Bi vacancies provide no such significant change regardless of the charge states (Fig. 7.8b), while iron vacancies have strong influence on magnetic moments (Fig. 7.8c). In particular, neutral Fe vacancies decrease magnetization for the three Fe atoms neighboring to the vacancy (0.74 μ_B), in addition to the direct loss of an iron atom (4.08 μ_B). As a result, V_{Fe} leads to a magnetic moment of 1.87 μ_B along [10-1]. In this way, only neutral oxygen and Fe vacancies can exhibit the extrinsic magnetic properties in addition to those of the host $BiFeO_3$ in contrast to the ferroelectricity of vacancies shown in the previous section.

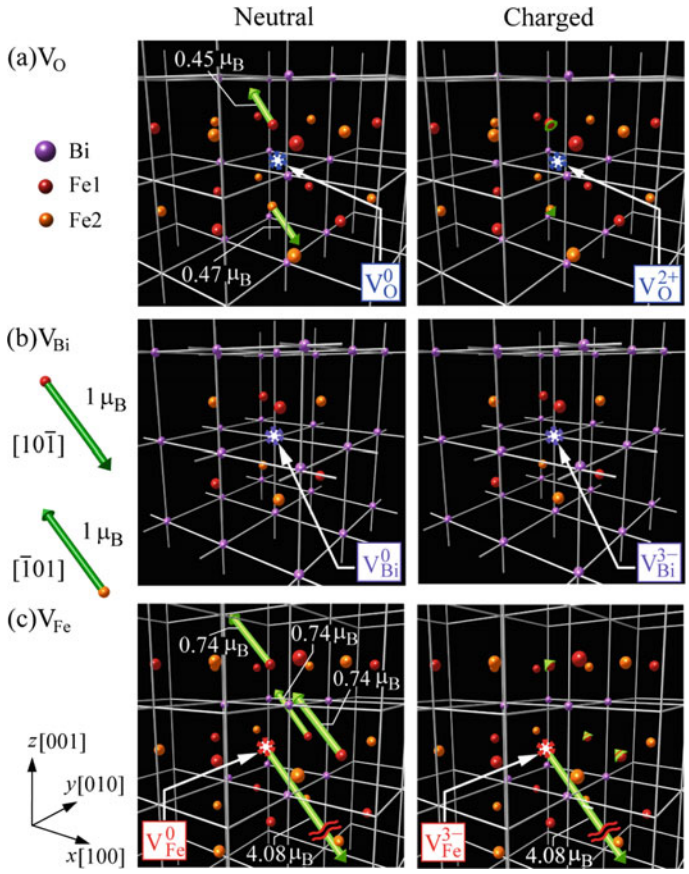

Fig. 7.8 Magnetic moment field near **a** V_O, **b** V_{Bi}, and **c** V_{Fe} in multiferroic $BiFeO_3$. Reprinted with permission from Ref. [22]. Copyright 2016 by American Physical Society

7.2 Multiferroicity in Nanostructures

7.2.1 *Nanofilms and Surface Properties: Two-Dimensional Structure*

Surfaces

As two-dimensional geometry, surfaces provide intriguing properties due to low coordinate numbers being different from those of bulk system. It is therefore worthwhile to extensively investigate surface effects on multiferroic properties. However, in contrast to purely ferroelectrics or magnets, there have been little ab initio studies on surface structures and properties for multiferroics, probably due to complexity of phenomena and difficulty in modeling the surface structures of such low-symmetry crystal structure of $BiFeO_3$ under periodic boundary conditions, which usually applied to the DFT calculation method as was mentioned in the previous section. As a successful example of very few studies on multiferroic surfaces, hereafter we review $BiFeO_3$ (110) surface properties [12].

Figure 7.9 shows the spontaneous polarization $\boldsymbol{P}$ at $BiFeO_3$ (110) surfaces [12]. Due to rhombohedral crystal lattice and G-type AFM configuration of $BiFeO_3$, the surface layer consists of two independent unit cells (Cell A and Cell B). At the O-terminated surface, the planar component of polarizations P_x and P_y is smaller than those for the bulk and the out-of-plane P_z becomes nonzero (Fig. 7.9b), while in the bulk counterpart, $\boldsymbol{P}$ aligns purely parallel to the (110) surface, i.e., the surface rotates spontaneous polarization (Fig. 7.9a). The rotational angle, θ, from the original direction is 11.7°, and its direction is approximately [8–55]. Moreover, each component of $\boldsymbol{P}$ in Cells A and B is slightly different from each other, suggesting an antiferroelectric (AFE) state emerging in addition to the ferroelectric (FE) state at the surface layer. Similarly, such polarization rotation can also be observed in the FeO- and Bi-terminated surfaces with different angles and orders (Fig. 7.9c, d): θ = 4.3° for the FeO-termination ([6-67]), and θ = 12.0° for the Bi-termination ([4-75]). These surface effects range from 8 to 9 layers below the surface layer (approximately 1 nm). As a consequence, spontaneous polarization is rotated and suppressed at the (110) surfaces due to low coordination number and unique AFE configuration. In fact, experimental observations showed that the spontaneous polarization of (110) $BiFeO_3$ thin films is smaller than those of bulk [23, 24], which corresponds to theoretical results. It is also noted that the effects of surface charges and the resulting depolarizing field should be little at $BiFeO_3$ (110) surfaces due to the small canting of polarization. In contrast, $BiFeO_3$ thin films with (001) or (111) surfaces are expected to possess much larger surface charges. Thin films with these surfaces may exhibit the significant suppression of polarization or may form polydomain configuration in order to compensate the surface charges. This effect will be presented in the following section.

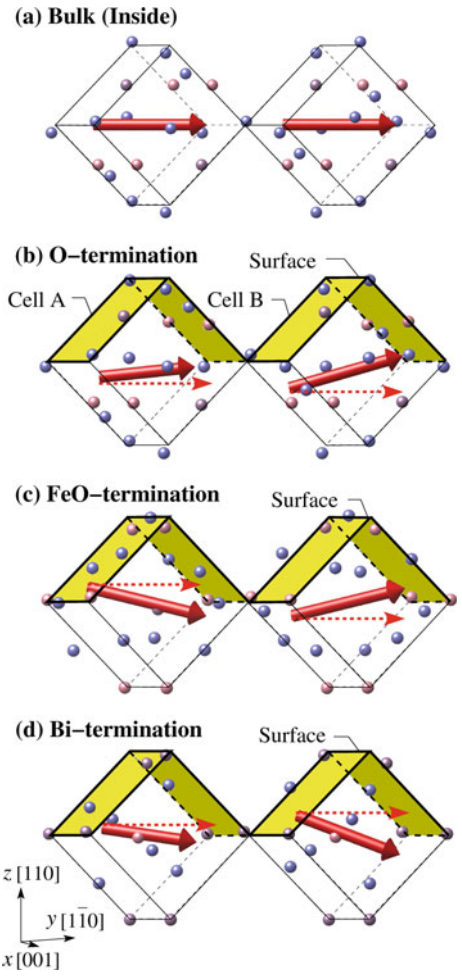

Fig. 7.9 Polarization $\boldsymbol{P}$ distribution in **a** $BiFeO_3$ bulk, and at **b** O-terminated, **c** FeO-terminated, and **d** Bi-terminated surfaces. The *dashed red arrows* in **b**–**d** show the bulk polarization for comparison. Reprinted with permission from Ref. [12]. Copyright 2014 by American Physical Society

Magnetism is also coupled to rotate at the surface in $BiFeO_3$. Figure 7.10 shows the magnetic moment at $BiFeO_3$ (110) surfaces [12]. In bulk $BiFeO_3$, magnetic moment of each iron atom $\boldsymbol{M}_{\mathrm{Fe}}$ lies almost along the z (out-of-plane) direction with small spin-canting (canting angle ϕ: 1.0°), and the total magnetic moment, $\boldsymbol{M}_{\mathrm{tot}}$, of the two Fe atoms is on the x-y plane, and its direction is [1] (Fig. 7.10a). On the other hand, the in-plane component of magnetic moment M_x and M_y is several times larger at the O-terminated surface than those for the bulk, leading to the much larger spin-canting (canting angles ϕ of Fe A and B are 4.8° and 3.7°, respectively) (Fig. 7.10b). Such relatively large canting of $\boldsymbol{M}_{\mathrm{Fe}}$ at the surface leads to the rotation of total magnetization $\boldsymbol{M}_{\mathrm{tot}}$ to [11-6]. A similar trend can be also observed at the different surface terminations (Fig. 7.10c, d). These surface effects on the magnetic moments diminish as the distance from the surface increases, disappearing about 11

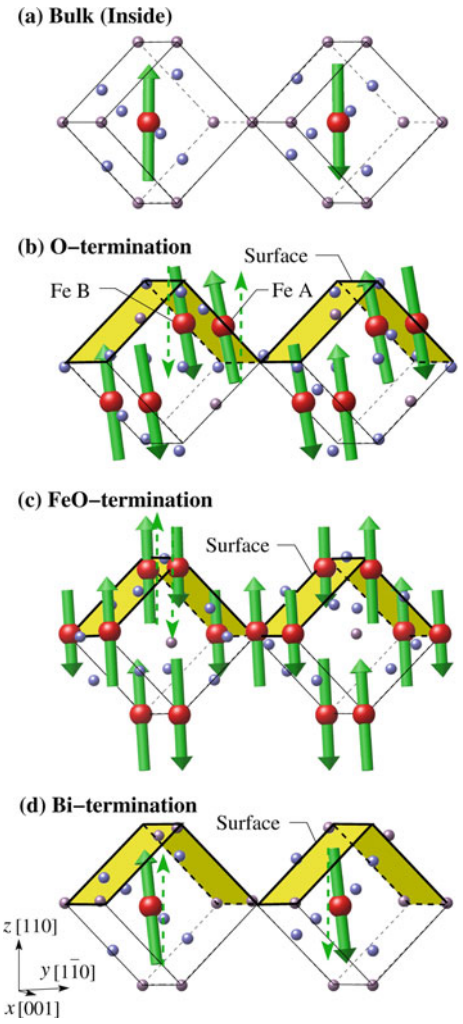

Fig. 7.10 **a** Magnetic moment of iron atoms $\boldsymbol{M}_{\mathrm{Fe}}$ in **a** $BiFeO_3$ bulk, and at **b** O-terminated, **c** FeO-terminated, and **d** Bi-terminated surfaces. The *dashed red arrows* in **b**–**d** show the bulk magnetic moments for comparison. Reprinted with permission from Ref. [12]. Copyright 2014 by American Physical Society

layers beneath the surface. In such way, surfaces drive the spin-canting and rotation of magnetic moments. As a whole, surfaces tend redirect multiferroic polarization and moment, which is distinct from the surface characteristics observed in pure ferroelectrics or magnetism as reviewed in previous chapters.

Nanofilms

For multiferroic thin films with a ferroic polar axis perpendicular to their surfaces, the termination of polarization at the surface or interface generates surface charges, which leads to the formation depolarization (or demagnetization) field that destabilizes the ferroic order. The thin-film system is naturally partitioned into domains to compensate these effects, leading to an exotic multiferroic domain configuration.

An interesting example is a $BiFeO_3$ ultrathin film with polar (001) surfaces. An ultimately thin thickness of three unit cells confines and enhances the intrinsic surface effects and depolarization, leading to a polydomain configuration. Figure 7.11 shows the ferroelectric distribution of $\boldsymbol{\delta}_{FE}$ in the three-unit-cell thick ($m = 3$) $BiFeO_3$ ultrathin film with polar (001) surfaces [25]. To achieve a very high resolution of polarization, ferroelectric distortion $\boldsymbol{\delta}_{FE}$ is introduced for each atomic layer,

$$\boldsymbol{\delta}_{FE} = \begin{cases} \boldsymbol{d}(\mathrm{Bi}) - \boldsymbol{d}(\mathrm{O}), & (\mathrm{BiO\ layer}) \\ \boldsymbol{d}(\mathrm{Fe}) - \boldsymbol{d}(\mathrm{O}), & (\mathrm{FeO_2\ layer}) \end{cases} \tag{7.1}$$

where $\boldsymbol{d}(i)$ is the unit-layer-averaged displacement of atom i (i = Bi, Fe, O) relative to ideal lattice sites. The thin-film system is mainly decomposed into two domains: [111] polarization at the left-side and [1] polarization at the right-side domains. These domains are periodically repeated along the thin film, so that the film consists of polydomains with a domain period of 8 unit cells. To see more details, one can find an interesting polar order at junction of these domains: At junction A, the direction of ferroelectric distortion gradually changes from [111] to [1] forms a pure 109° DW. On the other hand, at the opposite side of junction B, ferroelectric distortion locally rotates with shifting its direction from [111], through [1] and [-11-1], to [-1-11]. As a result, the film locally exhibits vortex domain structure which consists of several 109° DWs at junction B (see Fig. 7.11b–2). Note that such characteristic vortex structure can also be observed in the thinner films, as well. Such a vortex-like domain configuration can also be seen in ferroelectric $PbTiO_3$ thin films as reviewed in the previous chapter. However, vortex character is different from each other; vortices in $BiFeO_3$ consist of 109° DWs, while vortices (or closure domain) in $PbTiO_3$ consist of 180° and 90° DWs.

Through the formation of periodic polydomain with polarization vortices, magnetism also forms corresponding domains via multiferroic coupling. Figure 7.12a shows the distribution of magnetic moment in the same $BiFeO_3$ ultrathin film with eight-unit-cell domain period. As the net magnetic moment in $BiFeO_3$ is induced by the small canting of local magnetic moment, M_{Fe}, of each Fe atom from the G-type AFM order, the total magnetic moment, M_{tot}, of the two adjacent Fe atoms along the same y axis is also shown in Fig. 7.12b. At the middle of the domains, M_{tot} (shown as blue arrows) has the magnitude which is no more than 1–3 times compared to that in the bulk, and it is oriented to the similar direction to the bulk value, meaning the bulk-like ordering of magnetization inside of the film. In contrast, M_{tot} near the junction of the surfaces and DWs (shown as red arrows) is 3–9 times larger than the bulk value, and its direction differs from the original one (e.g., the largest M_{tot} in the left-side domain is oriented to almost [01-4], while [2-1-1] in the bulk). This indicates that the junction of the surfaces and DWs remarkably enhances the magnetic anisotropy. Either surface or DW alone can cause the change of magnetization, but its effect is relatively small, as reviewed above. In fact, M_{tot} near the 109° DWs in the bulk slightly increases by several dozen percent, while its direction remains almost the same as the middle of domains

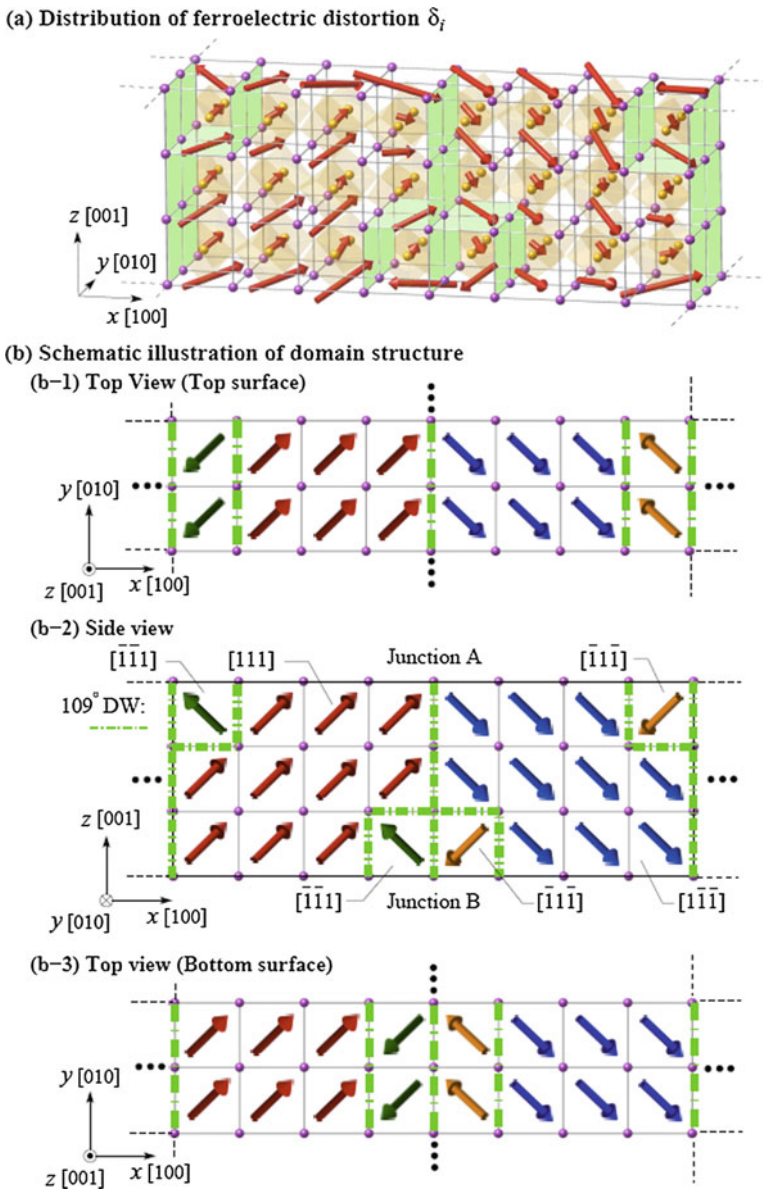

Fig. 7.11 **a** Distribution of ferroelectric distortion in the three-unit-cell thick $BiFeO_3$ ultrathin film with favorable domain period of eight unit cells. **b** Schematic illustration of domain structure in the film. *Arrows* indicate the ferroelectric distortions [25]

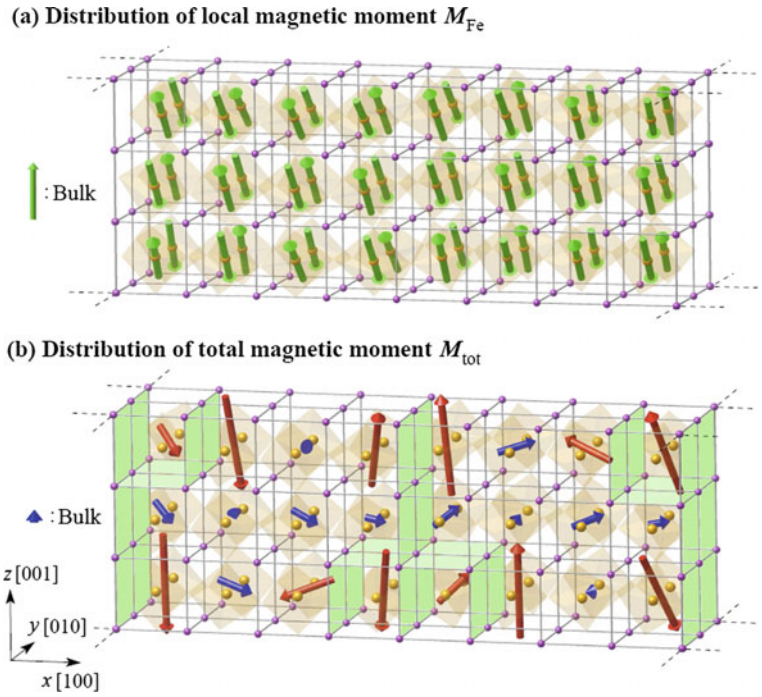

Fig. 7.12 **a** Distribution of local magnetic moment, M_{Fe}, of each Fe atom and **b** total magnetic moment M_{tot} over the two adjacent Fe atoms along the y axis in the three-unit-cell thick $BiFeO_3$ ultrathin film with favorable domain period of eight unit cells. *Green* and *blue arrows* indicate M_{Fe} and M_{tot}, respectively (M_{tot} of which magnitude is more than triple compared to that in the bulk is shown as *red arrows*) [25]

in the bulk [15, 26]. The (110) surfaces without any DWs notably induce the change of the direction of M_{tot}, but its magnitude is not so much different [25]. These facts suggest that the coupling of outer geometry (surfaces) and under-structure (DWs) synergistically affects the magnetic order in the $BiFeO_3$ ultrathin films.

7.2.2 *Nanowires: One-Dimensional Structure*

Multiferroic nanowires have drawn much attention as a one-dimensional nanostructure. Nanowires with perovskite structure usually possess a sharp edge where two surfaces with different orientation meet with each other at the end [27]. Due to the further change in coordination number at $BiFeO_3$ edges and low-dimensionality, both atomistic structure and electronic structure are expected to differ from those of the bulk or surface, leading to unique multiferroic properties. In recent years, multiferroic $BiFeO_3$ nanowires with a diameter of 8–10 nm have been fabricated

[28], and smaller nanowires with several lattice spacings are anticipated in the near future. Such ultrathin nanowires have the extremely high ratio of edges with respect to the entire volume. Thus, multiferroicity in the nanowires could be predominantly governed by the edge and surface interactions. Moreover, polarization which is normal to surfaces gives rise to surface charges and the resulting depolarizing field as shown in previous sections, which also leads to the disarrangement of the ferroic field. Since the nanowires are surrounded by surfaces with different orientation, such an effect can become more remarkable than that of the surface. Besides, the drastic change in magnetization in the nanowires can be accompanied by that in the polarization due to the interaction between the two ferroic orderings. Here, we review multiferroic properties in ultrathin $BiFeO_3$ nanowires with edge structure consisting of (100) and (010) surfaces.

Figure 7.13 shows the polarization distribution in the $BiFeO_3$ nanowire with a cross section of 4 × 4 unit cells [25]. Although the $BiFeO_3$ bulk exhibits the polarization along [111], the polarization in the nanowire lies along its surfaces, which results in the in-plane vortex-like structure in the circumferential (cross-sectional) direction (see Fig. 7.13b). To evaluate the strength of this polarization vortex, the toroidal moment $\boldsymbol{G}_{\mathrm{pol}}$ is introduced with a formulation of

$$\boldsymbol{G}_{\mathrm{pol}} = \frac{1}{V}\sum_i \boldsymbol{r}_i \times \boldsymbol{P}_i \Delta V_i, \tag{7.2}$$

where index i denotes each perovskite unit cell, $\boldsymbol{r}_i$ is the position vector from the central axis to the center of unit cell i, $\boldsymbol{P}_i$ is the local polarization, ΔV_i is the volume, and V is the total volume of the wire. $\boldsymbol{G}_{\mathrm{pol}}$ in the nanowire with a cross section of 4 × 4 unit cells is evaluated to be (0, 0, −0.234) e/Å. This also demonstrates that the nanowire has the large in-plane polarization vortex in the circumferential direction. On the other hand, the nanowire simultaneously possesses the axial polarization component as shown in Fig. 7.13c. Averaged polarization P_z as a representative of the axial polarization is written as

$$\overline{P_z} = \frac{1}{V}\sum_i P_{iz} \Delta V_i, \tag{7.3}$$

where P_{iz} is the z-component of polarization in i unit cell. P_z is 48.7 μC/cm^2 in the nanowire with a cross section of 4 × 4 unit cells, which reinforce the fact that there left a nontrivial axial polarization. Owing to the superposition of the circumferential and the axial components of the polarization, the $BiFeO_3$ nanowire has the spiral polarization. Chirality of the spiral in this nanowire is left-handed. Such spiral structure of polarization cannot be obtained in nanowires of other ferroelectric perovskite oxides. For the conventional ferroelectric $PbTiO_3$ nanowires with the same (100) surfaces, for example, the polarization lies only along the axial direction [29]. Thus, the spiral structure of polarization can be referred to as the unique characteristic of multiferroic $BiFeO_3$ nanowires.

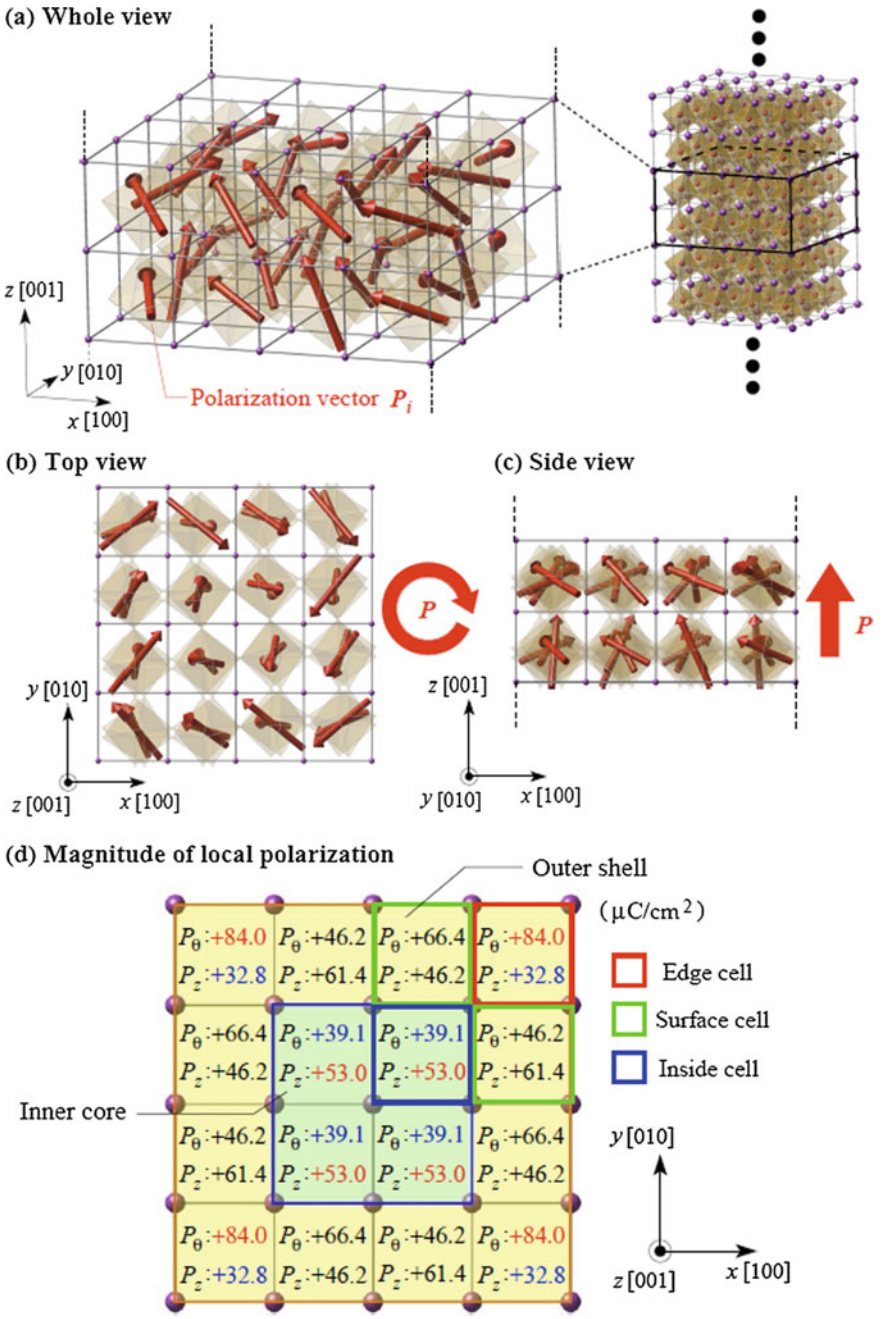

Fig. 7.13 **a**–**c** Polarization distribution in the $BiFeO_3$ nanowire with a cross section of the 4 × 4 unit cells. **d** Magnitude of the local polarization. The parts which are filled with *blue* and *yellow* indicate the inner core and the outer shell of the nanowire, respectively. For comparison, the spontaneous polarization in bulk $BiFeO_3$ is $P = 92.8\ \mu C/cm^2$ [25]

Coupled with the formation of spiral polarization in ferroelectric phase of the multiferroic nanowires, magnetism also exhibits a spiral character. Figure 7.14 shows the magnetic moment distribution in the $BiFeO_3$ nanowire with a cross section of 4 × 4 unit cells [25]. As shown in the Fig. 7.14c, the nanowire exhibits

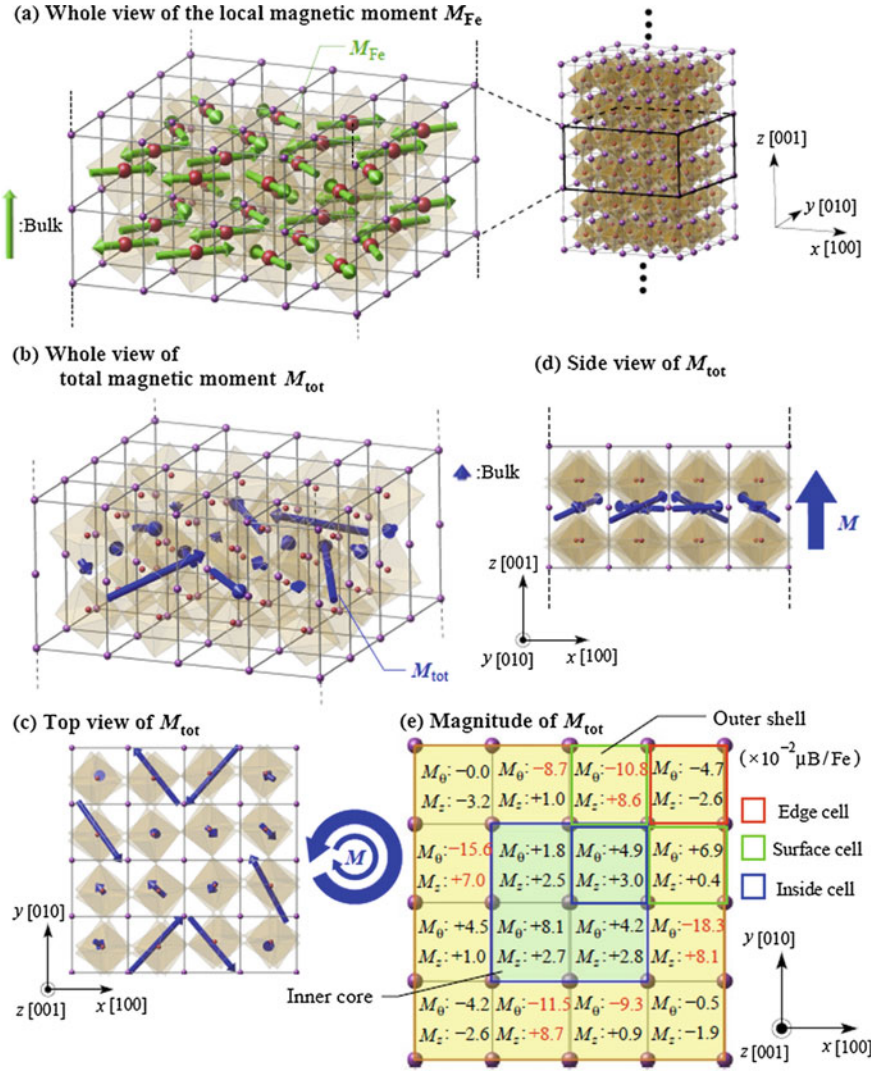

Fig. 7.14 Distribution of **a** local magnetic moment M_{Fe} of each Fe atom, **b–d** total magnetic moment M_{tot} of two adjacent Fe atoms along the z axis, and **e** magnitude of M_{tot} in the $BiFeO_3$ nanowire with a cross section of the 4 × 4 unit cells. M_θ and M_z indicate the components of M_{tot} in the circumferential and the axial directions, respectively. For comparison, the total magnetic moment in bulk $BiFeO_3$ is $M_{tot} = 3.9 \times 10^{-2}$ μ_B/Fe [25]

the complex magnetic ordering, but it creates the in-plane vortex-like structure as a whole. In the same way as polarization vortices, the toroidal moment $\boldsymbol{G}_{\mathrm{mag}}$ can be used to quantitatively evaluate the intensity of magnetic vortices as

$$\boldsymbol{G}_{\mathrm{mag}} = \frac{1}{V}\sum_i \boldsymbol{r}_i \times \boldsymbol{M}_i \Delta V_i, \tag{7.4}$$

where $\boldsymbol{M}_i$ is the local magnetic moment in i unit cell. $\boldsymbol{G}_{\mathrm{mag}}$ is evaluated to be (0, 0, +0.26) μ_B/Å. This further convinces that the nontrivial in-plane magnetization vortex exists in the nanowire. On the other hand, the nanowire simultaneously exhibits the axial magnetic moment as in the Fig. 7.14d. The average axial component M_z, which is formulated by

$$\overline{M_z} = \frac{1}{V}\sum_i M_{iz} \Delta V_i, \tag{7.5}$$

is +0.023 μ_B that is comparable to the in-plane component of the magnetic vortex. Again, the superposition of the in-plane and the axial components of the magnetic moment results in the characteristic spiral magnetic structure, which is totally different from the bulk where the magnetic moment is uniform in the [-1-12] direction. Chirality of the magnetic spiral is a right-handed system, which is opposite from that of the polarization. As a whole, the $BiFeO_3$ nanowires exhibit "double helix" of multiferroic order parameters with opposite chirality, like DNA. Such double helix multiferroic polarization and moment may lead to unique applications.

Finally, it should be remarked that, as a sequence of reducing dimensionality, nanodots are considered as a zero-dimensional nanostructure and are expected to have more remarkable effect of low-dimensionality. Studies on multiferroic nanodots have, however, not done yet and remain an exciting and promising future work.

7.3 Extrinsic (Defect-Induced) Multiferroics in Atomic Scales

As reviewed above, low-dimensional nanostructures provide unique and exotic multiferroic properties including closure or vortex domains. At the same time, increasing effect of depolarization or demagnetization field with shrinking the size destabilizes the ferroic order in the materials, leading to a critical size where ferroic order disappears [30]: In a thin-film geometry, for example, the critical thickness where the system loses ferroic instability was reported to be about 2 nm [31]. For nanodots, the critical size was also reported to be 3–5 nm [32, 33]. These critical sizes are inevitable so that further miniaturization and ultrahigh-density integration

of multiferroic components in devices are physically impossible by the direct scale of components.

In very recent years, however, a novel concept has been proposed to design atomic-scale multiferroics, clearly beyond the physical size limit of intrinsic multiferroics, by engineering lattice defects in nonmagnetic ferroelectrics like $PbTiO_3$ [34–36]. Hereafter, we review cutting-edge studies on extrinsic (defect-induced) multiferroics with an atomic dimension.

7.3.1 Multiferroic Grain Boundaries with Oxygen Vacancies

First preliminary work was motivated by an experimental observation that polycrystalline $PbTiO_3$ with nanoscale grains showed very weak but dilute ferromagnetism [37–39]. To understand the detailed origin and underlying mechanism of this unexpected magnetism in nonmagnetic ferroelectrics, vacancies coupled with grain boundaries are modeled using an ab initio calculations based on hybrid Hartree–Fock density-functional theories [40].

The magnetic moments distribution around oxygen vacancies at Σ5 grain boundaries (GBs) in $PbTiO_3$ is visualized in Fig. 7.15 [34]. Two inequivalent oxygen vacancies due to tetragonal lattice of $PbTiO_3$ are distinguished and denoted as V_{O1} and V_{O2}. In general, when oxygen vacancies are formed inside grains or bulk, no spin moment emerges regardless of their charge states, and the corresponding system keeps nonmagnetic as host (defect-free) $PbTiO_3$ [41]. On the other hand, a nonzero and nontrivial magnetic moment emerges when V_{O1} is formed at GB (see GB-V_{O1}; 2.0 μ_B). This indicates that the defect coupling of oxygen vacancies and GBs surely causes magnetism into nonmagnetic $PbTiO_3$. The emerged magnetization is highly localized at the Ti atom neighboring to the vacancy site (see Fig. 7.15a), which formed a strong covalent bond with O1 atom in the perfect (defect-free) Σ5 GB via the orbital hybridization between Ti 3*d* and O 2*p*. This suggests that the emergent magnetization arises from the unbound Ti 3*d* electrons due to the formation of an oxygen vacancy. In a similar way, another oxygen vacancy of V_{O2} located on the GB plane exhibits nonzero magnetization, as well, which are, however, separated and localized at the two Ti atoms neighboring to the vacancy site (see Fig. 7.15b). These magnetic moments at the paired Ti atoms are positive and negative each, meaning a locally antiferromagnetic configuration (i.e., the net magnetic moment is zero). In this way, the defect interplay between oxygen vacancies and GBs indeed induces localized magnetizations with a various magnetic ordering such as ferromagnetic and antiferromagnetic depending on the site of GBs.

This phenomenon tells us how to design atomic-scale multiferroics. Grain boundaries are well known to be the trapping site of oxygen vacancies, so that the high concentration of vacancies is expected at the GB plane. In fact, an experimental observation showed that more than 30% of oxygen sites are vacant in the

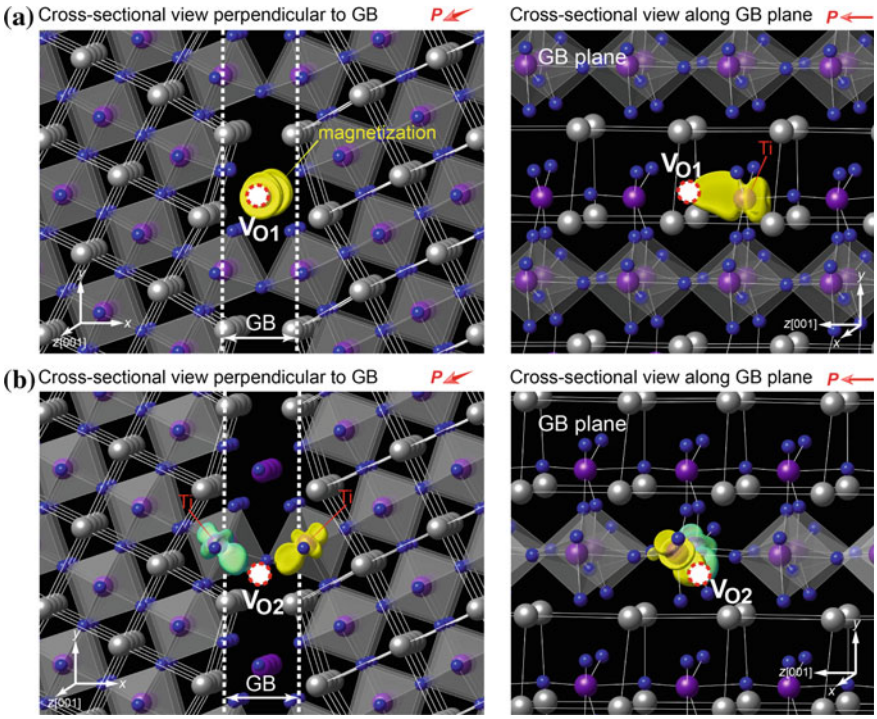

Fig. 7.15 Spatial magnetization distribution near the **a** GB-V_{O1} and **b** GB-V_{O2}. The views from perpendicular and parallel direction to the grain boundary (GB) are visualized in the *left* and *right*, respectively. Isosurfaces of +0.04 and −0.04 $\mu_B/Å^3$ are visualized by *yellow* and *green* colors, respectively. *P* denotes the direction of the spontaneous ferroelectric polarization. Reprinted with the permission from Ref. [34]. Copyright 2015 American Chemical Society

perovskite oxides [42, 43]. This means that GBs covered by oxygen vacancies, which provides magnetizations at their sites, are almost entirely spin polarization and magnetized. Emergence magnetism is arranged in a planar way along the GB plane coupled with ferroelectricity of host $PbTiO_3$, allowing an O-deficient GB to behave as a multiferroic monolayer. Clearly, such single atomic layer multiferroics made by GBs are beyond the critical thickness of ferroics [31–33]. Therefore, this suggests us the possibility of atomic-scale multiferroics via engineering lattice defects.

This defect-induced magnetism can be understood from electronic and orbital points of view. Figure 7.16a, c present the spin-polarized density of states (DOS) for the V_{O1} inside the grain (Bulk-V_{O1}) and at the grain boundary (GB-V_{O1}), respectively [34]. For the case of Bulk-V_{O1}, there is only one electronic defect state between the valence band maximum (VBM) and conduction band minimum (CBM) (Fig. 7.16a), which is absent in defect-free $PbTiO_3$. This state is intrinsic to the formation of oxygen vacancy (so-called defect state). The defect state is located at 2.56 eV above the VBM and unspin-polarized, which leads to simply

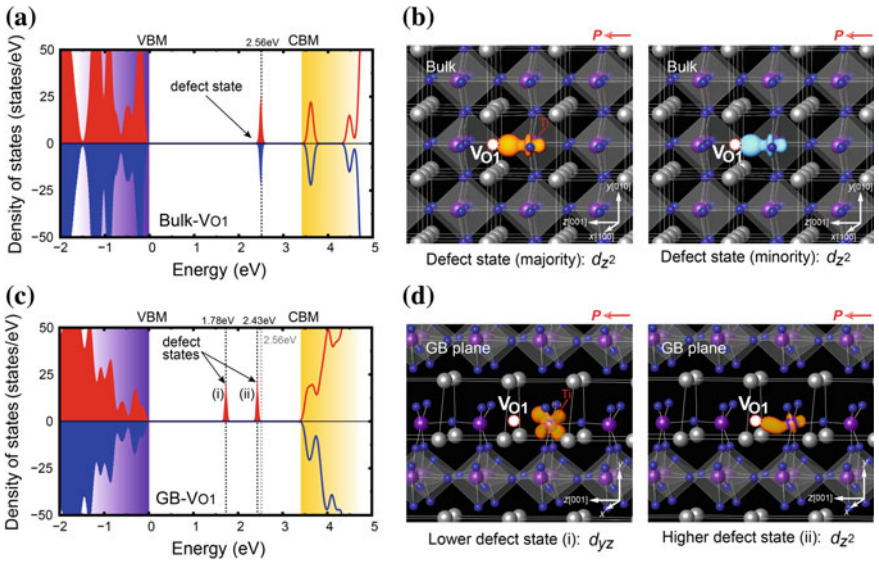

Fig. 7.16 Spin-polarized density of states (DOS) for **a** Bulk-V_{O1} and **c** GB-V_{O1} in left panels. The *right panels* show the squared wave functions of vacancy states in **b** Bulk-V_{O1} and **d** GB-V_{O1}. The *orange* and *blue* isosurfaces show the majority-spin and minority-spin densities, respectively. Reprinted with the permission from Ref. [34]. Copyright 2015 American Chemical Society

nonmagnetic nature of oxygen vacancies in grains or bulk. The spatial distribution of this defect state is shown in Fig. 7.16b. The defect state is distributed between the Bulk-V_{O1} and the neighboring Ti atom and is contributed by a d_{z2} orbital due to a loss of the Ti d—O1 p hybridized bond.

In contrast to the character of Bulk-V_{O1}, the oxygen vacancy at the GB (GB-V_{O1}) possesses two split defect states (i) and (ii) (Fig. 7.16c). The higher level state (ii) is found at 2.43 eV above the VBM (almost the same position as that of Bulk-V_{O1}), whereas the lower state (i) can be found at 1.78 eV. Each defect state is partially occupied by a single majority (up) spin, leading to the spin polarization of the defect states. This spin polarization gives GB-V_{O1} an emerging magnetic moment. The higher state (ii) is d_{z2} contribution, i.e., same as that in bulk, while the lower defect state (i) is a unique d_{yz} orbital (Fig. 7.16d). The appearance of this lower d_{yz} state splits the fully occupied higher d_{z2} state of Bulk-V_{O1} into the two partially occupied defect states of GB-V_{O1}, which leads to the spin polarization at the GB. This new defect state is due to the symmetry breaking at the GB. The GBs break the orbital symmetry of the original d_{z2}-dominant defect state and allow the additional d_{yz} contribution. Since such structural change and the resulting orbital symmetry break at GBs are common among GBs, a similar defect-induced ferromagnetism can be expected in other oxygen-deficient GBs.

It should be remarked that such clear and explaining picture of defect electronic structure cannot be obtained from normal density-functional theory calculations

based on standard exchange–correlation functional such as local density approximation (LDA) or generalized gradient approximation (GGA). LDA or GGA provides a better structural prediction but severely underestimated band gap, which hides the defect state into conduction bands and calculations capture the conduction band edge instead of correct defect states [44, 45]. Due to this inaccuracy, vacancies incorrectly show a delocalized state and no net magnetization. On the other hand, hybrid Hartree–Fock density-functional theory used here provides an accurate band structure with reproducing an experimental band gap, successfully providing a correct picture of defect electronic states [34]. Therefore, the use of hybrid functional method is essential for studies on lattice defects in ferroelectrics or large band gap insulators.

7.3.2 *Multiferroic Vacancies at Ferroelectric Oxide Surfaces*

The previous section shows the possibility of designing atomic-scale multiferroics through engineering lattice defects. Surface structures also provide the similar platform to design multiferroics in an atomic level.

Figure 7.17 shows various vacancies at (001) $PbTiO_3$ surfaces that give rise to magnetizations [35]. Oxygen vacancies are the most abundant and common point defects in oxide materials, and they also provide a nonzero and nontrivial magnetization when they are formed at TiO_2-terminated surface, which brings about magnetism to the intrinsically nonmagnetic $PbTiO_3$, as similarly observed at GBs. The magnetic ordering is, however, different depending on a type of oxygen vacancies. For V_{O1}, the emerging positive magnetization (i.e., ferromagnetism) is localized at the vacancy site and at the neighboring Ti atom (Fig. 7.17b) due to the dangling bond formation by the vacancy, which breaks hybridization of the Ti-3*d* and O1-2*p* orbitals. On the other hand, another oxygen-vacancy V_{O2} at the TiO_2-terminated surface provides two split magnetizations localized at the two neighboring Ti atoms (Fig. 7.17c). These localized magnetizations are positive and negative each, i.e., an antiferromagnetic configuration due to a superexchange interaction of spins via indirect *d*–*d* hopping between the two equivalent Ti ions paired by the oxygen vacancy [46]. The mechanism of this emergence of magnetization can be explained in the same manner as observed at GBs. As shown in Fig. 7.18, the single defect state intrinsic to an oxygen vacancy in a crystal can be split into two distinct states when it is formed at the surface [35]. This is due to the surface symmetry breaking by the absence of atoms in a direction perpendicular to the surface plane. Again, the key to induce the magnetization is symmetry breaking of defect-state orbitals.

For cation vacancies, a similar phenomenon is observed; while the internal V_{Pb} is nonmagnetic, V_{Pb} on the PbO-terminated surface provides antiferromagnetic magnetic moments paired at two O2 atoms neighboring to the vacancy site (Fig. 7.17d). Although V_{Ti} is hardly observed experimentally, at most in very low

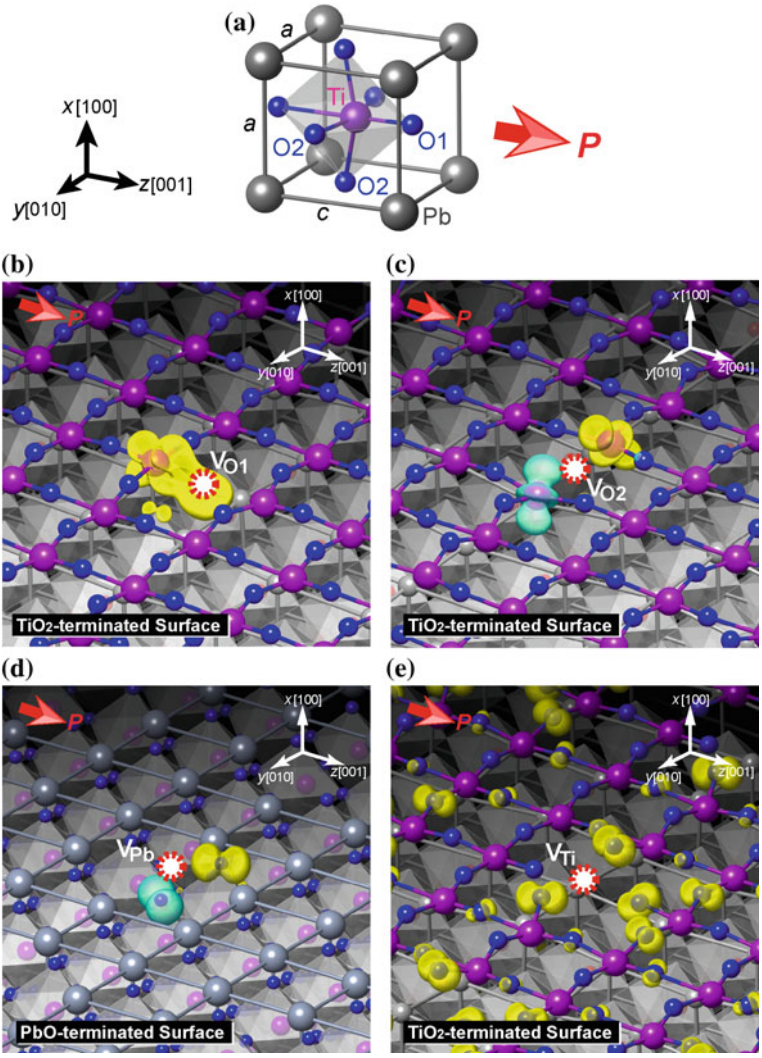

Fig. 7.17 **a** Atomic configuration of $PbTiO_3$. *P* is the spontaneous polarization. Magnetization distribution near **b** V_{O1} and **c** V_{O2} at the TiO_2-terminated surfaces, **d** V_{Pb} at the PbO-terminated surface, and **e** V_{Ti} at the TiO_2-terminated surface. The isosurfaces of +0.035 and $-0.035\mu_B/Å^3$ are shown by the *yellow* and *green* colors, respectively. Reprinted with permission from Ref. [35]. Copyright 2015 by American Physical Society

concentrations, V_{Ti} gives a net magnetic moment of 4.0 μ_B and ferromagnetism on the TiO_2-terminated surface; however, these magnetizations are delocalized away from the vacancy site (Fig. 7.17e). In this way, a surface provides an interesting platform to give various characters and orders of magnetic moments to atomic vacancies.

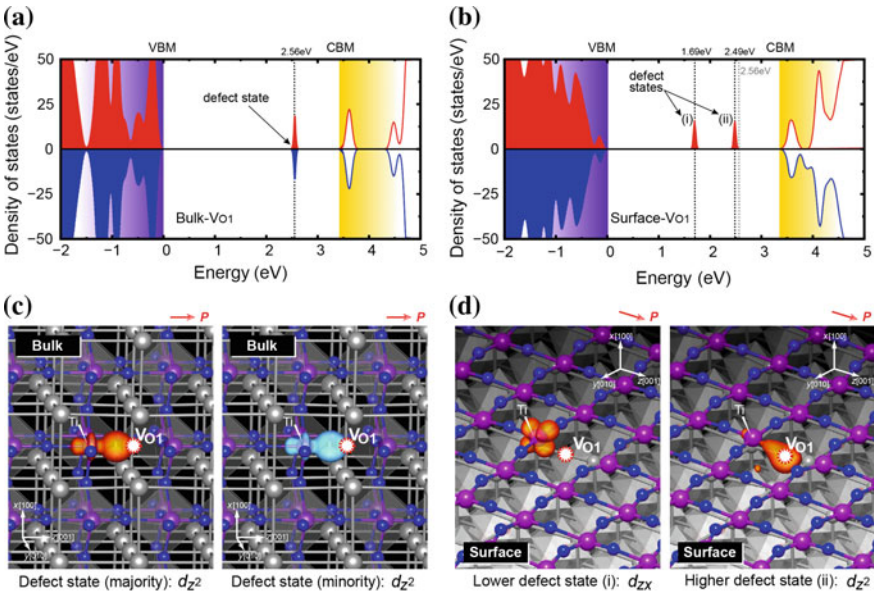

Fig. 7.18 Spin-polarized density of states (DOS) of V_{O1} **a** inside bulk and **c** at the surface. The spatial distribution of the defect state of V_{O1} **b** inside bulk and **d** at the surface. Reprinted with permission from Ref. [35]. Copyright 2015 by American Physical Society

Above, you can find the most of defect-induced magnetization highly localized at the vacancy site, meaning the atomic-scale magnetization. Considering that these localized magnetizations are coupled with host ferroelectricity of $PbTiO_3$, atomic vacancies at surfaces can be regarded as a "monoatomic" multiferroics, clearly beyond the critical size limit of nanodots [31–33]. It is also noted that concentration of each vacancy can be tuned by environment or growth conditions such as temperature or pressure [47, 48]. This indicates that appropriate choice of fabrication conditions allows us to control such defect-induced magnetic properties and resulting atomic-scale multiferroicity.

Another important feature is coupling of ferroelectricity and emerged magnetism, i.e., magnetoelectric coupling. Figure 7.19 shows the magnetic moment distribution of V_{O1} at the TiO_2-terminated surface depending on various spontaneous polarization configurations. Initially, the localized magnetization on the single adjacent Ti is ferromagnetic in the case of the polarization along the [001] direction (Fig. 7.19a) [35]. When the polarization is rotated to the another [010] polar axis, the initial ferromagnetic state changes into the antiferromagnetic state with both the negative and positive spin moments localized at each of the two neighboring Ti atoms (Fig. 7.19b). On the other hand, when the polarization is switched to a direction perpendicular to the surface layer (Fig. 7.19c), the vacancy transformed into a simply nonmagnetic state. As a whole, an oxygen vacancy at the

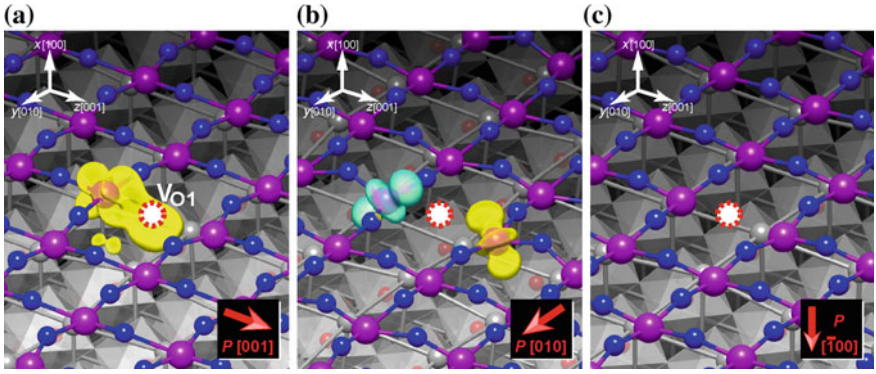

Fig. 7.19 Magnetic configurations of oxygen-vacancy V_{O1} at the TiO_2-terminated surface in various ferroelectric polarization directions. Ferromagnetic, antiferromagnetic, and nonmagnetic phase transitions with the polarization of the in-plane **a** [001] and **b** [010], and **c** out-of-plane directions, respectively. Reprinted with permission from Ref. [35]. Copyright 2015 by American Physical Society

surface undergoes the ferromagnetic, antiferromagnetic, and nonmagnetic phase transitions with a polarization switching, i.e., the existence of a nonlinear magnetoelectric response in this multiferroic oxygen vacancy. It is also noted that in the paraelectric phase, there is no magnetization and the system is simply nonmagnetic, indicating that not only surface–defect interaction but also the host ferroelectricity are both necessary to induce the magnetoelectric multiferroicity in atomic scale.

7.3.3 Strain-Induced Multiferroic Transitions

As reviewed above, a key to extrinsic multiferroicity with defect-induced magnetism should be an orbital symmetry breaking of defect states, leading to the split of state and spin polarization. Considering the fact that lattice symmetry in crystals can be distorted by mechanical strain, e.g., phase transformations triggered by pressure or epitaxial strains. In terms of multiphysics point of view, there is a possibility that applied strain can be also a trigger of defect-induced magnetism and multiferroicity.

Such multiphysics behaviors have been demonstrated in epitaxial (110) $PbTiO_3$ with oxygen vacancies. Figure 7.20a shows a magnetic phase diagram of three inequivalent (or independent) oxygen vacancies V_{O1}–V_{O3} in epitaxially strained (110) $PbTiO_3$ films [49]. Initially, oxygen vacancies have no magnetization, and the system remains purely ferroelectric (FE) at zero misfit strain, owing to the unspin-polarized nature of all the oxygen vacancies. On the other hand, when the misfit strain is applied above 0.5%, some or all of oxygen vacancies begin to be magnetized with ferromagnetic or antiferromagnetic orders, and the epitaxially

strained system undergoes a multiferroic phase transition. Similar phenomena can be also observed under relatively high compression (below 1.8%). In addition, these emerging multiferroic phases exhibit magnetoelectric effects as well. Figure 7.20b presents the magnetic phase transitions of an oxygen-vacancy V_{O2} at a strain of $\varepsilon = 3.0\%$ before and after polarization switching from [100] to [010] [49]. Via the switching, the vacancy V_{O2} exhibits a rapid change of magnetic moment and its ordering from ferromagnetic and antiferromagnetic configurations. This means the defect-mediated magnetism can be electrically controlled by applying electric field, and as a consequence, the oxygen vacancy shows the ME effect in epitaxial (110) $PbTiO_3$. In a similar way, a magnetic phase transition from ferromagnetism to nonmagnetism can be possible under compression (below 1.8%), and the coincident ME coupling is also present. These suggest another strategy for design of multiferroics by engineering defect–strain interaction. In fact, the epitaxial strain considered here (−4 to 4%) is practically feasible by the use of common perovskite

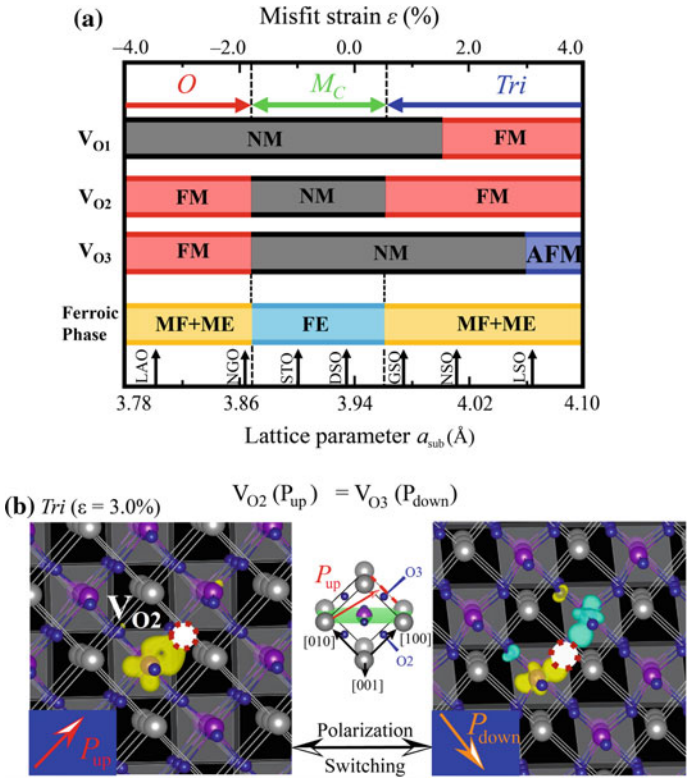

Fig. 7.20 **a** Multiferroic and magnetoelectric phase diagram in epitaxially strained (110) $PbTiO_3$. The lattice parameters for some candidate substrates, viz., $LaAlO_3$ (LAO), $NdGaO_3$ (NGO), $SrTiO_3$ (STO), $DyScO_3$ (DSO), $GdScO_3$ (GSO), $NdScO_3$ (NSO), and $LaScO_3$ (LSO), are also shown. **b** Magnetic phase transition of V_{O2} via polarization switching in the triclinic phase. Reprinted with permission from Ref. [49]. Copyright 2015 by American Physical Society

substrates, such as $LaAlO_3$, $NdGaO_3$, $SrTiO_3$, $NdScO_3$, and $LaScO_3$ [50, 51], as also shown in Fig. 7.20a.

In addition, the mobile nature of oxygen vacancies can provide a control of MF and ME properties in terms of their concentration and distribution [52–56]. This concept makes it possible to spatially modulate the multiferroic properties via oxygen-vacancy engineering; e.g., the migration and diffusion of oxygen vacancies with an external electric field [54] makes it possible to design functionally graded multiferroics by concentration gradient of vacancies.

7.4 Conclusion

In this chapter, we reviewed a series of studies on intrinsic multiferroics such as $BiFeO_3$ from bulk to nanostructures. In Sect. 7.1, we first introduced multiferroicity and their coupling (i.e., magnetoelectric effect) in $BiFeO_3$ and their sensitive response to strain through ferroelectric and magnetic interactions. We also briefly reviewed first-principles studies on the effect of understructures of multiferroic domain morphology and lattice defects.

After that, in Sect. 7.2, we discussed the multiferroic and magnetoelectric properties in low-dimensional nanostructures. By the reduction of dimensionality, ferroelectric polarization and magnetic moment form an interesting and unusual domain configuration such as closure, vortex, and spiral forms through interaction of geometry and multiferroicity.

In Sect. 7.3, extrinsic multiferroics with atomic scale were discussed by engineering lattice defects. Beyond the physical size limit of ferroic orders, interactions of lattice defects and/or strains successfully provide a platform to fabricate an atomic-scale multiferroics. Such a new material design concept will provide a promising avenue for multifunctional oxides at atomic scales.

References

1. N.A. Hill, J. Phys. Chem. B **104**, 6694 (2000)
2. M. Fiebig, Th Lottermoser, D. Fröhlich, A.V. Goltsev, R.V. Pisarev, Nature (London) **419**, 818 (2002)
3. J. Wang, J.B. Neaton, H. Zheng, V. Nagarajan, S.B. Ogale, B. Liu, D. Viehland, V. Vaithyanathan, D.G. Schlom, U.V. Waghmare, N.A. Spaldin, K.M. Rabe, M. Wuttig, R. Ramesh, Science **299**, 1719 (2003)
4. T. Kimura, T. Goto, H. Shintani, K. Ishizaka, T. Arima, Y. Tokura, Nature (London) **426**, 55 (2003)
5. H. Ohno, Science **281**, 951 (1998)
6. See M. Fiebig, J. Phys. D Appl. Phys. **38**, R123 (2005)
7. R. Ramesh, *Thin Film Ferroelectric Materials and Devices* (Kluwer Academic, Boston, 1997)
8. J.F. Scott, *Ferroelectric Memories* (Springer, Berlin, 2000)

9. N.A. Spaldin, M. Fiebig, Science **309**, 391 (2005)
10. J.F. Scott, Nat. Mater. **6**, 256–257 (2007)
11. A. Pimenov, A.A. Mukhin, V.Y. Ivanov, V.D. Travkin, A.M. Balbashov, A. Loidl, Nat. Phys. **2**, 97 (2006)
12. T. Shimada, K. Arisue, J. Wang, T. Kitamura, Phys. Rev. B **89**, 245437 (2014)
13. C. Ederer, N.A. Spaldin, Phys. Rev. B **71**, 060401(R) (2005)
14. K. Arisue, T. Shimada, T. Kitamura, J. Soc. Mater. Sci. Jpn. **63**, 168 (2014)
15. A. Lubk, S. Gemming, N.A. Spaldin, Phys. Rev. B **80**, 104110 (2009)
16. T. Shimada, Y. Umeno, T. Kitamura, Phys. Rev. B **77**, 094105 (2008)
17. B.J. Rodriguez, Y.H. Chu, R. Ramesh, S.V. Kalinin, Appl. Phys. Lett. **93**, 142901 (2008)
18. M.-W. Chu, I. Szafraniak, R. Scholz, C. Harnagea, D. Hesse, M. Alexe, U. Gsele, Nat. Mater. **3**, 87 (2004)
19. D. Kan, T. Terashima, R. Kanda, A. Masuno, K. Tanaka, S. Chu, H. Kan, A. Ishizumi, Y. Kanemitsu, Y. Shimakawa, M. Takano, Nat. Mater. **4**, 816 (2005)
20. S.R. Basu, L.W. Martin, Y.H. Chu, M. Gajek, R. Ramesh, R.C. Rai, X. Xu, J.L. Musfeldt, Appl. Phys. Lett. **92**, 091905 (2008)
21. Z. Zhang, P. Wu, L. Chen, J. Wang, Appl. Phys. Lett. **96**, 232906 (2010)
22. T. Shimada, T. Matsui, T. Xu, K. Arisue, Y. Zhang, J. Wang, T. Kitamura, Phys. Rev. B **93**, 174107 (2016)
23. W. Ratcliff II, D. Kan, W. Chen, S. Watson, S. Chi, R. Erwin, G.J. Mclntyre, S.C. Capelli, I. Takeuchi, Adv. Funct. Mater. **21**, 1567 (2011)
24. S.K. Lee, B.H. Choi, D. Hesse, Appl. Phys. Lett. **102**, 242906 (2013)
25. K. Arisue, Master's Dissertation, Kyoto University (unpublished)
26. C.T. Nelson, B. Winchester, Y. Zhang, S.J. Kim, A. Melville, C. Adamo, C.M. Folkman, S.H. Baek, C.B. Eom, D.G. Schlom, L.Q. Chen, X. Pan, Nano Lett. **11**, 828 (2011)
27. Y. Yamashita, K. Mukai, J. Yoshinobu, M. Lippmaa, T. Kinoshita, M. Kawasaki, Surf. Sci. **514**, 54 (2002)
28. Y. Zhao, J. Miao, X. Zhang, Y. Chen, X.G. Xu, Y. Jiang, J. Mater. Sci.: Mater. Electron. **23**, 180 (2012)
29. G. Pilania, R. Ramprasad, Phys. Rev. B **82**, 155442 (2010)
30. R. Kretschmer, K. Binder, Phys. Rev. B **20**, 1065 (1979)
31. L. Despont, C. Koitzsch, F. Clerc, M.G. Garnier, P. Aebi, C. Lichtensteiger, J.-M. Triscone, F.J. Garcia de Abajo, E. Bousquet, P. Ghosez, Phys. Rev. B **73**, 094110 (2006)
32. I.I. Naumov, L. Bellaiche, H. Fu, Nature (London) **432**, 737 (2004)
33. M.J. Polking, M.-G. Han, A. Yourdkhani, V. Petkov, C.F. Kisielowski, V.V. Volkov, Y. Zhu, G. Caruntu, A.P. Alivisatos, R. Ramesh, Nat. Mater. **11**, 700 (2012)
34. T. Shimada, J. Wang, T. Ueda, Y. Uratani, K. Arisue, M. Mrovec, C. Elsässer, T. Kitamura, Nano Lett. **15**, 27 (2015)
35. T. Shimada, J. Wang, Y. Araki, M. Mrovec, C. Elsässer, T. Kitamura, Phys. Rev. Lett. **115**, 107202 (2015)
36. T. Xu, T. Shimada, Y. Araki, J. Wang, T. Kitamura, Nano Lett. **16**, 454 (2016)
37. R.V.K. Mangalam, N. Ray, U.V. Waghmare, A. Sundaresan, C.N.R. Rao, Solid State Commun. **149**, 1–5 (2009)
38. M. Wang, G.-L. Tan, Q. Zhang, J. Am. Ceram. Soc. **93**, 2151–2154 (2010)
39. Z. Zhang, J. Hu, Z. Xu, H. Qin, L. Sun, F. Gao, Y. Zhang, M. Jiang, Solid State Sci. **13**, 1391–1395 (2011)
40. J. Heyd, G.E. Scuseria, M. Ernzerhof, J. Chem. Phys. **124**, 219906 (2006)
41. T. Shimada, T. Ueda, J. Wang, T. Kitamura, Phys. Rev. B **87**, 174111 (2013)
42. C.L. Jia, K. Urban, Science **303**, 2001–2004 (2004)
43. P. Prabhumirashi, V.P. Dravid, A.R. Lupini, M.F. Chisholm, S.J. Pennycook, Appl. Phys. Lett. **87**, 121917 (2005)
44. T. Shimada, Y. Uratani, T. Kitamura, Appl. Phys. Lett. **100**, 162901 (2012)
45. T. Shimada, Y. Uratani, T. Kitamura, Acta Mater. **60**, 6322–6330 (2012)
46. K. Yang, Y. Dai, B. Huang, Y.P. Feng, Phys. Rev. B **81**, 033202 (2010)

47. J.H. Cravford, L.M. Slifkin, *Point Defects in Solids* (Plenum, New York, 1972)
48. F. Tuomisto, K. Saarinen, D.C. Look, G.C. Farlow, Phys. Rev. B **72**, 085206 (2005)
49. T. Xu, T. Shimada, Y. Araki, J. Wang, T. Kitamura, Phys. Rev. B **92**, 104106 (2015)
50. D.G. Schlom et al., Annu. Rev. Mater. Res. **37**, 589 (2007)
51. N. Setter, D. Damjanovic, L. Eng, G. Fox, S. Gevorgian, S. Hong, A. Kingon, H. Kohlstedt, N.Y. Park, G.B. Stephenson, I. Stolitchnov, A.K. Taganstev, D.V. Taylor, T. Yamada, S. Streiffer, J. Appl. Phys. **100**, 051606 (2006)
52. H.L. Tuller, S.R. Bishop, Annu. Rev. Mater. Res. **41**, 369 (2011)
53. G.Y. Yang, G.D. Lian, E.C. Dickey, C.A. Randall, D.E. Barber, P. Pinceloup, M.A. Henderson, R.A. Hill, J.J. Beeson, D.J. Skamser, J. Appl. Phys. **96**, 7500 (2004)
54. S.J. Park et al., Naotechnology **24**, 295202 (2013)
55. G.L. Yuan, L.W. Martin, R. Ramesh, A. Uedono, Appl. Phys. Lett. **95**, 012904 (2009)
56. W.L. Warren, K. Vanheusden, D. Dimos, G.E. Pike, B.A. Tuttle, J. Am. Ceram. Soc. **79**, 536 (1996)

Chapter 8
Ferroic Nanometamaterials and Composites

Abstract Metamaterials, which possess artificially designed lattice-shaped internal structures, gain remarkably increasing attention owing to their unique, controllable, and often unprecedented properties. We review a series of pioneering studies on ferroelectric nanometamaterials and nanocomposites. After introducing the methodology of the phase-field model, we discuss domain configurations in nanometamaterials. Then, recent investigations of multiferroic (ferroelectric/ferromagnetic) nanocomposites are reviewed.

Keywords Metamaterial · Phase-field model · Ferroic · Ferroelectrics · Nanocomposite

8.1 Ferroic Nanometamaterials from Phase-Field Modeling

Metamaterials, which include an artificially designed lattice-shaped internal structure inside, are rapidly appearing at the cutting edge and frontier of current science and engineering due to their unique, controllable, and often unprecedented properties, which originate from periodic arrangements of rationally designed structures rather than from the intrinsic properties of the host constituents. From the beginning to realize negative refraction just over a decade ago [1, 2], the metamaterial concept rapidly covers a broader field of not only electromagnetics or optics [3–8] but also other properties such as mechanical and electromechanical properties [9–13]. In this way, the metamaterial concept is still expanding and broadening its coverage of fields, and a new class of metamaterials is further expected to appear.

As shown in previous chapters, ferroic properties are quite sensitive to dimension and geometry of materials and forms a unique polar configurations at the nanoscale dimensions such as polarization vortices. Considering the fact that the nanometamaterials can be regarded as a combination of nanoscale components with various shapes (i.e., accommodation of a hierarchical nanostructure), ferroelectric nanometamaterials provide a promising platform to design a plethora of unique

Y. Umeno et al., *Multiphysics in Nanostructures*, Nanostructure Science and Technology, DOI 10.1007/978-4-431-56573-4_8

polar configurations with complex domain patterns due to the dual advantages of variable length-scale and dimensionality distinct from the counterpart of the bulk or simple nanocomponent. Therefore, the ferroic metamaterials are promising for the advance of nanoscale ferroelectricity or multiferroicity.

Along the stream, very recently, a first attempt has been made to extend the metamaterial concept to ferroelectrics or multiferroics fields [14–16]. To capture their hierarchical polar and magnetic configurations in different length scales, not an ab initio approach but a phase-field method based on Landau free energy theory is used to study. In this chapter, we first introduce the methodology of phase-field modeling of ferroelectrics and multiferroics, and review the recent attempt on ferroic metamaterials.

8.1.1 Phase-Field Modeling of Ferroelectrics

The mechanical and ferroelectric behaviors in a ferroelectric system can be simulated using real-space phase-field modeling based on the Ginzburg–Landau theory [17, 18]. The phase-field modeling achieves self-consistency of the electrostatic and elastic interactions, which is essential to describe the electric and mechanical behavior near a crack tip. In the phase-field model of ferroelectric materials, the polarization vector $\boldsymbol{P} = (P_1, P_2, P_3)$ is taken as the order parameter to describe the free energies of ferroelectric systems. The total free energy of the ferroelectric system F can be described by

$$F = \int_V f \mathrm{d}V = \int_V \left(f_{\mathrm{Land}} + f_{\mathrm{grad}} + f_{\mathrm{elas}} + f_{\mathrm{coup}} + f_{\mathrm{elec}}\right)\mathrm{d}V, \tag{8.1}$$

where f, f_{Land}, f_{grad}, f_{elas}, f_{coup}, and f_{elec} denote the total free energy density, the Landau energy density, the gradient energy density, the elastic energy density, the coupling energy density, and the electrostatic energy density, respectively. V is the entire volume of the ferroelectric system.

The Landau energy density is expressed by a six-order polynomial of the spontaneous polarization [19] as

$$\begin{aligned} f_{\mathrm{Land}} = {} & \alpha_1\left(P_1^2 + P_2^2 + P_3^2\right) + \alpha_{11}\left(P_1^4 + P_2^4 + P_3^4\right) \\ & + \alpha_{12}\left(P_1^2P_2^2 + P_2^2P_3^2 + P_3^2P_1^2\right) + \alpha_{111}\left(P_1^6 + P_2^6 + P_3^6\right) \\ & + \alpha_{112}\left[P_1^4\left(P_2^2 + P_3^2\right) + P_2^4\left(P_1^2 + P_3^2\right) + P_3^4\left(P_1^2 + P_2^2\right)\right] \\ & + \alpha_{123}P_1^2P_2^2P_3^2 \end{aligned} \tag{8.2}$$

where α_1 is the dielectric stiffness, and α_{11}, α_{12}, α_{111}, α_{112}, and α_{123} are higher-order dielectric stiffness. The dielectric stiffness α_1 is given a linear temperature dependence based on the Curie–Weiss law:

$$\alpha_1 = (T - T_0)/2\kappa_0 C_0, \tag{8.3}$$

where T and T_0 denote the temperature and the Curie–Weiss temperature, respectively, C_0 denotes the Curie constant, and κ_0 is the dielectric constant of a vacuum. The gradient energy density [20, 21] in the second term of Eq. (8.1) is given by

$$\begin{aligned} f_{\text{grad}} = {} & \frac{1}{2}G_{11}\left(P_{1,1}^2 + P_{2,2}^2 + P_{3,3}^2\right) + G_{12}\left(P_{1,1}P_{2,2} + P_{2,2}P_{3,3} + P_{3,3}P_{1,1}\right) \\ & + \frac{1}{2}G_{44}\left[\left(P_{1,2} + P_{2,1}\right)^2 + \left(P_{2,3} + P_{3,2}\right)^2 + \left(P_{1,3} + P_{3,1}\right)^2\right] \\ & + \frac{1}{2}G'_{44}\left[\left(P_{1,2} - P_{2,1}\right)^2 + \left(P_{2,3} - P_{3,2}\right)^2 + \left(P_{3,1} - P_{1,3}\right)^2\right], \end{aligned} \tag{8.4}$$

where G_{11}, G_{12}, G_{44}, and G'_{44} are the gradient energy coefficients, and $P_{i,j} = \partial P_i/\partial x_j$ denotes the derivative of the ith component of the polarization vector P_i, with respect to the jth coordinate x_j, and $i, j = 1, 2, 3$. The gradient energy represents the energy penalty for inhomogeneous spatial distribution of the polarization in a ferroelectric system, such as domain walls. The elastic energy density is a pure mechanical strain energy:

$$\begin{aligned} f_{\text{elas}} = {} & \frac{1}{2}c_{11}\left(\varepsilon_{11}^2 + \varepsilon_{22}^2 + \varepsilon_{33}^2\right) + c_{12}\left(\varepsilon_{11}\varepsilon_{22} + \varepsilon_{22}\varepsilon_{33} + \varepsilon_{33}\varepsilon_{11}\right) \\ & + 2c_{44}\left(\varepsilon_{12}^2 + \varepsilon_{23}^2 + \varepsilon_{31}^2\right), \end{aligned} \tag{8.5}$$

where c_{11}, c_{12}, and c_{44} are the elastic constants, and ε_{ij} is the elastic strain. The fourth term of Eq. (8.1) represents the energy density of piezoelectric coupling between the spontaneous polarization and mechanical strain and is given by [20, 21]

$$\begin{aligned} f_{\text{coup}} = {} & -q_{11}\left(\varepsilon_{11}P_1^2 + \varepsilon_{22}P_2^2 + \varepsilon_{33}P_3^2\right) \\ & - 2q_{12}\left[\varepsilon_{11}\left(P_2^2 + P_3^2\right) + \varepsilon_{22}\left(P_1^2 + P_3^2\right) + \varepsilon_{33}\left(P_1^2 + P_2^2\right)\right] \\ & - 2q_{44}\left(\varepsilon_{12}P_1P_2 + \varepsilon_{13}P_1P_3 + \varepsilon_{23}P_2P_3\right), \end{aligned} \tag{8.6}$$

where q_{11}, q_{12}, and q_{44} are electrostrictive constants. The electrostatic energy density, which is obtained through Legendre transformation, is expressed as

$$f_{\text{elec}} = -\frac{1}{2}\kappa_c\left(E_1^2 + E_2^2 + E_3^2\right) - E_1P_1 - E_2P_2 - E_3P_3, \tag{8.7}$$

where E_i is the electric field due to the electrostatic potential distribution.

The temporal evolution of the spontaneous polarization field, and thereby the domain structure evolution, is described by the time-dependent nonlinear Ginzburg–Landau equation:

$$\frac{\partial P_i(\mathbf{r},t)}{\partial t} = -L\frac{\delta F}{\delta P_i(\mathbf{r},t)}, \tag{8.8}$$

where L is the kinetic coefficient, t denotes time and $\mathbf{r} = (x_1, x_2, x_3)$ is the spatial vector. $\delta F/\delta P_i(\mathbf{r}, t)$ represents the thermodynamic driving force for the spatial and temporal evolution of the ferroelectric system. Neglecting the body force in the materials, the mechanical equilibrium equation and the Gauss's equation are

$$\frac{\partial}{\partial x_j}\left(\frac{\partial F}{\partial \varepsilon_{ij}}\right) = 0, \tag{8.9}$$

and

$$\frac{\partial}{\partial x_i}\left(-\frac{\partial F}{\partial E_i}\right) = 0. \tag{8.10}$$

Using the variation or principle of virtual work, the weak form (or integral form) of the governing equation can be obtained as [20, 21]

$$\begin{aligned}\int_V \left[\frac{\partial f}{\partial \varepsilon_{ij}}\delta\varepsilon_{ij} + \frac{\partial f}{\partial E_i}\delta E_i + \frac{\partial P_i}{L\partial t}\delta P_i + \frac{\partial f}{\partial P_{i,j}}\delta P_{i,j} + \frac{\partial f}{\partial P_i}\delta P_i\right] \mathrm{d}V \\ = \int_S \left[t_i\delta u_i - \omega\delta\phi + \left(\frac{\partial f}{\partial P_{i,j}}n_j\right)\delta P_i\right] \mathrm{d}A,\end{aligned} \tag{8.11}$$

in which t_i is the surface traction, ω is the surface charge, and $(\partial f/\partial P_{ij})\ n_j$ is the surface gradient flux. One can obtain the time evolution of polarization in ferroelectrics by numerically solving this equation, e.g., the Newton iteration method is used to solve the nonlinear equations and the Euler backward method is used for the time integration in the simulation. The phase-field method is often used to study the behavior of nanoscale ferroelectrics [22–28].

8.1.2 Phase-field Modeling of Ferromagnetic Systems

In the phase-field model of ferromagnetic materials, the local magnetization $\boldsymbol{M} = (M_1, M_2, M_3)$ is used as an order parameter to describe the evolution of domain structures [29]. The stable magnetization configuration is determined by the minimization of the total free energy of the materials. In general, the total free energy of ferromagnetic material includes the magnetocrystalline anisotropy energy, exchange energy, elastic energy, and magnetostatic energy. The magnetocrystalline anisotropy energy is dependent on the orientation of magnetization. For a cubic crystal, it can be expressed as

$$f_{\text{anis}} = \frac{K_1}{M_S^4}(M_1^2M_2^2 + M_1^2M_3^2 + M_2^2M_3^2) + \frac{K_2}{M_S^6}(M_1^2M_2^2M_3^2), \tag{8.12}$$

where K_1 and K_2 are the magnetocrystalline anisotropy constants, and M_S is the magnitude of saturation magnetization. The exchange energy is a penalty for the spatially inhomogeneous distribution of magnetization in the ferromagnetic materials. It is a function of magnetization derivatives as

$$f_{\text{exch}} = \frac{A}{M_S^2}(M_{1,1}^2 + M_{1,2}^2 + M_{1,3}^2 + M_{2,1}^2 + M_{2,2}^2 + M_{2,3}^2 + M_{3,1}^2 + M_{3,2}^2 + M_{3,3}^2), \tag{8.13}$$

where A is the exchange stiffness constant.

Due to the magnetostriction effect, the magnetization distribution depends on the deformation in the ferromagnetic materials. Therefore, the elastic energy should be considered in the total free energy. The elastic energy includes the pure elastic energy and the coupling energy between the magnetization and strain, which is given by

$$\begin{aligned} f_{\text{elas}} = {} & \frac{1}{2}c_{11}(\varepsilon_{11}^2 + \varepsilon_{22}^2 + \varepsilon_{33}^2) + c_{12}(\varepsilon_{11}\varepsilon_{22} + \varepsilon_{22}\varepsilon_{33} + \varepsilon_{11}\varepsilon_{33}) \\ & + 2c_{44}(\varepsilon_{12}^2 + \varepsilon_{23}^2 + \varepsilon_{13}^2) \\ & - \frac{3\lambda_{100}}{2M_S^2}(c_{11} - c_{12})(\varepsilon_{11}M_1^2 + \varepsilon_{22}M_2^2 + \varepsilon_{33}M_3^2) \\ & - \frac{6\lambda_{111}}{M_S^2}c_{44}(\varepsilon_{12}M_1M_2 + \varepsilon_{23}M_2M_3 + \varepsilon_{13}M_1M_3), \end{aligned} \tag{8.14}$$

where c_{11}, c_{12}, and c_{44} are the elastic constants, λ_{100} and λ_{111} are the magnetostrictive constants, and ε_{ij} are the strain components. The magnetostatic energy is related to the external magnetic field and demagnetization field. It has the form as

$$f_{\text{mag}} = -\frac{1}{2}\mu_0(H_1^2 + H_2^2 + H_3^2) - \mu_0(H_1M_1 + H_2M_2 + H_3M_3), \tag{8.15}$$

where μ_0 is the permeability of the vacuum and H_i are the components of the magnetic field in the materials.

It should be noted that the energy forms in Eqs. (8.12)–(8.15) are valid only when the temperature is far below the Curie point. In such case, it is usually assumed that the magnitude of the magnetization vector equals the saturation magnetization M_S, i.e., $M_1^2 + M_2^2 + M_3^2 = M_S^2$. As a result, the evolution of the domain structures only depends on the change of the orientation of magnetization. This assumption is widely adopted in most micromagnetic models in the literature [30]. But the constraint of $M_1^2 + M_2^2 + M_3^2 = M_S^2$ usually requires special techniques

in the numerical solutions to the governing equations. To replace the constraint of $M_1^2 + M_2^2 + M_3^2 = M_S^2$, the following energy is added to the total free energy

$$f_{\text{cons}} = A_S(M - M_S)^2, \tag{8.16}$$

where M is the magnitude of magnetization, and A_S is the constraint energy coefficient. When the magnitude of magnetization vector deviates from the M_S, it provides an energy penalty. The penalty energy can keep the magnetization magnitude close to M_S, which has certain physical ground [31]. The total free energy density in the materials should be the summation of all the energy terms in Eqs. (8.12)–(8.16) as

$$f = f_{\text{anis}} + f_{\text{exch}} + f_{\text{elas}} + f_{\text{mag}} + f_{\text{cons}}. \tag{8.17}$$

Based on the total free energy, the following time-dependent Ginzburg–Landau (TDGL) equation [17, 18] is employed to describe the evolution of magnetization

$$\frac{\partial \mathbf{M}(\mathbf{r}, t)}{\partial t} = -L\frac{\delta F}{\delta \mathbf{M}(\mathbf{r}, t)}, \tag{8.18}$$

where L is the kinetic coefficient and $F = \int f \mathrm{dv}$ is the total free energy of the system. Neglecting the body force in the materials, the mechanical equilibrium equation and the Maxwell's equation are

$$\frac{\partial}{\partial x_j}\left(\frac{\partial F}{\partial \varepsilon_{ij}}\right) = 0, \tag{8.19}$$

and

$$\frac{\partial}{\partial x_i}\left(-\frac{\partial F}{\partial H_i}\right) = 0. \tag{8.20}$$

To solve the above partial differential equations, a 3D finite element method is employed. Ignoring the detailed derivation, the weak form for the 3D finite element method reads

$$\begin{aligned} &\int_V \left[\frac{\partial F}{\partial \varepsilon_{ij}}\delta\varepsilon_{ij} + \frac{\partial F}{\partial H_i}\delta H_i + \left(\frac{\partial M_i}{\partial t} + L\frac{\partial F}{\partial M_i}\right)\delta M_i + L\frac{\partial F}{\partial M_{i,j}}\delta M_{i,j}\right]\mathrm{d}v \\ &\quad = \int_S \left[t_i \delta u_i - B_n \delta\phi + \left(\frac{\partial F}{\partial M_{i,j}} n_j\right)\delta M_i\right]\mathrm{d}s, \end{aligned} \tag{8.21}$$

in which t_i is the traction on the surface, u_i are the mechanical displacements, ϕ is a scalar magnetic potential, B_n is the normal component of the magnetic induction, and $(\partial F/\partial M_{i,j})\, n_j$ is the gradient flux of magnetization on the surface. For the finite

element solution of Eq. (8.21), the eight-node brick element is employed in the space discretization. There are seven degrees of freedom at each node, which are three displacement components, one magnetic potential component, and three magnetization components. The Newton iteration method is employed to solve the nonlinear equations, and the Euler backward method is used for the time integration in the simulation. The phase-field method is often used to study the behavior of nanoscale ferromagnets [29, 32–34].

8.1.3 Phase-field Modeling of Multiferroic Composites

In the phase-field modeling of FE/FM nanocomposites, the multiferroic system is described by two vector fields (order parameters): a polarization field $\mathbf{P}(\mathbf{r})$ and a magnetization field $\mathbf{M}(\mathbf{r})$. The total free energy of multiferroic nanocomposite is, then, expressed by [35, 36]

$$F = \int_V \left[(1-\eta)f_p + \eta f_{\mathrm{m}} + f_{\mathrm{elec}} + f_{\mathrm{mag}} + f_{\mathrm{elas}}\right] \mathrm{d}V, \tag{8.22}$$

where the phase factor $\eta = 0$ denotes FE phase while $\eta = 1$ shows the FM phase. f_p denotes the sum of the Landau and gradient free energy densities, while f_{m} denotes the sum of the magnetocrystalline anisotropy, exchange, and penalty free energy densities.

In contrast to purely ferroelectric or ferromagnetic systems, electric polarization and magnetic moment in multiferroic composites interact with each other via the interfacial strain between FE and FM materials through electrostrictive and magnetostrictive effects (Fig. 8.1a), the electric field generated FE phase (Fig. 8.1b), and the magnetic field generated by FM phase (Fig. 8.1c). To include these effects, the total free energy of multiferroic system contains the contributions of the elastic free energy densities including the electrostrictive and magnetostrictive effects, f_{elast}, electrostatic and magnetostatic free energy densities, f_{elec} and f_{mag}.

In detail, the elastic energy density of the system, f_{elas}, can be expressed as

$$f_{\mathrm{elas}} = c_{ijkl}(\varepsilon_{ij} - \varepsilon^0_{ij})(\varepsilon_{kl} - \varepsilon^0_{kl}), \tag{8.23}$$

where ε_{ij} is the total strain, ε^0_{ij} is the stress-free strain due to the electrostrictive effect or magnetostrictive effect, and c_{ijkl} are the elastic stiffness tensors of FE or FM materials. In FE/FM composites, the ε^0_{ij} can be written as

$$\varepsilon^0_{ij} = \begin{cases} \eta\left[\frac{3}{2}\lambda_{100}\left(m_i m_j - \frac{1}{3}\right)\right] + (1-\eta)Q_{ijkl}P_kP_l & (i=j) \\ \eta\left(\frac{3}{2}\lambda_{111}m_i m_j\right) + (1-\eta)Q_{ijkl}P_kP_l & (i\neq j) \end{cases}, \tag{8.24}$$

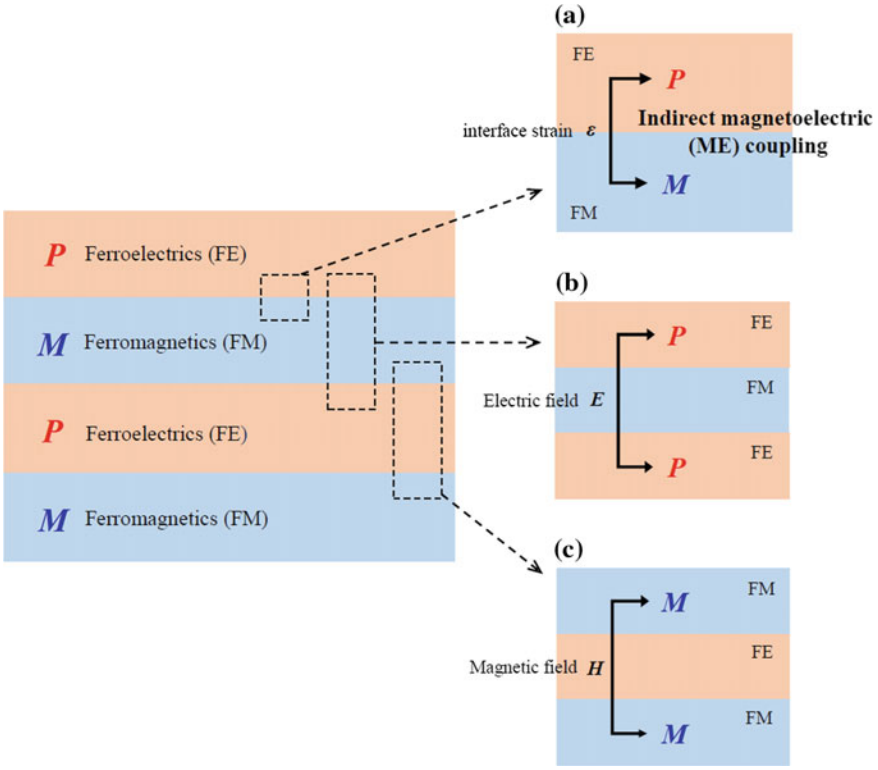

Fig. 8.1 Schematic illustration of interactions in the multiferroic composite among polarization P and magnetization M. **a** Interaction between polarization P and magnetization M via interface strain ε. **b** Interaction between polarization P in each of the ferroelectrics via electric field E. **c** Interaction between magnetization M in each of the ferromagnetics via magnetic field H. *Red* and *blue* areas indicate ferroelectrics and ferromagnetics, respectively

where Q_{ijkl} are the electrostrictive coefficients and λ_{100} and λ_{111} are the magnetostrictive constants.

The simultaneous evolution of the polarization and magnetization fields toward their thermodynamic equilibrium distributions is driven by the decrease of total free energy in the multiferroic nanocomposite system, which is described the time-dependent Ginzburg–Landau (TDGL) equations,

$$\frac{\partial P_i(x,t)}{\partial t} = -L_{\mathrm{FE}} \frac{\partial F}{\delta P_i(x,t)}, \tag{8.25}$$

$$\frac{\partial M_i(x,t)}{\partial t} = -L_{\mathrm{FM}} \frac{\partial F}{\delta M_i(x,t)}, \tag{8.26}$$

where L_{FE} and L_{FM} are kinetic coefficient for the time evolution of spontaneous polarization and magnetization, respectively. In addition to the TDGL equations, the following mechanical equilibrium equation

$$\frac{\partial \sigma_{ij}}{\partial x_j} = \frac{\partial}{\partial x_j}\left(\frac{\partial F}{\partial \varepsilon_{ij}}\right) = 0, \tag{8.27}$$

and the Maxwell's and Gauss equations

$$\frac{\partial D_i}{\partial x_i} = \frac{\partial}{\partial x_i}\left(-\frac{\partial F}{\partial E_i}\right) = 0, \tag{8.28}$$

$$\frac{\partial B_i}{\partial x_i} = \frac{\partial}{\partial x_i}\left(-\frac{\partial F}{\partial H_i}\right) = 0, \tag{8.29}$$

must be simultaneously satisfied for the body force, charge, and magnetic free in the multiferroic nanocomposites.

In order to solve the governing Eqs. (8.25)–(8.29) in real space, a nonlinear multifield coupling finite element method is employed. In the finite element method, the governing Eqs. (8.25)–(8.29) are expressed in the integral form (or weak form) as

$$\begin{aligned}\int_V &\left\{\begin{array}{l} \frac{\partial f}{\partial \varepsilon_{ij}}\delta\varepsilon_{ij} + \frac{\partial f}{\partial E_i}\delta E_i + \frac{\partial f}{\partial H_i}\delta H_i + \left(\frac{1}{L_{FE}}\frac{\partial P_i}{\partial t} + \frac{\partial f}{\partial P_i}\right)\delta P_i + \frac{\partial f}{\partial P_{i,j}}\delta P_{i,j} \\ + \left(\frac{1}{L_{\mathrm{FM}}}\frac{\partial M_i}{\partial t} + \frac{\partial f}{\partial M_i}\right)\delta M_i + \frac{\partial f}{\partial M_{i,j}}\delta M_{i,j} \end{array}\right\}\mathrm{d}v \\ &= \int_S \left\{t_i\delta u_i - w\delta\phi_{\mathrm{P}} - B_n\delta\phi_{\mathrm{M}} + \left(\frac{\partial f}{\partial P_{i,j}}n_j\right)\delta P_i + \left(\frac{\partial f}{\partial M_{i,j}}n_j\right)\delta M_i\right\}\mathrm{d}s,\end{aligned} \tag{8.30}$$

where t_i is the surface traction, u_i is the mechanical displacement, w denotes surface charge, ϕ_{P} is electrostatic potential, B_n is the normal component of the magnetic induction, ϕ_{M} is a scalar magnetic potential, and n_j denotes the components of normal unit vector of surfaces.

8.2 Nanometamaterials of Ferroelectrics

8.2.1 Unusual Domain Patterning in Nanometamaterials

As a preliminary work on ferroelectric nanometamaterials, two-dimensional (2D) nanometamaterials based on the Archimedean lattice concept are modeled to study

their ferroelectric properties [14]. Archimedean lattices proposed by Kepler are defined as a plane consisting of regular polygons in which all vertices are connecting each other. In 2D, there are 11 Archimedean lattices: Three Archimedean lattices consist of a specific polygon (squares, triangles, or hexagons), and eight Archimedean lattices consist of the combination of two or more different polygons, which gives rise to further complicated internal structures and resulting polarization configuration as follows:

In the square, honeycomb, and triangular specimens, spontaneous polarizations orient along the lattice and form a rectilinear or single domain in each lattice, as shown in Fig. 8.2a–c. The single domain of polarization is formed to decrease overall electrostatic energy through elimination or reduction of the depolarization fields at free surfaces. At lattice junctions, polarizations gradually change their orientations to connect all single domains and flow smoothly with a head-to-tail polarization. Such a continuous flow of polarization stabilizes the system by minimizing the gradient term of free energy and become energetically favorable in square, honeycomb, and triangular metamaterials.

In the SrCuBO, Star, and CaVO metamaterials, on the other hand, spontaneous polarizations in some lattices terminate the continuous connection and then curl their orientation to form microvortices while others form a single domain and make a continuous connection of single domains at the lattice junctions, as shown in Fig. 8.2d–f, respectively. The coexistence of microvortex and single domain exhibits an unusual polarization pattern and characterizes for the domain structure of SrCuBO, Star, and CaVO metamaterials.

Single domains are formed in the Bounce and Kagome metamaterials as shown in Fig. 8.2g–h, similar to the square, honeycomb, and triangular metamaterials. But, in these metamaterials, the single-domain polarization in each lattice consists a mesoscale flux-closure configurations as indicated by a black circular arrows, i.e., formation of a polarization mesovortex. Differing from the conventional polarization vortices in ferroelectric nanodots [37], this mesovortex lacks the vortex core, which provides locally high energy due to mismatch of mechanical and electric fields. The elimination of the energetically unfavorable vortex core may stabilize the mesovortex configuration [14].

The most complicated and exotic polarization patterns can be found in the respective trellis, maple-leaf, and SHD metamaterials (see Fig. 8.2i–k). Here, the polarization micro- and mesovortex simultaneously appears in these systems with two different length scales, i.e., the formation of hierarchical vortices. As a result of the coexistence of hierarchical rectilinear, micro- and mesovortex polarizations, this class of metamaterials simultaneously exhibits both the order parameters of polarization and toroidal moments.

In this way, the geometry of nanometamaterials provides a plethora of interesting domain configurations that differ from those observed in a simple nanoscale component. Although this pioneering study demonstrates a lot of possibilities of ferroelectric metamaterials with 2D samples, 3D metamaterials should be more

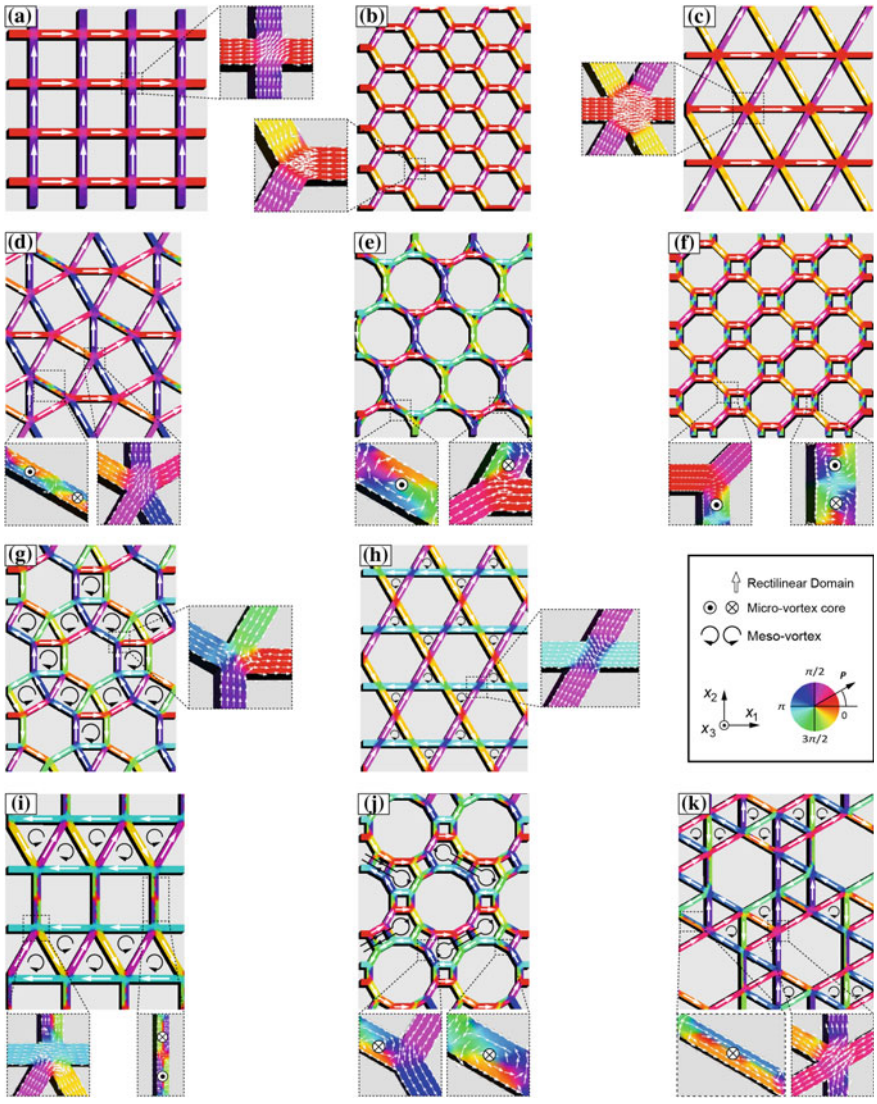

Fig. 8.2 Polarization distribution in ferroelectric nanometamaterials of **a** square, **b** honeycomb, **c** triangular, **d** SrCuBo, **e** Star, **f** CaVo, **g** Bounce, **h** Kagome, **i** Trellis, **j** SHD, and **k** Maple-leaf specimens. The contour colors indicate the polarization direction. Reprinted from Ref. [15], Copyright 2016, with permission from Elsevier

promising due to larger degrees of freedom in geometry and polar orientations. Hence, the domain engineering via the 3D metamaterial concept is expected to advance this field further in the near future.

8.2.2 *Tunable Polar and Toroidal Electromechanical Properties*

What is the role of unusual domain configuration appearing in ferroelectric nanometamaterials? This is quite an important question to what the metamaterial concept in ferroelectrics brings about in the materials science and engineering fields.

An answer has been provided through the evaluation of piezoelectric responses in ferroelectric metamaterials. Figure 8.3 shows the longitudinal and transverse piezoelectric coefficients, e_{11} and e_{12}, respectively, for Archimedean-type ferroelectric nanometamaterials [15]. Although the piezoelectric coefficient is a material constant, one can find a variety of piezoelectric responses depending on the geometry of nanometamaterials. This means that the metamaterial concept can tailor and design the piezoelectric properties: In particular, most of the nanometamaterials have positive e_{11} values while the Star, Bounce, and Trellis nanometamaterials show "zero" e_{11} value, i.e., the absence of piezoelectric coefficient. Since piezoelectricity is intrinsic property to all natural ferroelectric materials, the zero piezoelectric coefficient indicates an unusual longitudinal piezoelectric response in the artificial nanometamaterial of ferroelectric; that is, nonpiezoelectric ferroelectrics can be designed via the metamaterial concept. Another interesting feature can be found in the transverse piezoelectric coefficients: Although most of the nanometamaterials possess a negative value of e_{12}, which is commonly observed in homogenous ferroelectric [38, 39], the Honeycomb and Triangular metamaterials exceptionally exhibit "positive" e_{12}, leading to an anomalous transverse piezoelectric response beyond the homogeneous ferroelectrics. These unusual

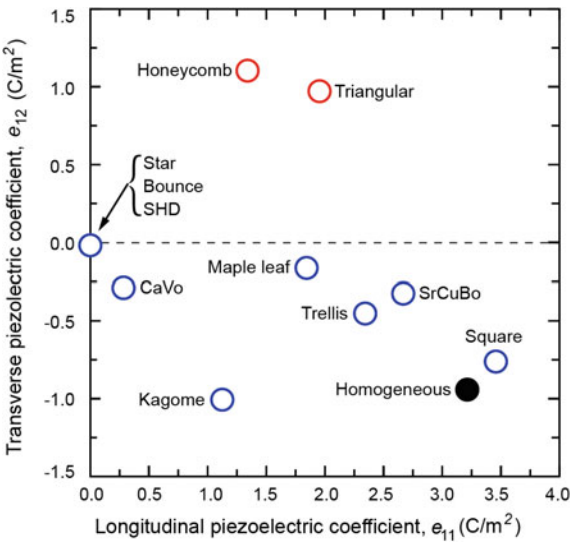

Fig. 8.3 Longitudinal and transverse piezoelectric coefficients in ferroelectric nanometamaterials with different internal structures. *Blue* and *red* circles correspond to the internal structures possessed negative and positive magnitudes of transverse piezoelectric coefficient, respectively. Reprinted from Ref. [15], Copyright 2016, with permission from Elsevier

piezoelectric responses originate from the unique polar domain configuration observed in ferroelectric nanometamaterials. Therefore, the metamaterial concept provides a platform to tailor and design piezoelectric properties beyond the conventional ferroelectrics.

As shown in the previous section, some of ferroelectric nanometamaterials simultaneously exhibit polarization (meso- or micro-)vortices. A new order parameter of ferrotoroidicity, "toroidal moment," can also be introduced to characterize strength of a vortex polarization [37, 40]. Although piezoelectricity, a response of polarization with respect to mechanical strain, has been widely considered, interaction between this new order parameter of toroidal moment and mechanical strain remains unknown.

A first attempt to characterize the response of toroidal moment in ferroelectric metamaterials to mechanical strain, as proposed as "piezotoroidicity," has been made in Ref. [15]: Fig. 8.4 shows toroidal moment G_3 in various nanometamaterials as a function of uniaxial strain ε_{11}, where the toroidal moment can be evaluated by

$$G_3 = \frac{\sum r_i \times P_i \Delta v_i}{V}. \tag{8.31}$$

Here, P_i and ΔV_i denote the local polarization in ith element at the position of r_i and the volume of ith element, respectively [37]. The toroidal moment is monotonically changes with increasing strain, where increasing trend is found in the Bounce, SHD, and Maple-leaf metamaterials while the moment keeps almost

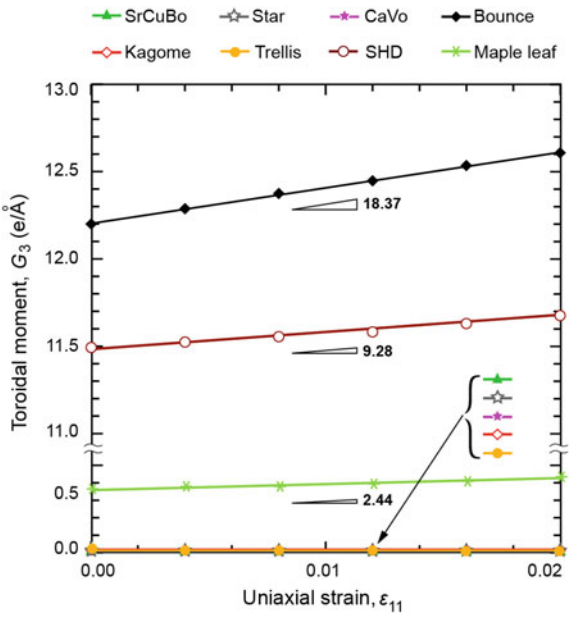

Fig. 8.4 Toroidal moment G_3 as a function of strain ε_{11} in ferroelectric nanometamaterials. The slope of G_3-ε_{11} relation presents the piezotoroidic coefficient. Reprinted from Ref. [15], Copyright 2016, with permission from Elsevier

constant value in the others. In the same manner to the piezoelectric coefficient, a new physical quantity of "piezotoroidic coefficient" can be defined as

$$e_{31}^{g} = \frac{\partial G_3}{\partial \varepsilon_{11}}. \tag{8.32}$$

Thus, the slope of G_3-ε_{11} corresponds to the piezotoroidic coefficient e_{31}^{g}. The slope for remarkable Bounce, SHD and Maple-leaf metamaterials is also shown in Fig. 8.4. The nonzero value indicates the presence of piezotoroidic effect in Bounce, SHD and Maple-leaf nanometamaterials, and the Bounce metamaterial shows the highest value and large piezotoroidic effects, while the rest shows zero since G_3 is suppressed due to opposite chirality of polarization vortices. Diversity of G_3-ε_{11} relations and e_{31}^{g} values means that the piezotoroidic effect strongly depends on the geometry and resulting domain configuration in ferroelectric nanometamaterials. Similar tendency is observed in the other piezotoroidic components [15]. Thus, a new physical quantity of "piezotoridicity" can be designed by metamaterial concept in ferroelectrics.

8.2.3 *Multiple Hysteresis Behaviors*

Figure 8.5a shows spontaneous polarization configuration in nanoporous specimen at thermodynamically equilibrium state without external field. Polarizations mostly align along the ligament lattice and, thereby, form a rectilinear domain accommodated in each ligament lattice. The formed rectilinear domains connect to each other at the lattice junctions, keeping the head-to-tail arrangement to form a continuous flow pattern. The connection preferably involves single domains in two neighboring lattices that are jointed at the junction to maintain the head-to-tail arrangement. This phenomenon is consistent with that observed in ferroelectric nanometamaterial and is discussed in the previous section. As a result, a zigzag chain of domain configuration is emerged in the ferroelectric triangular nanoporous, i.e., the zigzag chain flow along the x_1 direction.

Since the triangular nanoporous structure possesses sixfold rotational symmetry and similar to the graphene lattice configuration, there are six ways of polarization connection to keep the zigzag chain flow through the nanoporous structure. Figure 8.5b represents all possible stable states of spontaneous polarization, which are named by state I to state VI. Here, the number of equilibrium spontaneous states is increased to the structure's rotational symmetry. In each state, single domains of zigzag chains are contributed to flow of polarization which is specifically defined by average polarization. The equivalent states of spontaneous polarization are distinguished by their average polarization direction. It can be seen that average polarization possesses six symmetrical angles. The magnitude of the average polarization

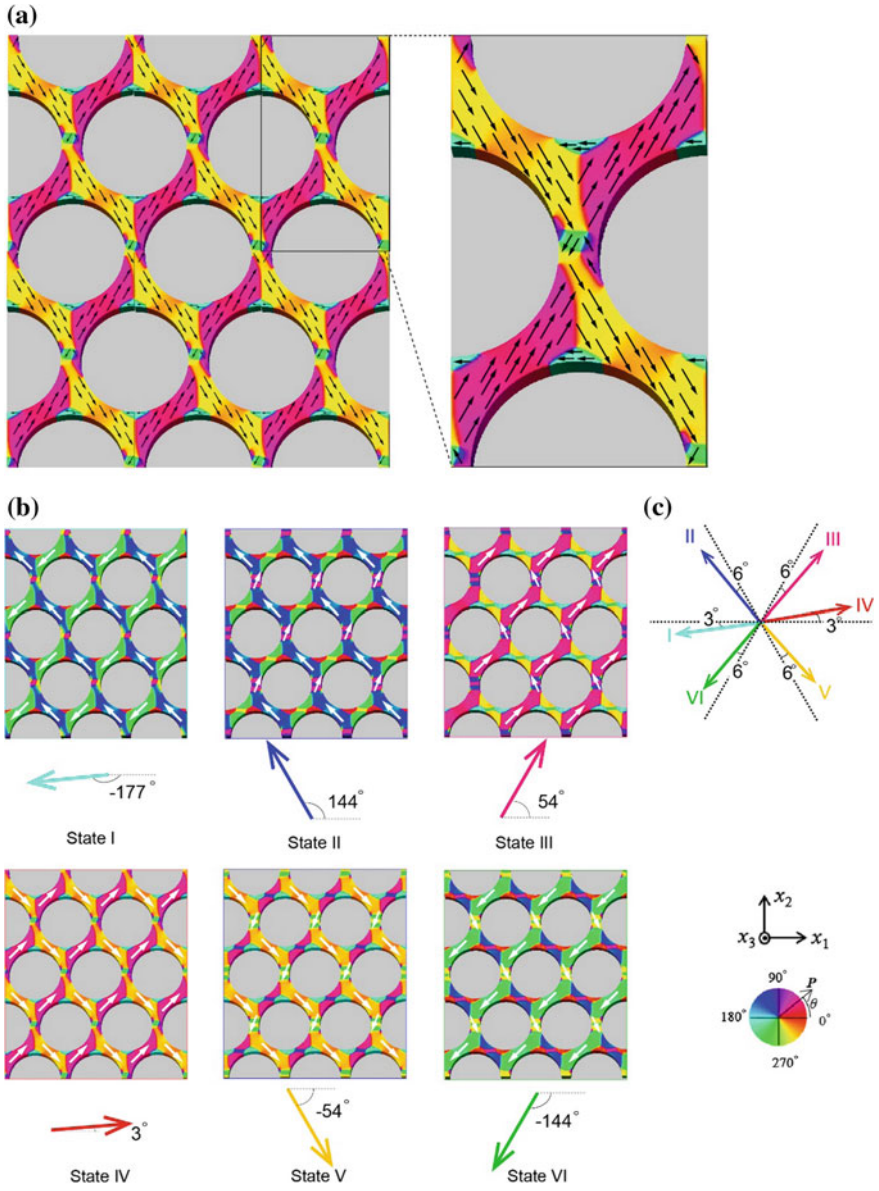

Fig. 8.5 **a** Representation of spontaneous polarization distribution in triangular nanoporous structure. **b** Six equivalent stable states of polarization and their average polarization orientation. **c** Distribution of average polarization in six stable states [41]

and its angle respect to x_1 direction θ are shown in Fig. 8.5c. Clearly, the average polarization vector is mostly aligned along the symmetry axes of triangular structure, where the slight deviation of the average polarization from ideal symmetrical

distribution is due to an effect of crystal orientation. This indicates that the internal shape is dominant over the intrinsic ferroelectric crystalline and it is different from common homogenous polarization distribution in which polarization direction highly depends on crystallographic axis. This result indicates that in nanoporous structure, pores induce multiple polar states more than the intrinsic one, and the geometry of pore arrangement determines the number of stable states. In consequence, in triangular nanoporous structure, six equivalent spontaneous states of polarization are obtained in which the average polarization is oriented along symmetrical axes of structure.

Due to the multiple stable states created by the nanoporous or metamaterial concept, this porous ferroelectrics exhibit unique hysteresis behavior by the application of electric field: Fig. 8.6a represents change of polarization with applied electric field. In a small value of electric field, magnitude of polarization stays almost constant. At $E = 6.83$ kV, magnitude of polarization is abruptly increased with keeping a negative value. Then, polarization is slightly increasing until $E = 10.06$ kV. At $E = 10.06$ kV, the magnitude of polarization is dramatically changed and it becomes positive. By applying a larger electric field, magnitude of polarization is slightly increasing until the other jump in magnitude of polarization occurs at $E = 14.28$ kV; however from this step, by increasing electric field more, no significant change is happened in magnitude of positive polarization. Considering the pattern of polarization in each of the steps shows that polarization goes from state I to state II, then switches to state III and finally reaches to state IV and 3-step switching is occurred. By reversing electric field direction, a similar incident is seen and polarization goes through state V, VI, and finally state I, hysteresis loop is thus completed as shown in Fig. 8.6a.

Given such unusual multilevel hysteresis loop in nanoporous structure, a comprehensive picture of 3-step hysteresis loop is provided by considering the local switching between neighboring states at an intermediate magnitude of electric field, at which the states II, III, V, and VI appear. As shown in 8.6b–d, an appropriate choice of electric field path (magnitude) can achieve a small local hysteresis loop, which can switch the intermediate state to another intermediate states, such as I-to-II, IV-to-V. As a whole, all six possible (meta)stable configurations of the nanoporous material can be switched by an electric field, leading to a multiple hysteresis loops, as summarized in Fig. 8.6f. It should be noted that all six states can be realized without electric field. According to Fig. 8.6f, six states can be switched from one to the other and the transformation path is defined by electric field. Neighboring states can be directly switched to each other through global loop or local loops. Furthermore, it shows possibility of shortcut switching from one state to its far state by means of local loop I-II-III and IV-V-VI in which it is not necessary to pass through intermediate states to reach desired state and it can increase the data storage and accelerate writing speed in memory devices such as FeRAM. It suggests a new way to not only increase the storage density of ferroelectric memory, but also freely tailor transition between states.

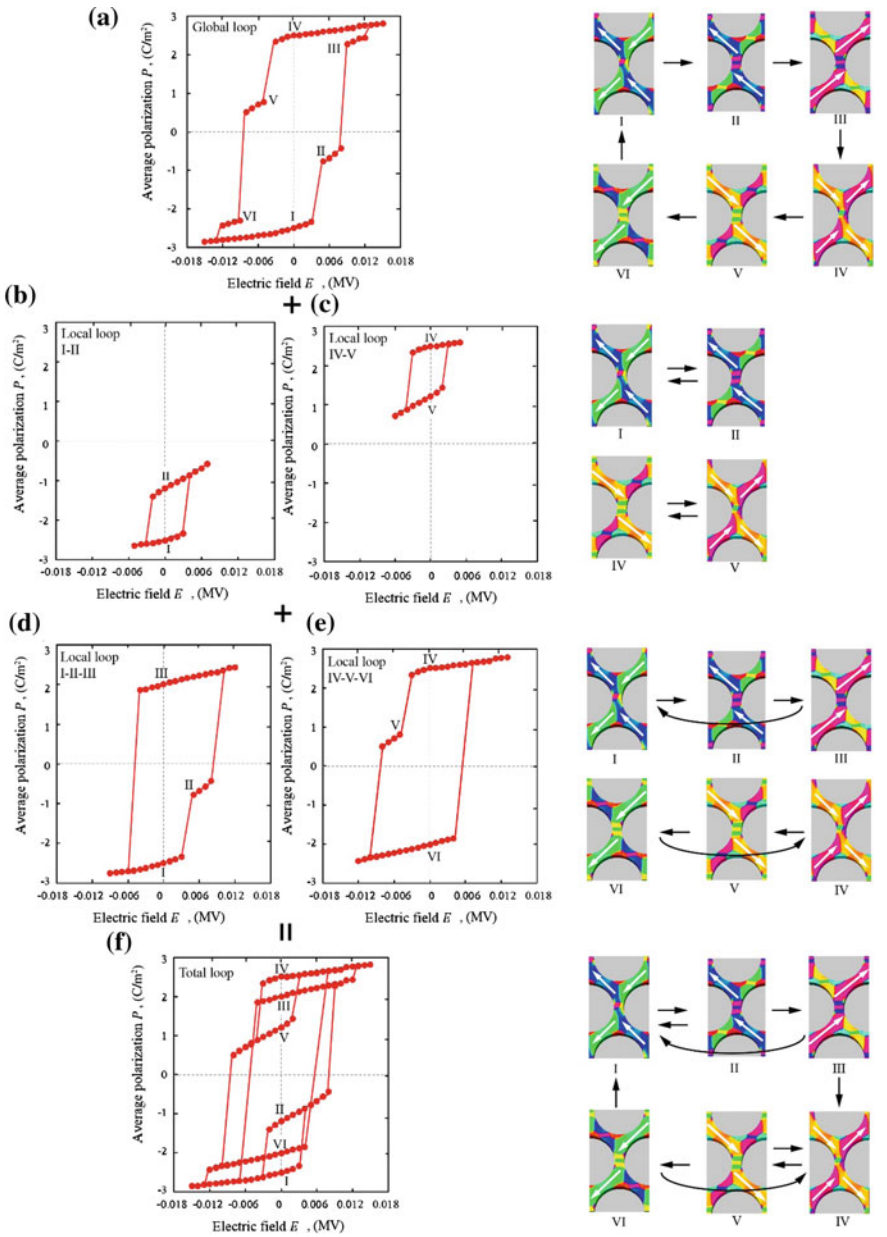

Fig. 8.6 P-E hysteresis loops of ferroelectric nanoporous subjected to electric field with $\alpha = 20°$ **a** outer (global) loop, **b–c** local loop of 2 states, **d–e** local loop of 3 states, and **f** total loops [41]

8.3 Multiferroic Nanocomposites

As shown above, nanometamaterials exhibit unique and interesting domain configurations depending on the geometry of internal structures. On the other hand, nanocomposites of ferroelectrics and ferromagnets have attracted much attention due to their large magnetoelectric effect and its significant dependence on geometry. The FE/FM (multiferroic) nanocomposites seem to have unique ferroelectric/ferromagnetic domain configurations, which can be engineered by the internal structure of nanocomposites, as similar as nanometamaterials. Here, we review a pioneering work to demonstrate the mutliferroic domain configurations in nanocomposites.

Figure 8.7 shows polar and magnetic domains in the square-arrayed FE ($BaTiO_3$)/FM($CoFe_2O_4$) nanocomposites [16]. An ultrafine FE domain with a stripe pattern is formed in the FE matrix. The thickness of the FE domain is approximately 10 nm. The domains elongated in the direction 45° from the x_1-axis, and inside of the ultrafine domains, a more complicated cross-hatched pattern of superdomains is formed, especially, between two neighboring FM nanowires (Fig. 8.7a_1–a_2). Such coexistence of ultrafine stripes and the cross-hatched pattern leads to a hierarchical FE domain structure in the nanocomposite. In contrast to the FE domains, the constituent FM nanowire forms a single magnetic domain with a uniform magnetization direction among the neighboring nanowires. Hence, the multiferroic nanocomposite consists of a hierarchical ultrafine FE and a single FM domains.

Similar to nanometamaterials, these unique FE/FM domain configurations in multiferroic composite show dependence on internal geometry of composites: Fig. 8.8a, b shows the FE and FM domain structures in different geometry of honeycomb- and triangular-arrayed nanocomposites, respectively. The honeycomb array nanocomposite exhibits a cross-hatched and bubble-like patterns of ultrafine FE domains (Fig. 8.8a) while the triangular-arrayed nanocomposite shows triangular nanodomains consisting of a sparse ultranarrow bands (Fig. 8.8b). In addition, FM domains also form a different patterns; multidomain in the honeycomb and quasi-single domain in the triangular nanocomposites. These also indicate that the multiferroic domain configuration in the FE/FM nanocomposites is significantly dependent on internal geometry, meaning a possible way of designing the domain patterns, reminiscent of nanometamaterial concepts. Such unique nanodomain configuration in FE/FM nanocomposites provides excellent functionalities, such as colossal magnetoelectric (ME) effect: Fig. 8.9 presents the ME coefficients for the various nanocomposites. The nanocomposites exhibit the ME coefficients in the order of 10^{-9}–10^{-8} s/m, which is one or two order higher than that of bulk composites with the same composition [42–44]. More interestingly, the ME coefficient also significantly depends on the internal geometry of the composites, e.g., triangular-arrayed nanocomposite shows the highest ME coefficient, approximately

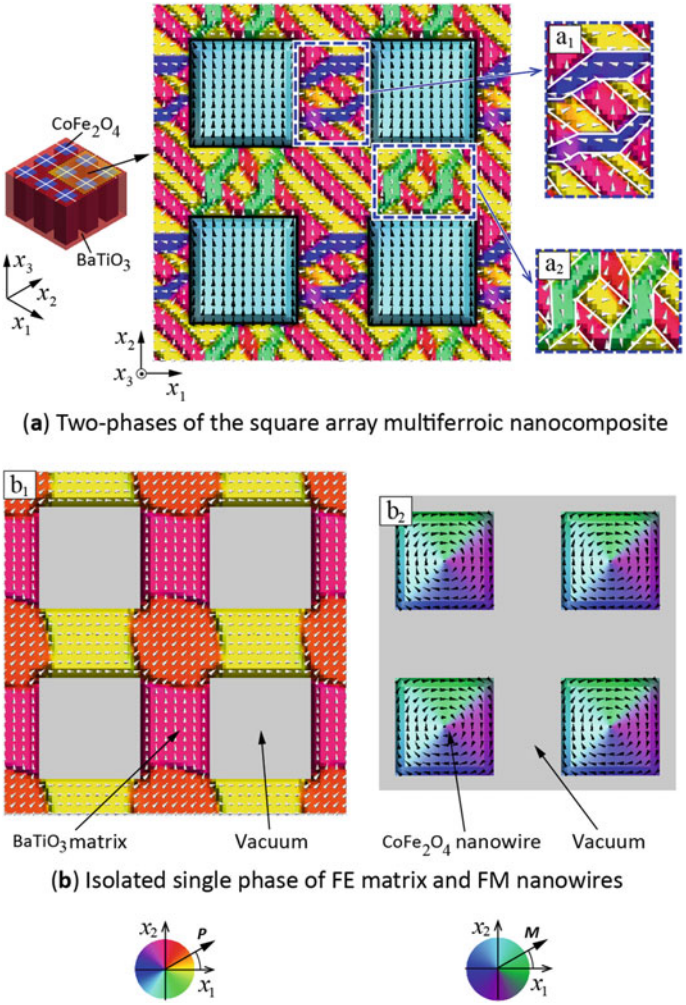

Fig. 8.7 Distribution of polarization and magnetization in **a** the two phases of the square-arrayed multiferroic nanocomposite and **b** isolated single phases of the ferroelectric matrix and ferromagnetic nanowires. The FE and FM domain variants are represented by different colors. The small *white* and *black arrows* indicate the directions of the polarization and magnetization vectors, respectively. Reprinted with permission from Ref. [16]. Copyright 2015, American Institute of Physics

one order of magnitude larger than in the other nanocomposites. Such colossal ME effect and its geometry-dependence suggests a new strategy to efficiently design multiferroic materials.

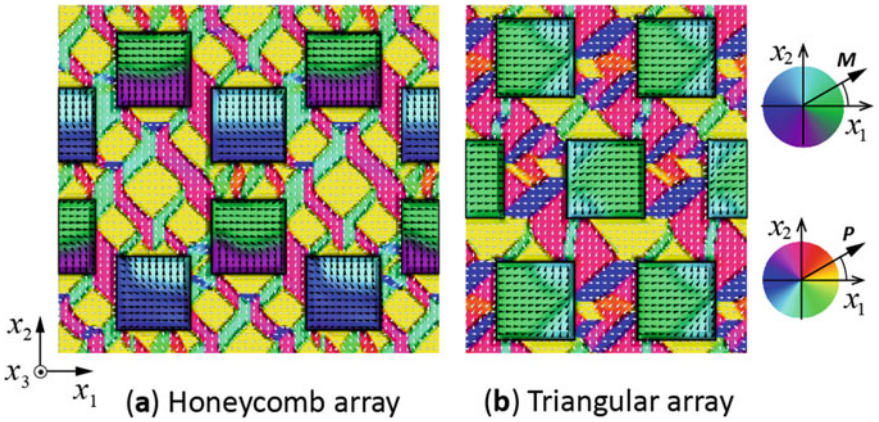

Fig. 8.8 Distribution of polarization and magnetization in multiferroic nanocomposites with **a** a honeycomb array and **b** a triangular array of ferromagnetic nanowires. Reprinted with permission from Ref. [16]. Copyright 2015, American Institute of Physics

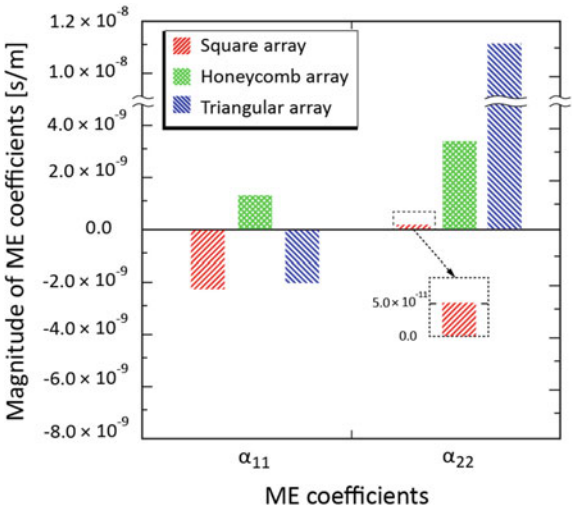

Fig. 8.9 Magnetoelectric coefficients for multiferroic nanocomposites with *square*, *honeycomb*, and *triangular* arrays of ferromagnetic nanowires. Reprinted with permission from Ref. [16]. Copyright 2015, American Institute of Physics

8.4 Conclusion

In this chapter, we reviewed pioneering studies on the next-coming ferroelectric nanometamaterials and nanocomposites. In Sect. 8.1, we first introduced a phase-field method that describes the ferroelectric and/or ferromagnetic domain evolutions in such nanometamaterials or composites.

With the phase-field method introduced in Sect. 8.1, we discussed how unique domain configurations emerge by the internal structure of metamaterials in

Sect. 8.2. At the same time, we introduced that the unique domain configurations designed by the nanometamaterial concept provide a new functionality that cannot be observed in homogeneous ferroelectrics or even simple nanocomponents.

As an extension of the nanometamaterial concept, In Sect. 8.3, we review multiferroic FE/FM nanocomposites. Not only the geometrical effects but also ferroelectric-ferromagnetic interactions through interfacial strain give rise to more complex domain configurations and lead to a colossal magnetoelectric effect. These characteristics and functionalities can be further tailored by geometry of internal structures, which is promising for the future design of ferroic properties.

References

1. D.R. Smith, W.J. Padilla, D.C. Vier, S.C. Nemat-Nasser, S. Schultz, Phys. Rev. Lett. **84**, 4184 (2000)
2. R.A. Shelby, D.R. Smith, S. Schultz, Science **292**, 77 (2001)
3. M. Choi et al., Nature **470**, 369 (2011)
4. S. Linden, et al. Science **306**, 1351 (2004)
5. M. Decker, M.W. Klein, M. Wegener, S. Linden, Opt. Lett. **32**, 856 (2007)
6. Z. Liu, H. Lee, Y. Xiong, C. Sun, X. Zhang, Science **315**, 1686 (2007)
7. N.I. Landy, S. Sajuyigbe, J.J. Mock, D.R. Smith, W.J. Padilla, Phys. Rev. Lett. **100**, 207402 (2008)
8. M. Kauranen, A.V. Zayats, Nat. Photonics **6**, 737 (2012)
9. M. Kadic, T. Buckmann, R. Schittny, M. Wegener, Rep. Prog. Phys. **76**, 126501 (2013)
10. T.A. Schaedler et al., Science **334**, 962 (2011)
11. T. Bückmann et al., Adv. Mater. **24**, 2710 (2012)
12. Z.G. Nicolaou, A.E. Motter, Nature Mater. **11**, 608 (2012)
13. T. Bückmann, M. Thiel, M. Kadic, R. Schittny, M. Wegener, Nat. Commun. **5**, 4130 (2014)
14. T. Shimada, L.V. Lich, K. Nagano, J. Wang, T. Kitamura, Sci. Rep. **5**, 14653 (2015)
15. L.V. Lich, T. Shimada, S. Sepideh, J. Wang, T. Kitamura, Acta. Mater. **113**, 81–89 (2016)
16. L.V. Lich, T. Shimada, K. Miyata, K. Nagano, J. Wang, T. Kitamura, Appl. Phys. Lett. **107**, 232904 (2015)
17. A. Gordon, I.D. Vagner, P. Wyder, Phys. Rev. B **41**, 658 (1990)
18. Y. Ni, L.H. He, A.G. Khachaturyan, J. Appl. Phys. **108**, 023504 (2010)
19. J. Wang, Appl. Phys. Lett. **97**, 192901 (2010)
20. A.F. Devonshire, Phil. Mag. (Suppl. 3) 85e130 (1954)
21. J. Wang, S.Q. Shi, L.Q. Chen, Y. Li, T.Y. Zhang, Acta Mater. **52**, 749–764 (2004)
22. J. Wang, M. Kamlah, Phys. Rev. B **80**, 012101 (2009)
23. J. Wang, T.Y. Zhang, Phys. Rev. B **73**, 144107 (2006)
24. J. Wang, T.Y. Zhang, Acta. Mater. **55**, 2465–2477 (2007)
25. J. Wang, T.Y. Zhang, Appl. Phys. Lett. **86**, 192905 (2005)
26. J. Wang, M. Kamlah, Smart Mater. Struct. **18**, 104008 (2009)
27. Y. Tong, M. Liu, H.M. Chen, G.P. Li, H. Fang, J. Wang, Z. Ma, J. Appl. Phys. **117**, 074102 (2015)
28. L.-Q. Chen, Annu. Rev. Mater. Res. **32**, 113 (2002)
29. J. Wang, J. Zhang, T. Shimada, T. Kitamura, J. Phys.: Condens. Matter **25**, 226002 (2013)
30. J.K. Ha, R. Hertel, J. Kirschner, Phys. Rev. B **67**, 224432 (2003)
31. C.M. Landis, J. Mech. Phys. Solids **56**, 3059 (2008)
32. J. Wang, G.-P. Li, T. Shimada, H. Fang, T. Kitamura, Appl. Phys. Lett. **103**, 242413 (2013)
33. J. Wang, J. Zhang, Int. J. Solids Struct. **50**, 3597 (2013)

34. J. Wang, G.P. Li, Comput. Mater. Sci. **108**, 316–322 (2015)
35. H.T. Chen, Y. Ni, A.K. Soh, J. Appl. Phys. **113**, 134102 (2013)
36. T.N. Yang, J.M. Hu, C.W. Nan, L.Q. Chen, Appl. Phys. Lett. **104**, 052904 (2014)
37. I.I. Naumov, L. Bellaiche, H. Fu, Nature **432**, 737 (2004)
38. V.I. Aleshin, I.P. Raevski, J. Appl. Phys. **112**, 114101 (2012)
39. V.I. Aleshin, I.P. Raevski, J. Appl. Phys. **113**, 224105 (2013)
40. T. Shimada, X. Wang, Y. Kondo, T. Kitamura, Phys. Rev. Lett. **108**, 067601 (2012)
41. S. Sepideh, Master's Dissertation, Kyoto University, (unpublished)
42. E. Ascher, H. Rieder, H. Schmid, H. Stossel, J. Appl. Phys. **37**, 1404 (1966)
43. T. Kimura, T. Goto, H. Shintani, K. Ishizaka, T. Arima, Y. Tokura, Nature **426**, 55 (2003)
44. N. Hur, S. Park, P.A. Sharma, J.S. Ahn, S. Guha, S.-W. Cheong, Nature **429**, 392 (2004)