WILLIAM & MARY
CHARTERED 1693

Dissertations, Theses, and Masters Projects   Theses, Dissertations, & Master Projects

2000

Structural instabilities in strontium titanate from first-principles calculations

Christopher A. LaSota
College of William & Mary - Arts & Sciences

Follow this and additional works at: https://scholarworks.wm.edu/etd

Part of the Condensed Matter Physics Commons

Recommended Citation

INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI®

Structural Instabilities in SrTiO$_3$ from First-Principles Calculations

———

A Dissertation

Presented to

The Faculty of the Department of Physics

The College of William and Mary

In Partial Fulfillment

Of the Requirements for the Degree of

Doctor of Philosophy

———

By

Christopher A. LaSota

1999
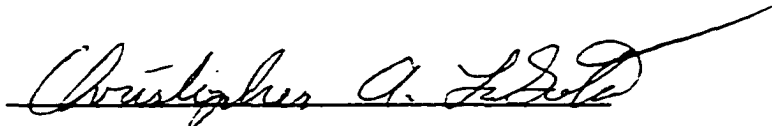
APPROVAL SHEET

This dissertation is submitted in partial fulfillment
of the requirements for the degree of

Doctor of Philosophy.

Christopher A. LaSota

Approved, August 1999

Henry Krakauer

Kenneth G. Petzinger

William J. Kossler

Shiwei Zhang

Gerald H. Johnson
Geology Department

ii

To my parents

Contents

iv

Acknowledgments

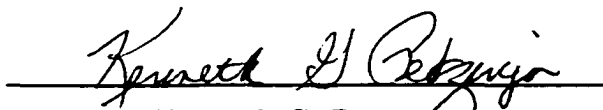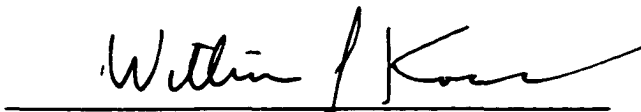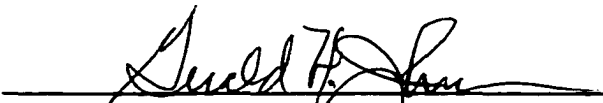These past years at William and Mary have been some of the most rewarding that I've ever known, and this is due in no small part to all of the great people I've met here. Coming from a relatively cloistered background, graduate school was an opportunity to develop my social skills and well as my scientific ones. To all of you not mentioned below, thanks for a memorable time in my life.

I have had the fortunate pleasure of working for Dr. Henry Krakauer, who, in my opinion, is an excellent role model for any aspiring academic. He is an active researcher who works as hard as his students, if not harder. He has made a consistent effort to make sure that his research and that of his students is relevant in the current scientific environment.

I would also like to thank my committee members, Profs. Kenneth G. Petzinger, William J. Kossler, Shiwei Zhang, and Gerald H. Johnson, who have taken the time to proofread this thesis, and provide their suggestions and comments.

In addition, I consider myself lucky to have worked with Dr. Cheng-Zhang Wang, who has also served as my second mentor and friend during the last five years. I'm sure he may have gotten annoyed with me from time to time, as I constantly bugged him with questions about theory and method. Without his help, I would have struggled a lot more than I did. Fellow researchers, Mark Wensell, Rici Yu, Matt Coffey, and Chree Haas were also generous with their time, especially when it came to computing issues.

It would be unforgiveable not to recognize the tremendous importance of the roles played by our secretaries, Paula Perry, Sylvia Stout, and Diane Fannin, who keep everything running smoothly, and see to it that we all stay on track and out of academic and financial trouble.

Among others, I begin with my fellow physics graduate students. I'll never forget the days spent living in the 'hood with Marco Brown, Justin McIntyre, and

Dave Viel. Which of us can forget the cheese-bread and cookies, the glop dinners, and the yap-dog next door. And man, those parties were something else. I'll miss the long hours of fat-chewing with them as well as with Jai Diggs, Darren Wah, Eric Allman, and Jon Curley. I received lots of good exercise playing IM sports with Marco, Al Greer, Jon Goetz, Jack Tucek, Scott Walton, Brian Peko, and Mike West, Among the lesser-funded disciplines, plenty of good times were also to be had with the likes of Matt and Stephanie Cohen, Joe Rainer, and hoards of other American Studies students. I've also met some great undergraduates during this era. Good friends like Natalie Uschner and Molly Schaefer helped me escape from the bittersweet routine of grad life. A few more names deserving mention are Shannon Davis, Steve Wandrai, "Bumbleberry" Bob Gaines, Sarah Pollara, Brinda Krishnan, and Anne Lipke.

Most importantly, I'd like to thank my parents. They each sacrificed their own dreams to raise their children, and have only just recently achieved some of their original goals. This thesis is dedicated to them.

List of Tables

.

List of Figures

x

Abstract

For some time now, first-principles calculation methods have proven to be an effective tool for investigating the physics of condensed matter systems. The additional use of density functional theory (DFT) and the local density approximation (LDA) has permitted even complex materials to be studied on desktop workstations with remarkable success. The incorporation of linear response theory into these methods has extended their power, allowing investigation of important dynamical properties.

Contained within the following pages are the results of a first-principles study of $SrTiO_3$. This transition metal oxide is often grouped with ferroelectric materials, due to its similar composition and perovskite structure. Although it behaves as if it were to become ferroelectric, it fails to do so, even at the lowest temperatures.

Using the LAPW method for bulk materials, the ground-state equilibrium properties for the cubic phase were found. Additional linear response calculations produced the phonon frequencies throughout the Brillouin zone. Imaginary values for these frequencies revealed two distinct regions of reciprocal space corresponding to structural instabilities of the ferroelectric (FE) and antiferrodistortive (AFD) types. A cell-doubling AFD transition to tetragonal structure is observed experimentally, so subsequent calculations were continued in this phase. Total energy calculations were performed for both FE and AFD distortions in this new phase, and it was found that the AFD instability is enhanced with decreasing lattice parameter, while the FE instability is diminished. Furthermore, these calculations suggest that this material is marginally stable against FE distortions, even at the 105 K volume.

Structural Instabilities in SrTiO$_3$ from

First-Principles Calculations

Chapter 1

Introduction

1.1 Ferroelectrics

When placed in an external electric field, a typical dielectric crystal develops a macroscopic polarization P, whose magnitude depends upon the strength of the applied field. For weak to moderate values of E, the relationship between the two is linear. With increased field strength, the polarization begins to saturate and eventually reaches its highest value just before breakdown. Usually, when the external field is removed, the polarization also disappears. However, a class of materials known as *pyroelectrics* can retain a residual polarization after the external electric field is removed, as long as their temperature is below some critical value T_c. *Ferroelectric* materials are a subset of pyroelectrics and are characterized by the fact that the direction of the remnant polarization may be reversed by the application of an external electric field of sufficient strength. A distinguishing sign of ferroelectric behavior is the appearance of a hysteresis loop on a plot of the polarization verses electric field

2
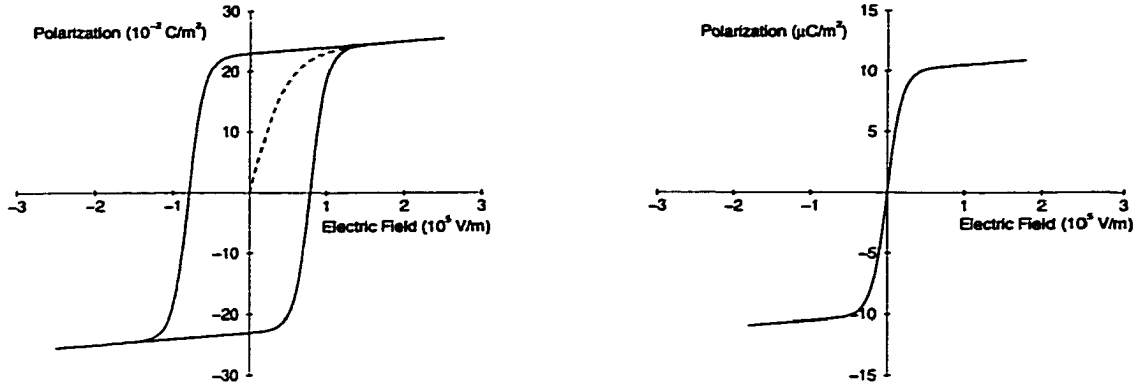
Figure 1.1: Ferroelectric materials exhibit a characteristic hysteresis loop (left), which is absent for ordinary dielectrics (right).

strength (Fig. 1.1). The magnitude of the remnant polarization is dependent upon the crystal's temperature, decreasing as the temperature rises, and disappearing altogether at and above T_c. In analogy to the temperature-dependent behavior of magnetic materials, this transition temperature is known as the *Curie temperature*, above which the material is said to be in its *paraelectric* state.

Even when no external field is present, a macroscopic *spontaneous* polarization will develop as the temperature drops below T_c; however, it is usually negligible or small. Depending upon the symmetry of the paraelectric phase, more than one polarization axis may develop during the transition, leading to the formation of domains. The polarization within each domain is homogeneous, but is randomly oriented along one of the possible polarization directions. If an electric field is applied after the transition to the ferroelectric state, many of these domains will grow or combine through the motion of domain walls, producing a nonzero macroscopic polarization which remains even after the external field is removed, as shown in Fig. 1.2. In this dissertation, we are chiefly concerned with the behavior of monodomain single crystals,

(a) $T > T_c$

$\bar{P}_{net} = 0$

(b) $T < T_c$

$\bar{P}_{net} = 0$

(c) $T < T_c$

$\longleftarrow \bar{P}_{net}$
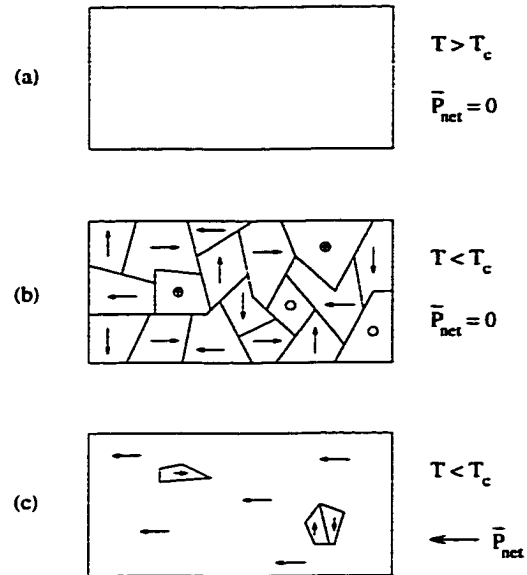
Figure 1.2: Development of spontaneous polarization in a multidomain ferroelectric crystal as T drops below T_c. (a) paraelectric phase; (b) ferroelectric phase with no net polarization; (c) ferroelectric phase after exposure to an external electric field.

and this should be assumed for the remainder of the work.

In order for a crystal to maintain a macroscopic polarization, the centers of positive and negative charge within it cannot overlap. As a consequence, ferroelectric crystals do not possess a center of symmetry below T_c. Because the paraelectric phase lacks a dipole moment, it is highly likely that the ferroelectric transition is a structural one. The reversibility of the remnant polarization also suggests that the crystal passes through the non-polar paraelectric structure on its way to the alternative ferroelectric state.

Ferroelectric structural transitions have traditionally been classified as belonging to two distinct types. When the transition is brought on by the individual ordering of ions or molecules, it is called an *order-disorder* transition. However, if a whole sublattice of ions becomes displaced relative to another sublattice, then the transition is referred to as a *displacive* transition. Examples of order-disorder transitions are seen in materials such as KH_2PO_4, wherein the location of the proton in each hydrogen bond becomes asymmetric below the Curie temperature (123 K). Perhaps the most famous displacive transition occurs in $BaTiO_3$. The high-temperature paraelectric phase of $BaTiO_3$ has the cubic perovskite structure shown in Fig. 1.3. When the temperature drops below 393 K, a shift occurs in the metal ion sublattices with respect to the oxygen sublattice, leading to the formation of a dipole in each unit cell. Since the high temperature phase is cubic, the spontaneous polarization can appear along one of three identical axes, making the new crystal symmetry tetragonal with the polarization parallel to the c-axis [001]. $BaTiO_3$ undergoes two more ferroelectric transitions to orthorhombic (278 K) and rhombohedral (183 K) structure as new
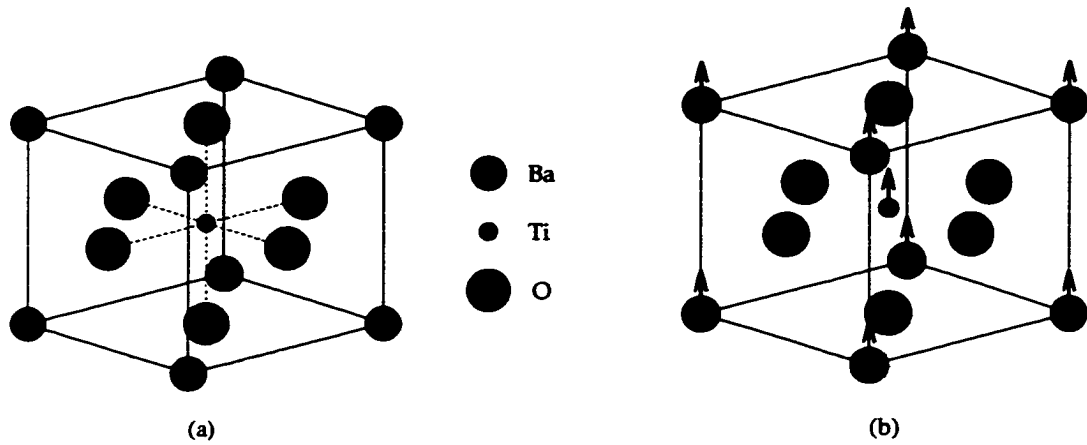
Figure 1.3: (a) The cubic perovskite structure of $BaTiO_3$. (b) Below T_c the Ba^{2+} and Ti^{4+} ions are slightly displaced relative to the O^{2-} ions giving the cell a dipole moment and a tetragonal structure.
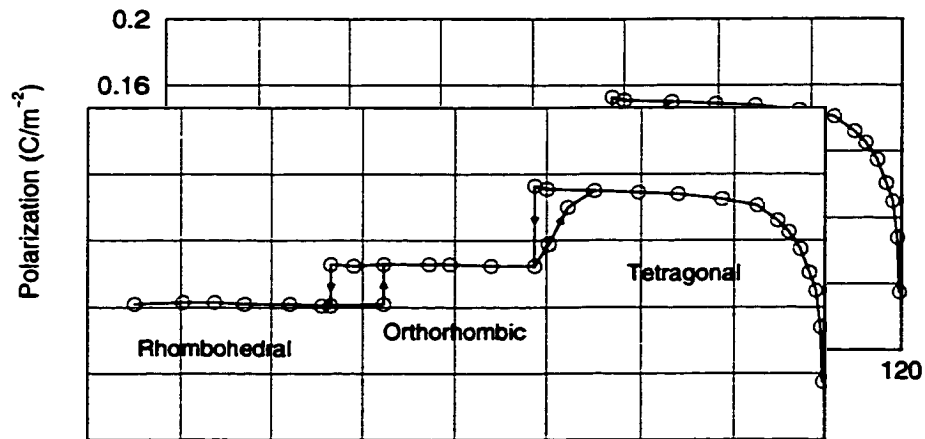
Figure 1.4: The spontaneous polarization in $BaTiO_3$ appears first along the [001] direction, and then along the [110] and [111] directions as the temperature is lowered. (After W. J. Merz.[77])

components are added to the polarization, appearing from the addition of displacive ordering along the other two possible polarization axes successively (Fig. 1.4). Similar behavior is seen in other perovskites as well. $KNbO_3$ undergoes the same sequence of transitions as $BaTiO_3$, whereas $PbTiO_3$ only transforms to a tetragonal structure. $SrTiO_3$, which is the subject of the present work, gives experimental evidence for a possible ferroelectric transition occurring near 0 K, even though it first undergoes a non-ferroelectric structural transition at 105 K.

Originally, it was thought that the relative simplicity of the 5-atom perovskite structure would quickly lead to an understanding of the ferroelectric transition. However, it is now known that the delicate interplay between the short-range repulsive and the long-range attractive forces at work can result in a complex array of observed phase transitions.[68] Even so, the perovskite structure serves as a good framework for a discussion of the mechanism behind displacive transitions. Although the next section discusses the general thermodynamic principles involved in ferroelectric transitions, the focus will be on how they are related to the behavior seen in the perovskites.

1.2 Thermodynamics of ferroelectric transitions

Prior to the onset of a ferroelectric transition, the low-frequency dielectric constant often rises dramatically. This behavior is consistent with a relaxation of the restoring forces acting on the ions due to a reduction in the temperature-dependent components of these forces. Consequently, ions will have a greater response to external electric fields. Above the transition temperature T_c, experimental measurements of the dielectric constant at different temperatures usually fit well to the empirical

form

$$\epsilon = \frac{\xi}{T - T_c}. \qquad (1.1)$$

Here, ξ is a constant factor determined by the specific material under investigation. For $BaTiO_3$ and $SrTiO_3$ (and their alloys), the fit is a remarkably good one.[89] The reason for this can be understood from an analysis of the thermodynamics of the transition via Landau's theory of phase transitions.[66] First, we conveniently expand the Landau free energy density F in one dimension in terms of the polarization

$$F(E, T; P) = -EP + c_0 + \frac{1}{2}\, c_2 P^2 + \frac{1}{4}\, c_4 P^4 + \frac{1}{6}\, c_6 P^6 + \cdots. \qquad (1.2)$$

When E=0, odd powers of P do not appear, since we are assuming that the unpolarized phase of the crystal has a center of inversion symmetry. The value of the polarization corresponding to a minimum in F will satisfy

$$\frac{\partial F}{\partial P} = 0$$
$$= -E + c_2 P + c_4 P^3 + c_6 P^5 + \cdots. \qquad (1.3)$$

In the paraelectric phase, the crystal behaves similarly to a normal dielectric, indicating that c_2 and higher-order coefficients are positive quantities above the transition temperature. Once the temperature drops below T_c, a spontaneous polarization P_s will appear, even if E=0. This would be expected if the value of c_2 were to become negative. Thus, let us assume that c_2 changes sign as the temperature passes through some value T_0, which we shall see may be equal to or lower than the transition temperature T_c. Near T_0, we take

$$c_2 = \gamma(T - T_0), \qquad (1.4)$$

with γ as a positive constant. Variations in c_2 with temperature are not unphysical, and can be accounted for by thermal expansion and other anharmonic lattice effects.

Although a negative value of c_2 can account for the appearance of a spontaneous polarization, it alone does not explain the temperature dependence of P_s near the transition. For this we must examine the sign of c_4. When $c_4 > 0$, we can ignore higher order terms, and solve for P_s in zero applied field via

$$c_4 P_s^3 + \gamma(T - T_0)P_s = 0, \tag{1.5}$$

resulting in either

$$P_s = 0 \qquad \text{or} \qquad |P_s| = (\gamma/c_4)^{\frac{1}{2}}(T_0 - T)^{\frac{1}{2}}.$$

Because the polarization drops to zero in a continuous fashion as T approaches T_0 from below, such a transition is known as a *second-order transition*. Since the free energy densities of both the paraelectric and ferroelectric phases are equal at the transition temperature, this implies that $T_0 = T_c$ for second-order transitions. If c_4 is negative rather than positive, then the spontaneous polarization in zero applied field will satisfy

$$c_6 P_s^5 - |c_4|P_s^3 + \gamma(T - T_0)P_s = 0, \tag{1.6}$$

meaning that either P_s is zero, or

$$c_6 P^4 - |c_4|P^2 + \gamma(T - T_0) = 0. \tag{1.7}$$

Here we have explicitly assumed $c_6 > 0$ to restrain the free energy density from becoming asymptotically negative. Consider the situation now where $T = T_0$. When
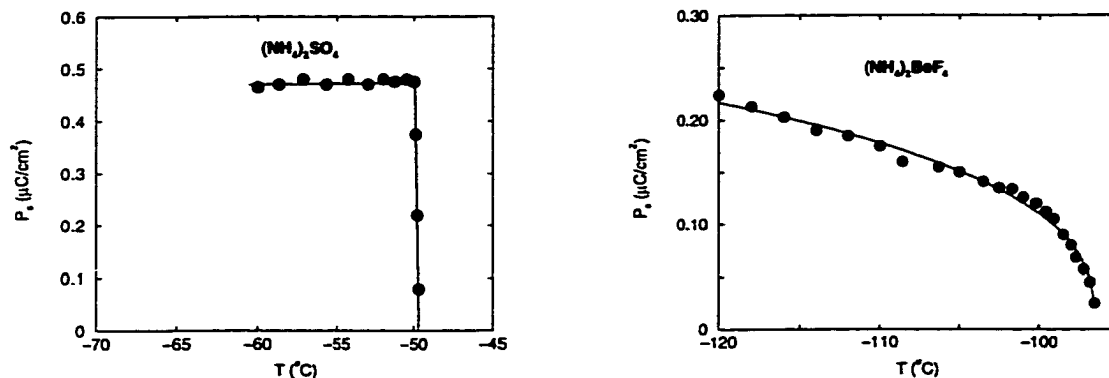
Figure 1.5: The behavior of the spontaneous polarization in first- and second-order ferroelectric transitions in chemically similar ammonium salts. After S. Hoshino *et al.*[52]

this occurs, Eq. (1.7) has solutions

$$P_s = 0 \qquad \text{or} \qquad P_s = \sqrt{\frac{|c_4|}{c_6}}.$$

If the free energy density is evaluated (to sixth order in the polarization) at both of these roots, one finds that the non-zero root yields a lower value. Since the phase transition takes place when the free energies of the polarized and unpolarized states are equal, it must be true that the actual transition temperature T_c is higher than T_0. Because of this, when $T=T_c$, the curvature of the free energy density function is positive about $P=0$, meaning that when the transition takes place, a non-zero polarization appears suddenly, rather than increasing continuously from zero value. For this reason, this type of transition is called a *first-order transition*. Both types of transition are illustrated in Fig. 1.5.

In either case, we can find the dielectric constant for temperatures above T_c

using

$$\epsilon = 1 + \frac{4\pi P}{E}. \tag{1.8}$$

In the paraelectric phase, for weak or moderate applied electric fields, we can ignore the nonlinear terms of Eq. (1.3) to give

$$\epsilon \approx 1 + \frac{4\pi}{\gamma(T - T_0)} \quad \text{for} \quad T > T_c. \tag{1.9}$$

Since the second-order term on the right is usually much greater than unity, the empirical form of Eq. (1.1) results.

1.3 The soft-mode concept

If a crystal is stable against small deformations, then as a consequence, all of its normal modes will have non-imaginary values for their frequencies.[10] Consequently, Cochran and Anderson took the approach of treating the ferroelectric transition in terms of lattice dynamics.[21, 3] They proposed that the ferroelectric transition in the perovskites and other ionic crystals could be explained by a strong temperature dependence of a particular normal mode of vibration. As the temperature approaches T_c, this normal mode "softens", i.e. its frequency decreases to zero, making the crystal unstable. The new atomic positions in the ferroelectric phase represent the frozen-in displacements of the unstable phonon. Since these displacements lead to a macroscopic spontaneous polarization, the soft phonon must necessarily be a polar mode of long wavelength ($\vec{q} \to 0$). In addition, for diatomic or cubic perovskite ferroelectrics, the rise in the low-frequency dielectric constant near T_c is predicted by

the generalized Lyddane-Sachs-Teller (LST) [71, 21, 22] relation

$$\frac{\epsilon_0}{\epsilon_\infty} = \prod_{i=2}^{n} \frac{(\omega_L^2)_i}{(\omega_T^2)_i},$$

(1.10)

where $(\omega_T)_i$ and $(\omega_L)_i$ are the frequencies of the transverse and longitudinal optic phonons at the Brillouin zone center ($\vec{q} = 0$), n is the number of atoms in the primitive cell, and the acoustic modes (i=1) have been excluded. The ϵ_0 and ϵ_∞ terms represent the dielectric constants of the material when subjected to low-frequency and high-frequency EM fields, respectively. Eq. (1.10) may also be derived from the Kramers-Kronig relations [55] where $(\omega_L)_i$ and $(\omega_T)_i$ are the zeros and the poles of the dielectric function $\epsilon(\omega)$, respectively.[9] Since ω_L is typically independent of temperature, the presence of a soft transverse phonon will result in a dramatic increase in ϵ_0. In fact, if the behavior of the dielectric constant follows a Curie-Weiss form, then we should expect that

$$\omega_T^2 \propto (T - T_c).$$

(1.11)

In Fig. 1.6, we see that experimental measurments performed in $SrTiO_3$ agree very well with the conclusions of the lattice dynamical approach. This and other experimental results have since established the validity of the soft phonon concept for perovskite and diatomic ferroelectrics.

1.4   Brief History of Ferroelectricity

Ferroelectricity was originally discovered in 1920 in Rochelle salt ($NaKC_6H_4O_6$ · $4H_2O$) [107] and later in KH_2PO_4 and related salts [12, 13, 101]. The ferroelectric
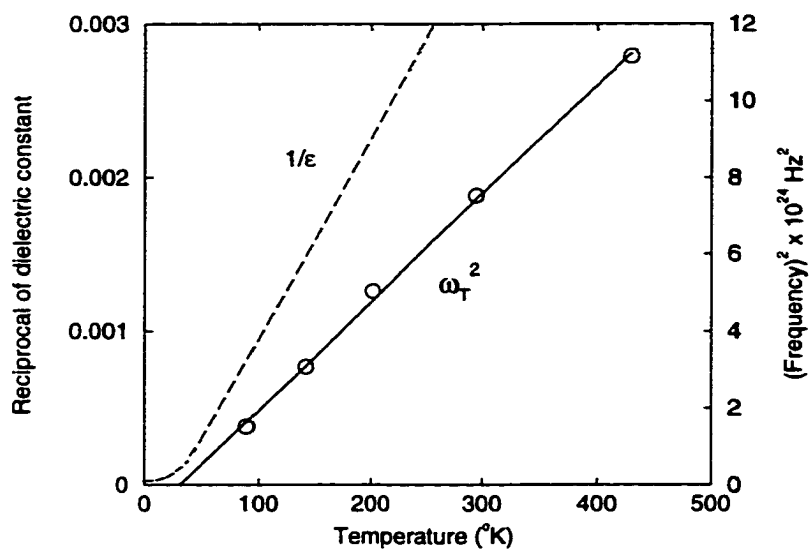
Figure 1.6: Experimental measurements by Cowley of the temperature dependence of the zone-center transverse optic mode in $SrTiO_3$.[28] The broken line is the reciprocal of the dielectric constant measurements of Mitsui and Westphal[79].

properties of $BaTiO_3$ were revealed near the end of World War II, and since that time it has become one of the most widely studied materials, helping to establish the modern understanding of ferroelectricity.[68] The relative simplicity of its perovskite structure led to searches for ferroelectricity in other transition metal oxides with the same ABO_3 formula unit. These efforts were rewarded by the discovery of a whole series of ferroelectric titanates ($PbTiO_3$, $SrTiO_3$), tantalates ($KTaO_3$, $NaTaO_3$, $LiTaO_3$, $RbTaO_3$), and niobates ($KNbO_3$, $LiNbO_3$), as well as mixed alloys ($Pb(Zr_xTi_{1-x})O_3$, also known as PZT), each with a single ferroelectric phase transition or a sequence of multiple transitions at successively lower temperatures.

The simple 5-atom formula unit permitted a microscopic analysis of the transition that later resulted in the lattice dynamical description mentioned above. The origin of the soft-mode instability is apparently caused by the anisotropic polarizability of the oxygen atoms, created by *dynamical* hybridization between the oxygen p-states and the transition metal d-states.[78, 8, 26, 14] Contemporary with these successful descriptions were the results of diffuse X-ray scattering in $BaTiO_3$ and $KNbO_3$ by Comes *et al*, which suggest that the local atomic structure might not be the same as the apparent structure seen in elastic X-ray and neutron scattering experiments.[27] The picture suggested by these results is one where short-range chains of similarly distorted unit cells exist *prior* to the actual transition, but only become macroscopically correlated below the transition temperature, similar to an order-disorder type of transition. Thus, although the results of the past few decades have greatly contributed to our knowledge of the ferroelectric transition in the perovskites, it appears that there is still more to be learned.

Much of the recent progress in our theoretical understanding of ferroelectric behavior has been largely due to a rapid increase in the computational power and availability of hardware; so much so that relevant research efforts may now be carried out on a desktop platform. Moreover, advances in the computational methods themselves have resulted in high-quality calculations. One such method, the Linear Augmented Plane-Wave (LAPW) method has emerged as the standard against which other theoretical calculations are compared. In the present work, this method is used to predict both the ground state and first-order dynamical properties of $SrTiO_3$. Before delving into the results of this investigation, the background theoretical basis and principal assumptions behind the LAPW and most other modern calculation methods will be presented in the next chapter.

Chapter 2

First-principles calculations in condensed matter

2.1 The Born-Oppenheimer approximation

Any attempt to accurately describe the physics of a condensed matter system must necessarily begin with the Schrödinger equation. Most often, we are concerned only with time-independent interactions, so that the ground-state of a system of nuclei and electrons can be found using the time-independent form of the Schrödinger equation

$$H(\vec{r}, \vec{R}) \, \Phi(\vec{r}, \vec{R}) \; = \; E \; \Phi(\vec{r}, \vec{R}), \tag{2.1}$$

where $\vec{r} = \{\vec{r}_1, \vec{r}_2, \ldots \}$ and $\vec{R} = \{\vec{R}_1, \vec{R}_2, \ldots \}$ are the positions of the N_e electrons and the N_a atomic nuclei, respectively. Φ is the many-body wavefunction and depends upon the positions of *all* of the particles. E is the total energy, and H is the

16

Hamiltonian operator, given by

$$H(\vec{r}, \vec{R}) = T_n(\vec{R}) + T_e(\vec{r}) + U_{n-n}(\vec{R}) + U_{e-e}(\vec{r}) + U_{e-n}(\vec{r}, \vec{R}). \tag{2.2}$$

The kinetic energy operators for the nuclei and electrons are, respectively,

$$T_n(\vec{R}) = -\sum_a \frac{1}{2M_a} \nabla^2_{\vec{R}_a}, \tag{2.3}$$

$$T_e(\vec{r}) = -\sum_i \frac{1}{2} \nabla^2_{\vec{r}_i}. \tag{2.4}$$

Here we have used atomic units, where $m_e=1$, $e=1$, and $\hbar=1$. The other terms, U_{n-n}, U_{e-e}, and U_{e-n}, are the electrostatic potential energy operators for the nuclei, the electrons, and the interactions between them:

$$U_{n-n}(\vec{R}) = \frac{1}{2} \sum_{a \neq a'} \frac{Z_a Z_{a'}}{|\vec{R}_a - \vec{R}_{a'}|}, \tag{2.5}$$

$$U_{e-e}(\vec{r}) = \frac{1}{2} \sum_{i \neq j} \frac{1}{|\vec{r}_i - \vec{r}_j|}, \tag{2.6}$$

$$U_{e-n}(\vec{r}, \vec{R}) = -\sum_{i,a} \frac{Z_a}{|\vec{r}_i - \vec{R}_a|}. \tag{2.7}$$

Here, both the electrons and the nuclei are treated as point charges. Due to the fact that the mass of a nucleus is much greater than that of an electron, the motion of the former is much slower than that of the latter, and T_n can be treated as a perturbation. Born and Huang [10] have shown that using this mass ratio in a perturbative expansion, one can write the many-body wavefunction as a product of an electronic wavefunction and a nuclear wavefunction. The physical interpretation of this is that during the nuclear motion, the electrons behave as if the nuclei were fixed. This is what is known as the *adiabatic* or *Born-Oppenheimer approximation*. With fixed

nuclei, U_{n-n} reduces to a scalar, and all that remains is to solve the corresponding *electronic* Schrödinger equation:

$$H_{el}(\vec{r}, \vec{R})\Psi(\vec{r}) = E_{el}\Psi(\vec{r}). \tag{2.8}$$

Here, H_{el} is the electronic Hamiltonian given by

$$H_{el}(\vec{r}, \vec{R}) = T_e(\vec{r}) + U_{e-e}(\vec{r}) + U_{e-n}(\vec{r}, \vec{R}), \tag{2.9}$$

and

$$E_{el} = K_e + E_{e-e} + E_{e-n} \tag{2.10}$$

is the electronic energy, composed of the total kinetic energy of the electrons and the electrostatic interaction energies. The many-body electronic wavefunction, $\Psi(\vec{r})$, is quite complicated and is the subject of the next section. As we shall see later, minimizing E_{el} with respect to $\Psi(\vec{r})$, under the constraint that it be antisymmetric and normalized, is equivalent to solving Eq. (2.8) for the ground state. Once the ground-state value of E_{el} is determined, the total energy of the system, $E_{el}+U_{n-n}$ can be evaluated. The nuclear positions can then be shifted, and a new total energy found. In this way, we can perform structural optimizations and determine linear response coefficients from first-principles.

2.2   The many-body electronic wavefunction

The physical interpretation of the many-body electronic wavefunction is that $|\Psi|^2$ is a probability density function in the following sense:

$$|\Psi|^2 \, d\vec{r}^N = \Psi^*(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N) \, \Psi(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N) \, d\vec{r}_1 d\vec{r}_2 \dots d\vec{r}_N$$

= the probability of finding electron 1 between $\vec{r}_1$ and   (2.11)

$\vec{r}_1 + d\vec{r}_1$;   electron 2 between $\vec{r}_2$ and $\vec{r}_2 + d\vec{r}_2$;   and so on.

Here, $N = N_e$, the number of electrons. Thus, $|\Psi|^2$ represents the probability density of finding the system in a particular configuration. Note that although we have not accounted for the spin variables, the concepts here can be easily extended to include them. The appropriate boundary conditions must be applied during the solution of Eq. (2.8). $|\Psi|^2$ must be periodic when dealing with an infinite crystal; or it must decay to zero at infinite distance for isolated atoms and molecules. Also, because the electrons are fermions, Ψ must be antisymmetric under the exchange of any pair:

$$\Psi(\dots, \vec{r}_i, \dots, \vec{r}_j, \dots) = -\Psi(\dots, \vec{r}_j, \dots, \vec{r}_i, \dots).$$ (2.12)

In addition, since the system will be found in *some* configuration, Ψ will be normalized:

$$\int |\Psi(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N)|^2 \, d\vec{r}_1 d\vec{r}_2 \dots d\vec{r}_N = 1.$$ (2.13)

The resulting solutions of Eq. (2.8) are the eigenfunctions Ψ_k, with corresponding eigenvalues E_k. Since this set is complete, we can also assert an orthogonalization condition:

$$\int \Psi_k^* \, \Psi_l \, d\vec{r} \equiv \langle \Psi_k | \Psi_l \rangle = \delta_{kl}.$$ (2.14)

The solution we are looking for is the ground-state wavefunction, Ψ_0, having eigenvalue E_0.

2.2.1 Density matrices

Even more general than the probability density is the *density matrix*, whose elements are defined as

$$\gamma_N(\vec{r}_1', \vec{r}_2', \ldots, \vec{r}_N'; \vec{r}_1, \vec{r}_2, \ldots, \vec{r}_N) \equiv \Psi(\vec{r}_1', \vec{r}_2', \ldots, \vec{r}_N')\,\Psi^*(\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_N). \quad (2.15)$$

If we set $\vec{r}_i = \vec{r}_i'$ for all i, then we get a diagonal element of this matrix, which is the original probability density shown in Eq. (2.11). This alternative formulation can be thought of as a coordinate representation of the density *operator*,

$$\hat{\gamma}_N = |\Psi\rangle\langle\Psi|, \quad (2.16)$$

because

$$(\vec{r}_1', \vec{r}_2', \ldots, \vec{r}_N'|\hat{\gamma}_N|\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_N) = (\vec{r}_1', \vec{r}_2', \ldots, \vec{r}_N'|\Psi\rangle\langle\Psi|\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_N)$$
$$= \Psi(\vec{r}_1', \vec{r}_2', \ldots, \vec{r}_N')\Psi^*(\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_N). \quad (2.17)$$

The normalization condition can now be expressed as a trace:

$$\mathrm{tr}\,(\hat{\gamma}_N) = \int \Psi(\vec{r}^N)\Psi^*(\vec{r}^N)\,d\vec{r}^N = 1. \quad (2.18)$$

The density operator can be used to evaluate the expectation values of observables. The expectation value of a Hermitian linear operator $\hat{A}$, is given by

$$\langle \hat{A} \rangle \equiv \frac{\langle\Psi|\,\hat{A}\,|\Psi\rangle}{\langle\Psi|\Psi\rangle} = \frac{\int \Psi^*\hat{A}\,\Psi\,d\vec{r}}{\int \Psi^*\Psi\,d\vec{r}}. \quad (2.19)$$

If Ψ is normalized, then $\langle \hat{A} \rangle$ can be written in the position representation:

$$\langle \hat{A} \rangle = \langle \Psi | \hat{A} | \Psi \rangle \tag{2.20}$$

$$= \int \langle \Psi | \vec{r} \rangle \langle \vec{r} | \hat{A} | \vec{r}' \rangle \langle \vec{r}' | \Psi \rangle \, d\vec{r}' d\vec{r} \tag{2.21}$$

$$= \int \Psi^*(\vec{r}) A(\vec{r}, \vec{r}') \Psi(\vec{r}') \, d\vec{r}' d\vec{r} \tag{2.22}$$

$$= \text{tr} \, (\hat{\gamma}_N \hat{A}) = \text{tr} \, (\hat{A} \hat{\gamma}_N). \tag{2.23}$$

Here, $\vec{r} = \{\vec{r}_1, \vec{r}_2, \ldots\}$ and $\vec{r}' = \{\vec{r}_1', \vec{r}_2', \ldots\}$. The matrix elements of $\hat{A}$ in Eq. (2.22) indicate that it is possible for an operator to be *nonlocal*. An operator is only *local* if

$$A(\vec{r}_1, \vec{r}_1') = A(\vec{r}_1) \delta(\vec{r}_1' - \vec{r}_1). \tag{2.24}$$

2.2.2   Reduced density matrices

Because we will be dealing with symmetric one-electron and two-electron operators, the expectation value given by Eq. (2.23) can be simplified by integrations over the position variables which do not appear in these operators. This will result in a *reduced density matrix*. Eq. (2.15) is called the Nth *order density matrix*. The *reduced density matrix of order m* is given by

$$\gamma_m(\vec{r}_1', \vec{r}_2', \ldots, \vec{r}_m'; \vec{r}_1, \vec{r}_2, \ldots, \vec{r}_m)$$

$$= \binom{N}{m} \int \gamma_N(\vec{r}_1', \ldots, \vec{r}_m', \vec{r}_{m+1}, \ldots, \vec{r}_N; \vec{r}_1, \ldots, \vec{r}_m, \vec{r}_{m+1}, \ldots, \vec{r}_N) \, d\vec{r}_{m+1} \ldots d\vec{r}_N,$$

$$\tag{2.25}$$

where $\binom{N}{m}$ is a binomial coefficent. The first-order density matrix γ_1 is given by

$$\gamma_1 (\vec{r}_1'; \vec{r}_1) = N \int \Psi(\vec{r}_1', \vec{r}_2, \ldots, \vec{r}_N) \, \Psi^*(\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_N) \, d\vec{r}_2 \ldots d\vec{r}_N, \tag{2.26}$$

which normalizes to the total number of electrons:

$$\text{tr}\left(\gamma_1(\vec{r}_1';\vec{r}_1)\right) = \int \gamma_1(\vec{r}_1;\vec{r}_1)\,d\vec{r}_1 = N. \tag{2.27}$$

Thus, the diagonal part of the first-order density matrix is just the ordinary electron density, $n(\vec{r}_1)$. Similarly, the second-order density matrix γ_2, which is written

$$\gamma_2\left(\vec{r}_1',\vec{r}_2';\vec{r}_1,\vec{r}_2\right)$$
$$= \frac{N(N-1)}{2} \int \Psi(\vec{r}_1',\vec{r}_2',\vec{r}_3,\dots,\vec{r}_N)\,\Psi^*(\vec{r}_1,\vec{r}_2,\vec{r}_3,\dots,\vec{r}_N)\,d\vec{r}_3\dots d\vec{r}_N, \tag{2.28}$$

normalizes to the number of electron *pairs*:

$$\text{tr}\left(\gamma_2\left(\vec{r}_1',\vec{r}_2';\vec{r}_1,\vec{r}_2\right)\right) = \iint \gamma_2\left(\vec{r}_1,\vec{r}_2;\vec{r}_1,\vec{r}_2\right)\,d\vec{r}_1 d\vec{r}_2 = \frac{N(N-1)}{2}. \tag{2.29}$$

Here, the diagonal part of the second-order density matrix $n(\vec{r}_1,\vec{r}_2)$ is the density of *pairs* of electrons, one at $\vec{r}_1$, and the other at $\vec{r}_2$. Both the first and second order densities are positive semidefinite and Hermitian, and obey antisymmetry requirements. They are related to each other by

$$n(\vec{r}_1) = \frac{2}{N-1} \int n(\vec{r}_1,\vec{r}_2)d\vec{r}_2. \tag{2.30}$$

It is now possible to evaluate the expectation values of the components of the electronic Hamiltonian using these results.

2.3 The electronic energy

The expectation value of the electronic Hamiltonian is dependent upon the wavefunction, and for this reason, the electronic energy is a *functional* of Ψ:

$$E_{el}[\Psi] = \langle H_{el}\rangle \equiv \frac{\langle \Psi|\,H_{el}\,|\Psi\rangle}{\langle \Psi|\Psi\rangle}. \tag{2.31}$$

To determine $E_{el}[\Psi]$, it is necessary to evaluate the expectation value of the individual contributing terms. Since we have imposed the normalization condition, these terms become:

$$K_e[\Psi] = \langle \Psi | T_e | \Psi \rangle = -\sum_{i=1}^{N} \langle \Psi | \frac{1}{2} \nabla_{\vec{r}_i}^2 | \Psi \rangle, \tag{2.32}$$

$$E_{e-e}[\Psi] = \langle \Psi | U_{e-e} | \Psi \rangle = \frac{1}{2} \sum_{i \neq j} \langle \Psi | \frac{1}{|\vec{r}_i - \vec{r}_j|} | \Psi \rangle, \tag{2.33}$$

$$E_{e-n}[\Psi] = \langle \Psi | U_{e-n} | \Psi \rangle = \sum_{i=1}^{N} \langle \Psi | v_{ext}(\vec{r}_i) | \Psi \rangle. \tag{2.34}$$

Here, $v_{ext}(\vec{r}_i)$ represents the external potential energy due to the fixed nuclei, and is therefore only explicitly dependent on the positions of the electrons. Each of these quantities can be evaluated using the reduced density matrix formalism. Both the kinetic energy and the ion-electron electrostatic energy are local, single-particle operators. Their contributions to the electronic energy can be written as:

$$\begin{aligned} \langle \Psi | T_e | \Psi \rangle &= \text{tr}\,(T_e \hat{\gamma}_N) \\ &= \int \delta(\vec{r}_1{}' - \vec{r}_1)\left(-\frac{1}{2}\nabla_{\vec{r}_1}^2\right) \gamma_1(\vec{r}_1{}', \vec{r}_1)\, d\vec{r}_1 d\vec{r}_1{}', \end{aligned} \tag{2.35}$$

$$\begin{aligned} \langle \Psi | U_{e-n} | \Psi \rangle &= \text{tr}\,(U_{e-n} \hat{\gamma}_N) \\ &= \int v_{ext}(\vec{r}_1)\, \gamma_1(\vec{r}_1; \vec{r}_1)\, d\vec{r}_1 = \int v_{ext}(\vec{r}_1)\, n(\vec{r}_1)\, d\vec{r}_1. \end{aligned} \tag{2.36}$$

It is important to note that it is not possible to evaluate the kinetic energy unless we know the off-diagonal terms of $\gamma_1(\vec{r}_1{}', \vec{r}_1)$. In contrast, U_{e-n} can be evaluated using only the diagonal terms, which make up the electron density $n(\vec{r}_1)$. The repulsive Coulomb interaction between electrons is a local, two-particle operator, and the

interaction energy can be written using the second-order density matrix:

$$\langle \Psi | U_{e-e} | \Psi \rangle = \frac{1}{2} \int \delta(\vec{r}_1{}' - \vec{r}_1) \, \delta(\vec{r}_2{}' - \vec{r}_2) \, \frac{1}{|\vec{r}_1 - \vec{r}_2|} \, \gamma_2(\vec{r}_1{}', \vec{r}_2{}'; \vec{r}_1, \vec{r}_2) \, d\vec{r}_1{}' d\vec{r}_2{}' d\vec{r}_1 d\vec{r}_2$$

$$= \frac{1}{2} \int \frac{1}{|\vec{r}_1 - \vec{r}_2|} \, \gamma_2(\vec{r}_1, \vec{r}_2; \vec{r}_1, \vec{r}_2) \, d\vec{r}_1 d\vec{r}_2$$

$$= \frac{1}{2} \int \frac{1}{|\vec{r}_1 - \vec{r}_2|} \, n(\vec{r}_1, \vec{r}_2) \, d\vec{r}_1 d\vec{r}_2.$$

$$(2.37)$$

If the electron-electron interaction were classical, one would expect to be able to write its energy as a function of the single-electron density:

$$E_{e-e}^{classical}[n(\vec{r})] = \frac{1}{2} \iint \frac{n(\vec{r}_1) \, n(\vec{r}_2)}{|\vec{r}_1 - \vec{r}_2|} \, d\vec{r}_1 d\vec{r}_2. \qquad (2.38)$$

From this observation, it would seem possible to use the classical form if an extra term is added to represent the non-classical deviations. Let us propose a *pair-correlation function* $w(\vec{r}_1, \vec{r}_2)$ such that

$$n(\vec{r}_1, \vec{r}_2) = \frac{1}{2} n(\vec{r}_1) \, n(\vec{r}_2) \, [1 + w(\vec{r}_1, \vec{r}_2)], \qquad (2.39)$$

which, when inserted into Eq. (2.30), gives

$$n(\vec{r}_1) = \frac{n(\vec{r}_1)}{N - 1} [N + \int n(\vec{r}_2) \, w(\vec{r}_1, \vec{r}_2) \, d\vec{r}_2]. \qquad (2.40)$$

This implies that

$$\int n(\vec{r}_2) \, w(\vec{r}_1, \vec{r}_2) \, d\vec{r}_2 = -1. \qquad (2.41)$$

The integrand above is called the *exchange-correlation hole*, due to the fact that it integrates to the absence of one electron. It can also be written as

$$n_{xc}(\vec{r}_1, \vec{r}_2) = n(\vec{r}_2) \, w(\vec{r}_1, \vec{r}_2). \qquad (2.42)$$

Using $n_{xc}(\vec{r}_1, \vec{r}_2)$, we can separate E_{e-e} into a classical part, known as the *Hartree energy*, and an exchange-correlation part:

$$\langle \Psi | U_{e-e} | \Psi \rangle = E_H[\Psi] + E_{xc}[\Psi], \tag{2.43}$$

where

$$E_H[\Psi] = \frac{1}{2} \int \frac{n(\vec{r}_1)\, n(\vec{r}_2)}{|\vec{r}_1 - \vec{r}_2|}\, d\vec{r}_1 d\vec{r}_2, \tag{2.44}$$

$$E_{xc}[\Psi] = \frac{1}{2} \int \frac{n(\vec{r}_1)\, n_{xc}(\vec{r}_1, \vec{r}_2)}{|\vec{r}_1 - \vec{r}_2|}\, d\vec{r}_1 d\vec{r}_2. \tag{2.45}$$

Here, E_H, is the repulsion energy of the electrons as if they were independent of each other's positions. The exchange-correlation energy is comprised of differences from E_H due to the presence of the exchange-correlation hole. We may now write the entire expression for the electronic energy functional:

$$
\begin{aligned}
E_{el}[\Psi] = \langle \Psi | H_{el} | \Psi \rangle = & \int \delta(\vec{r}_1{}' - \vec{r}_1) \left(-\frac{1}{2} \nabla^2_{\vec{r}_1} \right) \gamma_1(\vec{r}_1{}'; \vec{r}_1)\, d\vec{r}_1 d\vec{r}_1{}' \\
& + \int v_{ext}(\vec{r}_1)\, n(\vec{r}_1)\, d\vec{r}_1 + \frac{1}{2} \int \frac{n(\vec{r}_1)\, n(\vec{r}_2)}{|\vec{r}_1 - \vec{r}_2|}\, d\vec{r}_1 d\vec{r}_2 \\
& + \frac{1}{2} \int \frac{n(\vec{r}_1)\, n_{xc}(\vec{r}_1, \vec{r}_2)}{|\vec{r}_1 - \vec{r}_2|}\, d\vec{r}_1 d\vec{r}_2.
\end{aligned}
\tag{2.46}
$$

The task now is to find the ground state, Ψ_0, which will minimize the value of E_{el} under the constraints of Eqs. 2.12 and 2.13. This is done by making use of the *variational principle*.

2.4 The variational principle

When a system is in an eigenstate Ψ_k, or a linear combination of degenerate eigenstates, every measurement of H_{el} will result in the same eigenvalue E_k. However,

when Ψ is a linear combination of non-degenerate eigenstates, the eigenvalue produced upon measurement is not pre-determined. As a consequence, the expectation value of H_{el}, being a statistical average of the E_k, will always be greater than or equal to the ground-state value. This can be proven by expanding a general wavefunction Ψ in terms of the eigenstates Ψ_k:

$$|\Psi\rangle = \sum_k c_k |\Psi_k\rangle, \tag{2.47}$$

so that the energy becomes

$$E_{el}[\Psi] = \frac{\sum_{k=0} |c_k|^2 E_k}{\sum_{k=0} |c_k|^2} \tag{2.48}$$

$$= \frac{\sum_{k=1} |c_k|^2 (E_k - E_0)}{\sum_{k=0} |c_k|^2} + E_0 \geq E_0, \tag{2.49}$$

where E_0 is the ground-state energy and the orthogonality of the Ψ_k has been used. Obviously, the ground-state energy is reached when $\Psi = c_0\Psi_0$. Thus, minimization of $E_{el}[\Psi]$ with respect to all possible N-electron wavefunctions satisfying the conditions of normalization and antisymmetry will result in the true electronic ground-state:

$$E_0 = E_{el}[\Psi_0] = \min_\Psi E_{el}[\Psi]. \tag{2.50}$$

The task of actually finding this minimum can be achieved using the variational principle, written as

$$\delta E_{el}[\Psi] = 0. \tag{2.51}$$

This is equivalent to requiring that

$$\frac{\delta E_{el}[\Psi]}{\delta \Psi} = 0, \tag{2.52}$$

which is the *Euler-Lagrange* equation. Substituting Eq. (2.31) into Eq. (2.52) and carrying out the functional derivative explicitly yields Eq. (2.8). Consequently, solving Eq. (2.52) for Ψ is equivalent to solving the electronic Schrödinger equation. Eq. (2.52) is also an expression of the fact that all of the eigenstates of H_{el} are extrema of the electronic energy functional. Using the method of Lagrange multipliers, we can minimize $E_{el}[\Psi]$ under the constraint that $\langle \Psi | \Psi \rangle = 1$. This is the same as requiring that

$$\frac{\delta}{\delta \Psi} \left[\langle \Psi | H_{el} | \Psi \rangle - E_{el} \langle \Psi | \Psi \rangle \right] = 0, \qquad (2.53)$$

using E_{el} as the Lagrange multiplier, without imposing the constraint. This equation is then solved for Ψ as a function of E_{el}, and then E_{el} is modified to satisfy normalization. Typically, one constructs a trial wavefunction out of some linear combination of acceptable (antisymmetric, but not necessarily orthonormal) and completely determined basis functions:

$$|\Psi'\rangle = \sum_k c_k |\chi_k\rangle, \qquad (2.54)$$

so that

$$\langle \Psi' | H_{el} | \Psi' \rangle = \sum_{l,k} c_l^* c_k \langle \chi_l | H_{el} | \chi_k \rangle = \sum_{l,k} c_l^* c_k H'_{lk}, \qquad (2.55)$$

$$\langle \Psi' | \Psi' \rangle = \sum_{l,k} c_l^* c_k \langle \chi_l | \chi_k \rangle = \sum_{l,k} c_l^* c_k O'_{lk}, \qquad (2.56)$$

where H'_{lk} and O'_{lk} are the matrix elements of H_{el} and the overlap matrix in the trial basis, respectively. Applying the variational principle as in Eq. (2.53) gives

$$\frac{\delta}{\delta c_l^*} \left[\sum_{l,k} (c_l^* c_k H'_{lk} - E_{el} c_l^* c_k O'_{lk}) \right] = 0, \qquad (2.57)$$

resulting in the standard form for generalized matrix eigenvalue equations

$$\sum_{k} [H'_{lk} - E_{el}O'_{lk}] c_k = 0, \tag{2.58}$$

which have a nontrivial solution when

$$\det (H'_{lk} - E_{el}O'_{lk}) = 0. \tag{2.59}$$

The energy roots and the normalization condition are then used to find the expansion coefficients. This procedure will produce the set of c_k which give the minimum energy possible for a given choice of basis functions. In addition, it is guaranteed that as we use an increasingly flexible trial wavefunction, the error between E_{el} and E_0 is reduced. In fact, even a relatively poor trial wavefunction can give a good estimate of the ground-state energy because if

$$|\Psi'\rangle = |\Psi_0\rangle + |\delta\Psi\rangle, \tag{2.60}$$

then by Eq. (2.49),

$$E_{el}[\Psi'] = E_{el} + O[\delta\Psi^2]. \tag{2.61}$$

However, for the problem of electrons in a solid, Ψ is a complicated many-body wavefunction, and direct minimization as in Eq. (2.53) is impossible. From the results of Eq. (2.30) and Eqs. (2.35-2.37), it is evident that only the second-order density matrix γ_2 is needed to produce a solution. However, using a trial form for γ_2 presents its own difficulties. For any guessed γ_2, it must be derivable from some antisymmetric wavefunction Ψ. This is the *N-representability problem* for γ_2. It is not clear at this time what the conditions are for γ_2 to be representative of an acceptable wavefunction.[24, 25]

2.5 Density functional theory

Once one choses an external potential $v_{ext}(\vec{r})$, whether it be due to the positively charged nuclei or ions, or due to some other external field, the electronic Hamiltonian is fixed. Therefore, for an N-electron system, the ground-state energy and wavefunction are completely determined by the choice of $v_{ext}(\vec{r})$. In a landmark paper by Hohenberg and Kohn[51], it was demonstrated that in the place of N and $v_{ext}(\vec{r})$, one could use the electron density $n(\vec{r})$ as the basic variable. Obviously, N is determined by the density:

$$N = \int n(\vec{r}) \, d\vec{r}. \tag{2.62}$$

The proof that $n(\vec{r})$ comes from only one $v_{ext}(\vec{r})$ to within an additive constant is proven by showing that different external potentials must necessarily lead to different ground-state densities. Since $n(\vec{r})$ can determine N and $v_{ext}(\vec{r})$, it also produces all of the ground-state properties; in particular, the total energy

$$E\left[n(\vec{r})\right] = E_{ext} + \int v_{ext}(\vec{r}) \, n(\vec{r}) \, d\vec{r} + F[n(\vec{r})], \tag{2.63}$$

where E_{ext} is the classical Coulomb energy of the nuclei or ions. Here, $F[n(\vec{r})]$ contains the kinetic and electron interaction energies, and is a *universal functional* of the density in the sense that it has the same form regardless of the external potential. It forms the basis of *density functional theory*, which is an exact theory with no approximations. A second result of Hohenberg and Kohn, which was later extended by Levy,[67] proved that for $n(\vec{r}) \geq 0$ satisfying Eq. (2.62),

$$E_0 = \operatorname*{Min}_{n(\vec{r})} \left\{ F[n(\vec{r})] + \int v_{ext}(\vec{r}) n(\vec{r}) \, d\vec{r} \right\}. \tag{2.64}$$

Actually finding the electronic ground-state energy E_0 requires use of the variational procedure discussed previously. The most significant result of this development is that a problem of 3N variables has been reduced to one of only 3 variables. The only drawback is that the form for $F[n(\vec{r})]$ is as of yet unknown. This requires that some further approximation be made in order to perform practical caluclations.

2.5.1   The Kohn-Sham method

It is traditional to separate out the classical Coulomb repulsion term from $F[n(\vec{r})]$, so that all of the non-classical many-body effects are contained within another functional:

$$F[n(\vec{r})] = E_H[n(\vec{r})] + G[n(\vec{r})] \tag{2.65}$$

$$= \frac{1}{2} \int \frac{n(\vec{r}_1)\, n(\vec{r}_2)}{|\vec{r}_1 - \vec{r}_2|}\, d\vec{r}_1 d\vec{r}_2 + G[n(\vec{r})]. \tag{2.66}$$

Kohn and Sham[63] proposed the use of a *fictitious* non-interacting system of electrons, described by single-particle wavefunctions. This allows one to roughly calculate the kinetic energy, ignoring correlated motion. The corrections to this approximation could then be absorbed into another term which also accounted for the exchange energy of the real system given in Eq. (2.45). The density of this hypothetical system would simply be

$$n(\vec{r}) = \sum_i^{occ} |\psi_i(\vec{r})|^2, \tag{2.67}$$

where the sum is over the N occupied lowest-energy orbitals of both spins. The Kohn-Sham kinetic energy is just

$$T_s[n(\vec{r})] = -\sum_i^{occ} \langle \psi_i(\vec{r}) | \frac{1}{2}\nabla^2 | \psi_i(\vec{r}) \rangle, \tag{2.68}$$

so that the *exchange-correlation functional* is defined by

$$E_{xc}[n(\vec{r})] = G[n(\vec{r})] - T_s[n(\vec{r})]. \tag{2.69}$$

This is not the same quantity given in Eq. (2.45), but rather also contains the difference between the real kinetic energy and $T_s[n(\vec{r})]$:

$$E_{xc}[n(\vec{r})] \equiv T[n(\vec{r})] - T_s[n(\vec{r})] + U_{e-e}[n(\vec{r})] - E_H[n(\vec{r})]. \tag{2.70}$$

The total electronic energy of the real system then becomes

$$E_{el}[n(\vec{r})] = T_s[n(\vec{r})] + \int v_{ext}(\vec{r})\, n(\vec{r})\, d\vec{r} + E_H[n(\vec{r})] + E_{xc}[n(\vec{r})], \tag{2.71}$$

which, when minimized under the constraint of Eq. (2.62), yields the chemical potential,

$$\mu = \frac{\delta T_s[n(\vec{r})]}{\delta n(\vec{r})} + v_{ext}(\vec{r}) + \int \frac{n(\vec{r}')}{|\vec{r} - \vec{r}'|}\, d\vec{r}' + \mu_{xc}(\vec{r}), \tag{2.72}$$

where

$$\mu_{xc}(\vec{r}) = \frac{\delta E_{xc}[n(\vec{r})]}{\delta n(\vec{r})} \tag{2.73}$$

is defined to be the *exchange-correlation potential*. The key assumption in the Kohn-Sham method is that there exists some effective potential for the non-interacting system which can result in the *same* ground-state density as the real interacting system. That is to say if

$$E_{el,KS}[n(\vec{r})] = T_s[n(\vec{r})] + \int v_{eff}(\vec{r})\, n(\vec{r})\, d\vec{r} \tag{2.74}$$

is minimized under the constraint of Eq. (2.62), then

$$\mu_{KS} = \frac{\delta T_s[n(\vec{r})]}{\delta n(\vec{r})} + v_{eff}(\vec{r}). \tag{2.75}$$

As a result, the non-interacting system will give the same ground state density as the real interacting system when $\mu = \mu_{KS}$ or when

$$v_{eff}(\vec{r}) = v_{ext}(\vec{r}) + \int \frac{n(\vec{r}')}{|\vec{r} - \vec{r}'|} d\vec{r}' + \mu_{xc}(\vec{r}). \qquad (2.76)$$

The whole problem may now be recast, using this effective external potential, into a pair of single-particle equations:

$$[-\frac{1}{2}\nabla^2 + v_{eff}(\vec{r})]\,\psi_i = \varepsilon_i \psi_i, \qquad (2.77)$$

$$n(\vec{r}) = \sum_i^{occ} |\psi_i(\vec{r})|^2, \qquad (2.78)$$

which must be solved self-consistently, under the conditions of orthonormalization of the single-particle wavefunctions:

$$\langle \psi_i | \psi_j \rangle = \delta_{ij}. \qquad (2.79)$$

Iterative methods are used to perform the calculations, since the density, $n(\vec{r}')$, appearing in the integral operator for the electron interaction, is necessarily dependent upon the wavefunctions. It should be noted that the ground-state electronic energy of the real interacting system *is not* just the sum of the fictitious independent-particle eigenenergies, but may be obtained from the self-consistent ground-state density:

$$E_{el}[n_0(\vec{r})] = \sum_i^{occ} \varepsilon_i - \frac{1}{2} \int \frac{n_0(\vec{r}_1)\, n_0(\vec{r}_2)}{|\vec{r}_1 - \vec{r}_2|} d\vec{r}_1 d\vec{r}_2$$
$$+ E_{xc}[n_0(\vec{r})] - \int \left. \frac{\delta E_{xc}[n(\vec{r})]}{\delta n(\vec{r})} \right|_{n_0} n_0(\vec{r})\, d\vec{r}. \qquad (2.80)$$

2.5.2 The local-density approximation

The Kohn-Sham method is simple and elegant, but the exchange-correlation functional has still not been defined. Kohn and Sham suggested to use a form for

$E_{xc}[n(\vec{r})]$ which was a local function of the density:

$$E_{xc}[n(\vec{r})] = \int \varepsilon_{xc}(n(\vec{r}))\, n(\vec{r})\, d\vec{r}, \tag{2.81}$$

where $\varepsilon_{xc}(n(\vec{r}))$ is the exchange-correlation energy per particle of a uniform gas of electrons, and $n(\vec{r})$ is the density of the electrons for a particular problem. This is what is known as the *local density approximation* or the LDA. The exchange-correlation potential $\mu_{xc}(\vec{r})$ is now written as

$$\mu_{xc}(\vec{r}) \equiv \frac{d}{d\,n(\vec{r})}\,[n(\vec{r})\varepsilon_{xc}(n(\vec{r}))] \tag{2.82}$$

$$=\varepsilon_{xc}(n(\vec{r})) + n(\vec{r})\,\frac{\partial\varepsilon_{xc}(n(\vec{r}))}{\partial n(\vec{r})}. \tag{2.83}$$

Since density functional theory is exact, the LDA is the *only* approximation, and is therefore, in principle, the only source of error in the Kohn-Sham formalism. Even though the LDA was originally used for calculations in systems with slowly varying densities, it has performed surprisingly well in atomic and molecular systems where the density gradients can be rather large.[57] Despite this success, it should be remembered that the exchange-correlation energy is a actually a non-local function of the density *everywhere*, and for this reason, the LDA fails to accurately reproduce certain characteristics of materials. For example, the LDA has a tendency to overestimate the dielectric constant by about 10 to 30%. It also can underestimate the lattice parameters of crystals by about 1-2%. There have been numerous proposed forms for $\varepsilon_{xc}(n(\vec{r}))$; [63, 121, 93, 47, 108, 109] the most successful of which are the quantum Monte Carlo results of Ceperley and Alder.[17, 18] The form used in the present work is the Wigner interpolation formula[121, 44],

$$\varepsilon_{xc}(n(\vec{r})) = n^{\frac{1}{3}}\left[-0.984 - \frac{0.944 + 8.90\,n^{\frac{1}{3}}}{(1 + 12.57n^{\frac{1}{3}})^2}\right], \tag{2.84}$$

which attempts to reproduce both the high density data and low density Wigner crystal energy.

Two notable methods which attempt to improve the LDA are the weighted density approximation (WDA) and the generalized gradient approximation (GGA). The approach taken in the GGA is to modify ε_{xc} by making it a function of not only the local density, but also of $|\nabla n(\vec{r})|$. Some initial results of this approach can be found in Refs.[85, 97, 62, 34, 43]. The WDA is more complicated, incorporating the non-local effects by using model exchange-correlation holes.[42] It is exact in the limits of a uniform electron gas, and a single electron system. Both methods greatly improve results for the ground-state properties of low mass atoms, but the GGA errors in lattice parameters can be greater than the LDA for some materials containing heavy elements. The WDA can yield poor results for metallic surfaces, and the increased computational difficulty required to implement it has resulted in a preference for the GGA over the WDA.

2.5.3 General calculational procedure for solids

Most modern calculations for crystalline solids utilize the single-particle Kohn-Sham method along with the LDA, and are commonly known as *first-principles* or *ab-initio* methods. In solid crystalline materials, we have the advantage that the potential and the density (and consequently, the Hamiltonian) are periodic in space, allowing us to invoke Bloch's theorem. The single-particle wavefunctions therefore have the crystal momentum wavevector $\vec{k}$ as a good quantum number and satisfy the

boundary condition,

$$\psi_{n,\vec{k}}(\vec{r}+\vec{R}) = e^{i\vec{k}\cdot\vec{R}}\psi_{n,\vec{k}}(\vec{r}), \tag{2.85}$$

where n is the band index, and $\vec{R}$ is a direct lattice vector. These wavefunctions are formed from some linear combination of known basis functions, similar to what was done in Eq. (2.54) for the many-body wavefunction. The variational single-particle wavefunction is written as

$$\psi_{n,\vec{k}}(\vec{r}) = \sum_{\vec{G}} C_{n\vec{k}}(\vec{G})\,\phi_{\vec{k}+\vec{G}}(\vec{r}), \tag{2.86}$$

where $\vec{G}$ is a reciprocal-lattice wavevector, and $\vec{k}$ is restricted to the first Brillouin zone. The LDA total energy is minimized with respect to the unknown coefficients, $C_{n\vec{k}}(\vec{G})$, when the secular equations are satisfied:

$$\sum_{\vec{G}'} [H_{\vec{k}+\vec{G},\vec{k}+\vec{G}'} - \varepsilon_{n,\vec{k}}O_{\vec{k}+\vec{G},\vec{k}+\vec{G}'}]\,C_{n,\vec{k}}(\vec{G}') = 0. \tag{2.87}$$

Here, $H_{\vec{k}+\vec{G},\vec{k}+\vec{G}'}$ and $O_{\vec{k}+\vec{G},\vec{k}+\vec{G}'}$ are the Hamiltonian and overlap matrix elements, respectively,

$$H_{\vec{k}+\vec{G},\vec{k}+\vec{G}'} = \langle\phi_{\vec{k}+\vec{G}}(\vec{r})|\,H\,|\phi_{\vec{k}+\vec{G}'}(\vec{r})\rangle \tag{2.88}$$

$$O_{\vec{k}+\vec{G},\vec{k}+\vec{G}'} = \langle\phi_{\vec{k}+\vec{G}}(\vec{r})|\,\phi_{\vec{k}+\vec{G}'}(\vec{r})\rangle, \tag{2.89}$$

with $H = -\frac{1}{2}\nabla^2 + v_{eff}(\vec{r})$ as in Eq. (2.77). The rank of these matrices is equal to the number of basis functions used in Eq. (2.86), resulting in an equivalent number of eigenvalues and eigenfunctions at the wavevector $\vec{k}$. However, only the N lowest states will be occupied. The secular equation is solved on a 3-dimensional mesh of $\vec{k}$-points in reciprocal space, but in such a manner that these points lie only within the

symmetry-reduced wedge of the first Brillouin zone. The charge density is constructed using the resulting wavefunctions of the occupied states:

$$\rho(\vec{r}) \doteq \sum_{n}^{occ} \int |\psi_{n,\vec{k}}(\vec{r})|^2 d\vec{k}. \tag{2.90}$$

In practice, the integral above is replaced by a symmetry-weighted sum over the same grid points where the time-consuming diagonalizations take place. The location and density of these grid points is chosen so as to maximize the computational efficiency.[81] For an insulator with a large number of electrons, even a low density grid can perform rather well. The new charge density is "mixed" with the previous one in some algebraic fashion, and the resulting density is used to create a new $v_{eff}(\vec{r})$, which is incorporated into a subsequent Hamiltonian. The diagonalization is repeated, and the process continues until self-consistency is achieved. To start this iterative technique, it is necessary to create an inital "guess" for the charge density. This is typically produced from some preliminary calculations for the component atoms in the material.

The sum in Eq. (2.86) is over as many basis functions as are necessary for an accurate representation of $\psi_{n,\vec{k}}(\vec{r})$. If the true Kohn-Sham single-particle wavefunctions can be expressed exactly in terms of the basis functions, then finding the $C_{n\vec{k}}(\vec{G})$ will give the exact self-consistent solution to the problem. However, this may require a large number of basis functions. A larger basis set will produce larger matrices, requiring more time for diagonalization. Even if the choice of basis does not permit exact representation of the $\psi_{n,\vec{k}}(\vec{r})$, this procedure will yield an optimal set of coefficients which minimize the total energy. Thus, the quality of a given basis set is judged upon its ability to achieve an acceptable compromise between accurate results and

minimum computation time. Many initial LDA calculations used plane waves as basis functions, since they form a complete set, and do not treat any region of the crystal with preference (they are unbiased). However, the large number of basis functions required to mimic the atomic-like states near the nuclei can make this approach quite expensive computationally. A more complicated method uses linear combinations of atomic orbitals (LCAO). This leads to improved computational efficiency, especially for large systems. However, attempting to achieve good convergence by using more basis functions encounters problems due to the fact that the orbitals at each nuclear site are already complete, making the variational wavefunction overcomplete. The linearized augmented plane wave method (LAPW), described in the next chapter, with its dual representation of the wavefunction, overcomes many of the limitations of these earlier methods. It can efficiently deal with the core electronic states (including relativistic effects), and has no difficulty treating materials with localized orbitals. The development of a formalism for the accurate calculation of forces has made possible the investigation of dynamical properties from first-principles.[125] Because of these capabilities, the LAPW method has become perhaps the most accurate method for producing accurate electronic structure calculations for materials containing transition metal atoms (with their localized d orbitals), making it an ideal tool for first-principles studies of many modern ferroelectric materials.

Chapter 3

The Linearized Augmented Plane Wave (LAPW) Method

3.1   Background

The *linearized augmented plane wave* or LAPW method is a natural extension of Slater's original (APW) method.[99, 69] In the regions near the nucleus, the wavefunctions will be rapidly varying due to the steep potential gradient, while in the bonding regions, they will be much smoother, as will the potential. The idea is to use dual-representation basis functions which can express both of these characteristics. Inside non-overlapping *muffin-tin spheres*, centered on each atom, one uses a product of spherical harmonic functions multiplied by numerical solutions to a radial Schrödinger equation having the form

$$\left[-\frac{d^2}{dr^2} + \frac{l(l+1)}{r^2} + V(r) - E_l \right] r\, u_l(r) = 0. \tag{3.1}$$

38

The radial functions, $u_l(r)$, depend upon E_l as a parameter, and $V(r)$ is the spherical part of the potential within the sphere. Outside the spheres, plane waves are used, and the two representations are forced to match at the spherical boundary. Thus, the expansion coefficients inside the muffin-tin sphere are completely determined by the plane wave expansion coefficients outside and the energy parameters E_l. In the original APW method, the E_l were used as non-linear variational parameters, and resulted in a standard secular equation which was dependent upon E_l, requiring a costly root-finding algorithm. Another less severe problem, known as the asymptote problem, was that it was possible for some of the $u_l(r)$ to have a node on the sphere boundary, causing the secular determinant to vary wildly. These problems were later solved when Andersen, Koelling, and Arbman[2, 60] proposed including the derivative of the radial functions with respect to E_l in the basis, and forcing continuity of both the basis functions *and* their first radial derivatives at fixed E_l. This is the essence of the LAPW method. The addition of the energy derivative introduces errors of order $(\varepsilon - E_l)^2$ in the wavefunction, which translates into errors in the band energy of order $(\varepsilon - E_l)^4$. This high order for the error means that the LAPW basis functions become useable over a relatively wide range of energies, allowing all of the valence bands to be treated with a single set of E_l. Furthermore, a single diagonalization at a particular $\vec{k}$-point is able to produce accurate energy bands. The establishment of the LAPW method as one of the most powerful and accurate techniques to date began with with the use of model atomic potentials [60], and was extended to fully self-consistent calculation methods for slabs [56, 64, 46] and for bulk materials [44, 122, 115].

Since then, Yu, Singh and Krakauer have developed a formalism for force cal-

culations within the LAPW method,[125] and this work has recently been extended into the linear response regime, allowing accurate LAPW studies of the dynamical properties of materials from first principles.[126, 110, 111, 113] Details of the LAPW method have been reviewed by Singh[98], Wei[115], and Lu[70]. The material covered in the next few sections will give a basic overview of the LAPW basis functions, charge density, potential, and total energy. This will be followed by a discussion of pseudopotentials and their use within the LAPW method. Forces and linear-response calculations will be covered in the next chapter.

3.2 LAPW basis functions

In the LAPW method, the Kohn-Sham orbitals are expanded as in Eq. (2.86), with

$$
\phi_{\vec{k}+\vec{G}}(\vec{r}) = \begin{cases} \frac{1}{\sqrt{\Omega}} e^{i(\vec{k}+\vec{G})\cdot\vec{r}}, & \vec{r} \in \text{Interstitial} \\ \sum_{l,m} \left[a_{lm}(\vec{k}+\vec{G}) \, u_l(E_l^i; r_i) + b_{lm}(\vec{k}+\vec{G}) \, \dot{u}_l(E_l^i; r_i) \right] Y_{lm}(\hat{r}_i), & r_i \leq R_i, \end{cases}
$$

$$(3.2)$$

where Ω is the volume of the unit cell. The dot in the expression $\dot{u}_l(E_l^i; r_i)$ indicates the first derivative with respect to E_l^i. Also, $\vec{r}_i = \vec{r} - \vec{r}_i$, where $\vec{r}_i$ is the position of the i-th muffin-tin sphere of radius R_i in the unit cell. As usual, l and m are the orbital angular momentum quantum numbers. The expansion coefficients $a_{lm}(\vec{k}+\vec{G})$ and $b_{lm}(\vec{k}+\vec{G})$ are determined by forcing both the basis functions and their radial derivatives to be continuous at the muffin-tin sphere boundary. Not only do the radial

functions satisfy Eq. (3.1), but their energy derivatives solve a similar equation:

$$\left[-\frac{d^2}{dr^2} + \frac{l(l+1)}{r^2} + V(r) - E_l \right] r\, \dot{u}_l(r) = r\, u_l(r). \tag{3.3}$$

The u_l are chosen to be normalized within the muffin-tin sphere.

$$\int_0^{R_i} r^2 u_l^2(r)\, dr = 1, \tag{3.4}$$

which implies that $u_l(r)$ and $\dot{u}_l(r)$ are orthogonal:

$$\frac{d}{dE_l}\left[\int_0^{R_i} r^2 u_l^2(r)\, dr \right] = 2\int_0^{R_i} r^2 u_l(r)\dot{u}_l(r)\, dr = 0. \tag{3.5}$$

In general, the $\dot{u}_l(r)$ are non-normalized:

$$\int_0^{R_i} r^2\, \dot{u}_l^2(r)\, dr = N_l. \tag{3.6}$$

One of the identities from the APW method [15], which still holds in the LAPW method, is

$$R_i^2\left[\dot{u}_l(E_l^i, R_i)\, u_l'(E_l^i, R_i) - u_l(E_l^i, R_i)\, \dot{u}_l'(E_l^i, R_i) \right] = 1. \tag{3.7}$$

Here, the prime symbol indicates a radial derivative. From this we can see that even if $u_l(r)$ has a node at the sphere boundary, then in general, $u_l'(R_i)$ and $\dot{u}_l(R_i)$ will be non-zero. This removes the asymptote problem inherent in the APW method. Using the energy derivative gives the additional advantage of more variational freedom within the spheres. Even if we do not know the band energy ε_l, it may be possible to approximate the correct APW wavefunction using the energy derivative:

$$u_l(\varepsilon_l, r) = u_l(E_l, r) + (\varepsilon_l - E_l)\dot{u}_l(r) + O((\varepsilon_l - E_l)^2). \tag{3.8}$$

If the error in the wavefunction is second-order, then the error in the energy will be fourth-order. Thus, the LAPW basis can produce good results over a range of band energies. Usually, the E_l are chosen to lie near the center of the bands of interest. In this way, accurate energy bands can be obtained with only one diagonalization at each $\vec{k}$-point, which greatly improves the efficiency over the APW method. This is often the case, but problems can develop, especially with physically extended semi-core states having the same angular momentum as some of the valence states. This is usually dealt with by performing two independent calculations within separate energy windows. This approach, and the problems it can generate, are discussed by Mattheiss and Hamann,[74] as well as by Singh and Krakauer.[95, 98] If the kinetic energies of the electrons are large, as they will be for high-mass atoms, then relativistic effects cannot be ignored. In this case, the radial Schrödinger equations (Eq. (3.1) and Eq. (3.3)) are replaced by the j-weighted averaged Dirac equation and its energy derivative.[61] The spin-orbit interaction term is dropped in this treatment, but it can be added later on. When constructing the charge density and the Hamiltonian matrix elements, both the large and small radial components are used, whereas only the large component and its derivative are used to match the functions at the boundary, since the valence electrons surrounding the sphere surface are typically non-relativistic.

3.3 The charge density

The charge density is determined from the occupied Kohn-Sham orbitals,

$$\rho(\vec{r}) = \sum_i^{occ} n_i \, |\psi_i(\vec{r})|^2, \tag{3.9}$$

where n_i is the occupation number of the state i at the k-point, $\vec{k}_i$, times the symmetry weighting of the k-point. This summation replaces the integral in Eq. (2.90), and is performed over a mesh of $\vec{k}$-points, as mentioned earlier. The generation of this mesh is accomplished using a special k-point method,[19, 81] which optimizes the convergence of this sum for relatively smooth bands as a mesh of increasing density is used. The diagonalizations, which determine the wavefunctions used to construct $\rho(\vec{r})$, are necessarily also performed at each $\vec{k}$-point of this mesh. In the LAPW method, a dual representation is used for the charge density and the potential as well, but the natural crystalline symmetry can drastically reduce the number of expansion coefficients in both the interstitial and the muffin-tin regions. The real charge density is written as

$$\rho(\vec{r}) = \begin{cases} \sum_\mu \rho_\mu(r_i) K_\mu(\hat{\vec{r}}_i), & \vec{r}_i \in i^{th} \text{ sphere} \\ \sum_s \rho_s \phi_s(\vec{r}), & \vec{r} \in \text{interstitial} \end{cases} \tag{3.10}$$

where the functions $K_\mu(\hat{\vec{r}}_i)$, called *lattice harmonics*, are created from linear combinations of the ordinary spherical harmonics using the site symmetry of the ith atom:

$$K_{i,\nu}(\hat{\vec{r}}_i) = \sum_m c^i_{\nu,m} Y_{lm}(\hat{\vec{r}}_i). \tag{3.11}$$

Here, the coefficients $c^i_{\nu,m}$ are determined under the conditions that the lattice harmonics are real-valued, orthonormal, and invariant under the point symmetry of the ith atom site. Any function within the spheres may now be expanded in terms of the appropriate lattice harmonics. Symmetry-related atoms will have the same lattice harmonic expansions, differing only in a rotation of the local coordinate system. Similarly, any function which has the space group symmetry of the crystal can be

expanded in terms of *star functions*, $\phi_s(\vec{r})$, which are constructed from symmetrized plane waves:

$$\phi_s(\vec{r}_i) = \frac{1}{N_{op}} \sum_{\hat{R}} e^{i\hat{R}\vec{G}\cdot(\vec{r}-\vec{t}_{\hat{R}})}. \tag{3.12}$$

Here, $\hat{R}$ and $\vec{t}_{\hat{R}}$ are the rotation and translation components of each space group operation, of which there are N_{op}. The ρ_s are the expansion coefficients of the charge density in the star basis.

Within the muffin-tin spheres, the density and the potential are expanded on a discrete radial mesh. A logarithmic mesh is used so that more points cover regions of rapid variation near the center:

$$r_{j+1} = r_j e^{\Delta x}, \tag{3.13}$$

where typically, $\Delta x \sim 0.025$.

A single self-consistent calculation begins with the generation of the potential from an input charge density. This input density is created by mixing the most recent output density with the previous input density. The simplest scheme for doing this is straight mixing

$$\rho(\vec{r})_{in}^{m+1} = (1 - \alpha)\,\rho(\vec{r})_{in}^m + \alpha\,\rho(\vec{r})_{out}^m, \tag{3.14}$$

which will converge if the mixing parameter, α, is not too large. It is a crude mixing method which converges poorly for large systems. In this dissertation, the straight mixing method was used for the initial input charge density. For subsequent iterations, Broyden's mixing method,[11] which utilizes information about how the density is changing, yields much better convergence.

3.4 The potential

For crystals, the effective potential appearing in the Kohn-Sham equations is a combination of the exchange-correlation potential and the Coulomb potentials due to the nuclei and the electron charge density. The Coulomb terms can be determined from Poisson's equation,

$$\nabla^2 V_c(\vec{r}) = 4\pi\rho(\vec{r}), \tag{3.15}$$

which is usually calculated in reciprocal space, via Fourier transform, since the equation is diagonal in $\vec{G}$, giving

$$V_c(\vec{G}) = \frac{4\pi\rho(\vec{G})}{|\vec{G}|^2}. \tag{3.16}$$

However, since the charge density within the core (near the nucleus) is rapidly varying, a Fourier expansion in this region converges extremely slowly. A real-space solution to Poisson's equation is also complicated by the long range of the Coulomb potential. Fortunately, a method for solving Poisson's equation was proposed by Weinert,[118] which utilizes the fact that the potential outside of the muffin tin spheres is dependent only upon the interstitial charge density and the *multipole moments* of the charge density, q_{lm}^i, inside of the spheres. This means that a given exterior potential can come from a variety of charge densities within the sphere, as long as the multipole moments are the same for each. Because of this, a smoother *pseudocharge density* may be constructed so that it maintains the same real charge density in the interstitial region, while having the same multipole moments as the true charge density within the spheres. This is done by adding a *difference* charge density, which can easily be Fourier transformed, and which possesses multipoles equal to the difference between

the true sphere multipoles and the plane wave multipoles:

$$\Delta \tilde{q}_{lm}^i = q_{lm}^i - q_{lm}^{iPW}.$$

(3.17)

This difference charge density is somewhat arbitrary, and in the work presented in this dissertation, the Weinert polynomial form,

$$\Delta \tilde{\rho}^i(\vec{r}_i) = \sum_{lm} \Delta \tilde{q}_{lm}^i \left(\frac{(2l + 2N_l + 3)!!}{2^{N_l} N! \, (2l + 1)!!} \right) \frac{r^l}{R_i^{2l+3}} \left[1 - \frac{r^2}{R_i^2} \right]^{N_l} Y_{lm}(\hat{\vec{r}}),$$

(3.18)

was used since its N_l-1 derivatives are continuous and its Fourier transform is analytic. N_l is a parameter chosen to optimize the convergence of the Fourier transform of this difference charge density. Now the pseudocharge would normally be defined in terms of star functions and lattice harmonics

$$\tilde{\rho}(\vec{r}) = \sum_s \rho_s \phi_s(\vec{r}) + \sum_{i,\mu} \Delta \tilde{\rho}_\mu^i(r_i) K_\mu(\hat{\vec{r}}_i) \, \theta(r_i),$$

(3.19)

where $\theta(r_i)$ is the step function, which is unity when r_i lies within the ith sphere boundary, and zero otherwise. Choosing the $\Delta \tilde{\rho}_\mu^i(r_i)$ such that

$$\Delta \tilde{\rho}^i(\hat{\vec{r}}_i) = \sum_\mu \Delta \tilde{\rho}_\mu^i(r_i) K_\mu(\hat{\vec{r}}_i),$$

(3.20)

the pseudocharge density is written entirely in terms of star functions:

$$\tilde{\rho}(\vec{r}) = \sum_s \tilde{\rho}_s \phi_s(\vec{r}).$$

(3.21)

The interstitial Coulomb potential can then be simply obtained as

$$V_c(\vec{r}) = \sum_s \frac{4\pi \tilde{\rho}_s}{G_s^2} \phi_s(\vec{r}), \qquad \vec{r} \in \text{interstitial}.$$

(3.22)

Once the interstitial potential is computed, it can be used as a boundary condition for solving Poisson's equation inside the spheres with the true charge density. This is

done in real space using a standard Green's function approach for solving Poisson's equation inside a sphere.

The exchange-correlation potential must be calculated in real space from the charge density. The interstitial contribution to μ_{xc} is obtained by a complex fast Fourier transform (CFFT) to values on the real space FFT grid. $\mu_{xc}(\rho(\vec{r}))$ is calculated on each grid point, and is then back-transformed into a plane wave representation. Within the muffin-tin spheres, μ_{xc} is determined directly from the charge density at the radial mesh points. The exchange-correlation potential is then expanded in terms of the lattice harmonics via a least squares fitting technique.

3.5   The Hamiltonian and overlap matrices

The Hamiltonian and overlap matrix elements (Eqs. 2.88 and 2.89) are also broken into interstitial and muffin-tin components:

$$S_{\vec{G}\vec{G}'} = \frac{1}{\Omega} \int_{\Omega} e^{i(\vec{G}'-\vec{G})\cdot\vec{r}} \Theta(\vec{r})\, d\vec{r} + \sum_{a} S_a(\vec{G}, \vec{G}'), \qquad (3.23)$$

and

$$H_{\vec{G}\vec{G}'} = \frac{1}{\Omega} \int_{\Omega} e^{-i(\vec{G}+\vec{k})\cdot\vec{r}} [T + V_{PW}]\, e^{i(\vec{G}'+\vec{k})\cdot\vec{r}} \Theta(\vec{r})\, d\vec{r} \\ + \sum_{a} [H_a(\vec{G}, \vec{G}') + V_{NS}^a(\vec{G}, \vec{G}')]. \qquad (3.24)$$

The step function, $\Theta(\vec{r})$, in these expressions is zero if the position vector lies within a muffin-tin sphere, and unity in the interstitial. Note that this definiton is opposite of that used for the pseudocharge density. The muffin-tin portions of the overlap matrix are contained in $S_a(\vec{G}, \vec{G}')$. The interstitial potential is V_{PW}, and T is the kinetic

energy operator. The muffin-tin part of the Hamiltonian has been further separated into spherical and non-spherical terms, where $H_a(\vec{G}, \vec{G}')$ includes the kinetic energy and $l = 0$ potential terms, and $V^a_{NS}(\vec{G}, \vec{G}')$ represents the $l \neq 0$ contributions to the potential. Details of the construction of these matrix elements are described by Singh.[98]

3.6   The total energy

The calculation of the total energy is often the ultimate objective of any first-principles calculation. Its minimization with respect to structural perturbations yields the ground state structure and its corresponding physical properties. In addition, the forces on the atoms may be determined, giving information about the linear response of the system. The total energy in the Kohn-Sham formulation is not just the sum of the single particle eigenvalues, for if it were, it would double count for the electron-electron interactions. They can be used however, to rewrite the total energy per unit cell in the LDA as:

$$E_{tot} = \sum_i^{occ} n_i\, \varepsilon_i - \frac{1}{2} \int_\Omega \int \frac{\rho(\vec{r})\, \rho(\vec{r}')}{|\vec{r} - \vec{r}'|}\, d\vec{r}\, d\vec{r}' - \int_\Omega \rho(\vec{r})\, \mu_{xc}(\vec{r})\, d\vec{r}$$
$$+ \int_\Omega \rho(\vec{r})\, \varepsilon_{xc}(\vec{r})\, d\vec{r} + \frac{1}{2} \sum_{a \neq b} \sum_{\vec{R}} \frac{Z_a Z_b}{|\vec{R} + \vec{r}_a - \vec{r}_b|}, \tag{3.25}$$

where again, the sum over the single-particle eigenvalues is only for the occupied states, with n_i defined as it was for the charge density. The difficulty with this form for the total energy is the inherent Coulomb singularities present in the Hartree term.

These may be forced to cancel, if we write the total energy per unit cell as

$$E_{tot} = \sum_i^{occ} n_i \, \varepsilon_i - \frac{1}{2} \int_\Omega \rho(\vec{r}) \, \chi(\vec{r}) \, d\vec{r} - \frac{1}{2} \sum_\nu Z_\nu \, V_M(\vec{\gamma}_\nu),$$ (3.26)

where $\vec{\gamma}_\nu$ locates only the nuclei within the unit cell,

$$\chi(\vec{r}) = \int \frac{\rho(\vec{r}')}{|\vec{r} - \vec{r}'|} \, d\vec{r}' - \sum_a \sum_{\vec{R}} \frac{Z_a}{|\vec{r} - \vec{r}_a - \vec{R}|} + 2(\mu_{xc}(\vec{r}) - \varepsilon_{xc}(\vec{r})),$$ (3.27)

and $V_M(\vec{\gamma}_\nu)$ is a generalized Madelung potential [119] due to all of the charges in the crystal except for the nuclear charge at the site $\vec{\gamma}_\nu$:

$$V_M(\vec{\gamma}_\nu) = \int \frac{\rho(\vec{r})}{|\vec{r} - \vec{\gamma}_\nu|} \, d\vec{r} - \sum_{a \neq \nu} \sum_{\vec{R}} \frac{Z_a}{|\vec{R} + \vec{r}_a - \vec{\gamma}_\nu|}.$$ (3.28)

Expanding $\chi(\vec{r})$ in terms of the lattice harmonics and star functions, the form actually used for the present LAPW calculations is [115]:

$$\begin{aligned}
E_{tot} = \sum_i^{occ} n_i \, \varepsilon_i &- \frac{1}{2} \sum_a Z_a \left[\frac{2(Z_a - Q_{MT}^a)}{R_a} + V_{MT}^a(R_a) \right] \\
&- \frac{1}{2} \int_I \rho(\vec{r}) \, \chi(\vec{r}) \, d\vec{r} - \frac{1}{2} \sum_{a,\mu} \int_0^{R_a} \chi_\mu(r_a) \, \rho_\mu(r_a) \, r_a^2 \, dr_a \\
&- \frac{1}{2} \sum_a \int_0^{R_a} \left[\frac{2 Z_a \, \sigma(r_a)}{r_a} + \chi_{MT}(r_a) \, \sigma(r_a) \right] dr_a,
\end{aligned}$$ (3.29)

where Q_{MT}^a is the number of electrons inside sphere a, and $\sigma(r_a) = 4\pi \, r^2 \rho(r_a)$. The quantity $V_{MT}^a(R_a)$ is the spherical average of the Coulomb potential on the surface of sphere a. Subscript I indicates integration in the interstitial region only.

3.7 Pseudopotentials

Before the development of the LAPW method, the majority of computational techniques were purely planewave-based. Not only is the use of a plane wave basis

set conceptually simpler, but it also posesses numerical advantages. Fast Fourier transform (FFT) techniques permit real-space and reciprocal space representations to be determined easily, and the solution of Poisson's equation is simple (Eq. (3.15)). In addition, the matrix elements are much easier to construct. Practical calculations use a finite set of plane waves up to some cutoff value for $|\vec{G}|$, such that the accuracy is maximized while the computational time is kept to a minimum. The chief drawback to plane wave methods is that the valence wavefunctions oscillate rapidly near the atomic nuclei as a result of othogonalization to the core states. The high cutoff energy in such systems makes the basis set too large to be practical. One way to get around this limitation is to replace the strong atomic potential with a weaker *pseudopotential* that projects out the core electron states. By doing so, the original Hamiltonian is replaced by an effective Hamiltonian having pseudo-states corresponding to the valence electrons only. The success of the method is based on the validity of the *frozen core approximation*, which is justified by the fact that the chemical properties of materials are determined by the valence electrons only.

Early pseudopotentials were empirical in nature, but have since been superceded by the development of first-principles pseudopotentials, introduced by Hamann, Schlüter, and Chiang (HSC).[45] These are constrained to have the following characteristics: (i) the pseudowavefunctions are nodeless for all r; (ii) the pseudopotential and the real potential are identical for r greater than or equal to some core radius r_c; (iii) the eigenvalues of the pseudo valence states are the same as the true valence states; (iv) the pseudo valence wavefunctions and the real valence wavefunctions are identical beyond r_c; (v) the integrals of the real charge density and the pseudocharge
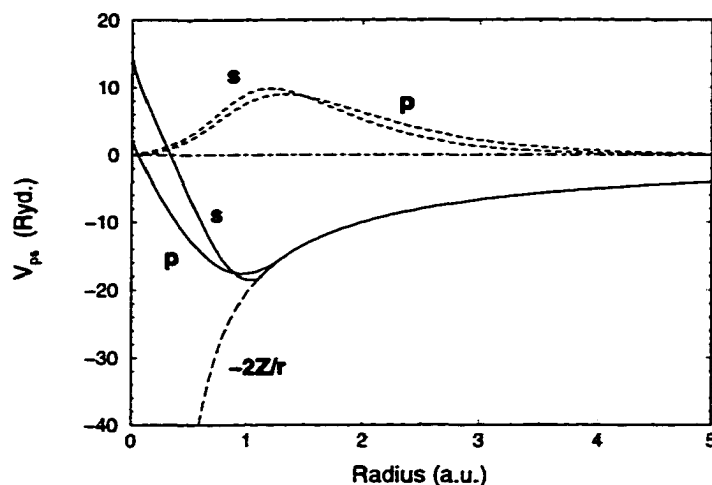
Figure 3.1: Kerker non-local ionic pseudopotential and pseudovalence wavefunctions for $4s$ and $4p$ states in Sr^{+2}.

density up to the core radius are equal for each valence state. This last condition, known as *norm conservation*, ensures that the electrostatic and scattering properties of the pseudopotential are reproduced with minimum error as the pseudo-atom is placed in different chemical environments. Since the true valence wavefunctions have different eigenvalues for each angular momentum value, l, the pseudopotential must necessarily be dependent upon l as well. Thus, these kinds of pseudopotentials are called *non-local*. Each l value will have a corresponding r_c, at which the matching conditions stated above are to be satisfied. If the core radius is chosen to be relatively small, then the pseudowavefunction can more accurately reproduce the true wavefunction near the nucleus, and the resulting pseudopotential is acceptable for use in a larger variety of environments. This is the issue of *transferability*. However, this means that the pseudowavefunctions will have sharper features, and may not necessarily lend themselves to efficient plane wave representation. The choice of r_c is
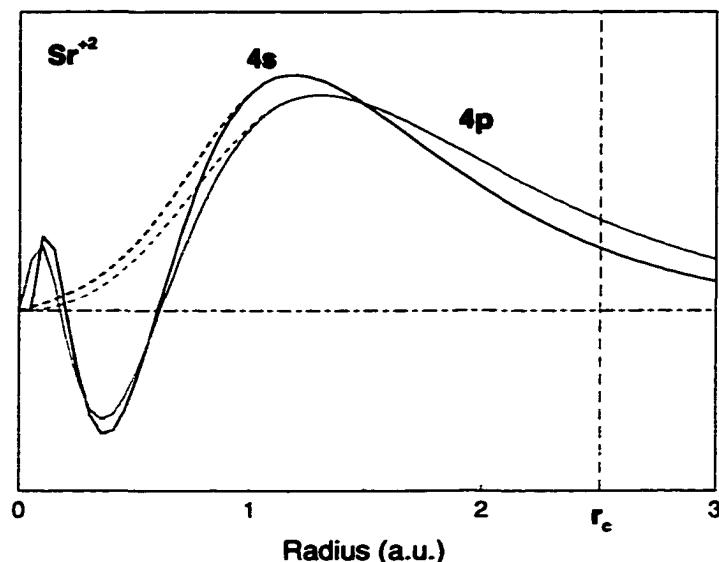
Figure 3.2: The pseudowavefunctions for $4s$ and $4p$ states of Sr^{+2} are nodeless and identical to the true wavefunctions beyond r_c in the Kerker scheme.

one of quality and transferability, and should not be thought of as a free parameter. The HSC method was later refined by Bachelet, Hamann, and Schlüter [4] (BHS) and yielded high-quality pseudopotentials with good transferability. Kerker [58] proposed a scheme which was simpler than that of BHS, yet produced pseudopotentials of the same quality. In the Kerker construction, the pseudopotential is strictly local for

$r > r_c$.

Although the LAPW method can easily handle the characteristics of the core regions, the utilization of pseudopotentials means that the chemically inactive core states can be ignored. Hamann,[44] compared full potential and pseudopotential results, illustrating that the pseudopotential approach was workable within the LAPW method and gave good results. In order that the method not be complicated by the use of pseudopotentials, it is desirable to keep the l-dependence within the muffin-tin

sphere, such that the pseudopotential is local and equal to the screened all-electron value beyond the surface of the sphere. This can be assured by setting the muffin tin radius to be larger than the largest value of r_c. With this in mind, the simplicity of Kerker pseudopotentials makes them a natural choice for use within the LAPW method. More recently, LAPW force calculations using these pseudopotentials were performed by Yu, Singh, and Krakauer,[125] which showed excellent agreement with the all-electron forces.

In order to incorporate the pseudopotential into the LAPW formalism, it is split into local and non-local parts

$$V_{ps}^a(\vec{r}) = \sum_{l=0}^{l_{max}} \widetilde{V}_{NL}^{l,a}(\vec{r}) \tag{3.30}$$

$$= V_L^a(r) + \left[\sum_{l=0}^{l_{max}} \widetilde{V}_{NL}^{l,a}(\vec{r}) - \widetilde{V}_{NL}^{l',a}(r) \right] \tag{3.31}$$

$$= V_L^a(r) + \sum_{l=0}^{l_{max}} v_l^a(r) \, \hat{P}_l, \tag{3.32}$$

where an arbitrarily chosen l' angular momentum component of the atomic pseudopotential is used as a local potential, $V_L^a(r)$. The term $\hat{P}_l$ is an operator which projects out the component of the wavefunction having angular momentum l. The local part is then rewritten using the ionic potential, Z_{ion}^a/r, to define a much smoother local potential $v_L^a(r)$:

$$V_L^a(r) = -\frac{Z_{ion}^a}{r} + \left[V_L^a(r) - \left(-\frac{Z_{ion}^a}{r} \right) \right] \tag{3.33}$$

$$= -\frac{Z_{ion}^a}{r} + v_L^a(r). \tag{3.34}$$

The ionic potential then replaces the nuclear potential of the all-electron problem. Since the redefined local potential will extend beyond the muffin-tin sphere slightly,

it is further decomposed as:

$$v_L^a(r) = w_a(r) + \bar{v}_a(r),$$

(3.35)

where

$$w_a(r) = \begin{cases} v_L^a(r) - (A_a + B_a r^2) \ , \ r \le R_a \\ 0 \ , \ r \ge R_a, \end{cases}$$

(3.36)

and

$$\bar{v}_a(r) = \begin{cases} A_a + B_a r^2 \ , \ r \le R_a \\ v_L^a(r) \ , \ r \ge R_a. \end{cases}$$

(3.37)

The coefficients A_a and B_a are obtained by imposing the condition that $\bar{v}_a(r)$ and its derivative are continuous at $r = R_a$. Now $\bar{v}_a(r)$ is smooth and slowly-varying, allowing the expansion of

$$\tilde{V}(\vec{r}) = \sum_a \sum_{\vec{R}} \bar{v}_a(|\vec{r} - \vec{r}_a - \vec{R}|)$$

(3.38)

in terms of plane waves only. It is then further written in terms of the star functions and lattice harmonics. The total expression for the pseudopotential can now be written as

$$V_{ps}(\vec{r}) = \sum_a \sum_{\vec{R}} \left\{ w_a(|\vec{r} - \vec{r}_a - \vec{R}|) + \sum_l v_l^a(|\vec{r} - \vec{r}_a - \vec{R}|) \hat{P}_l \right\} + \tilde{V}(\vec{r}).$$

(3.39)

The Kohn-Sham equations are now solved using a new Hamiltonian

$$H = T + V_{ps}(\vec{r}) + V_{eff}(\vec{r}),$$

(3.40)

where V_{eff} is given by

$$V_{eff}(\vec{r}) = -\sum_a \sum_{\vec{R}} \frac{Z_{ion}^a}{|\vec{r} - \vec{\tau}_a - \vec{R}|} + \int \frac{\rho(\vec{r}')}{|\vec{r} - \vec{r}'|} \, d\vec{r}' + \mu_{xc}(\vec{r}). \qquad (3.41)$$

The radial functions and their energy derivatives, $u_l(r)$ and $\dot{u}_l(r)$, are also determined using a new muffin-tin portion of the Hamiltonian:

$$H_{MT}^a = T + w_a(r) + \sum_l v_l^a(r) \, \hat{P}_l + \widetilde{V}^S(r) + V_{eff}^S(r), \qquad (3.42)$$

where S refers to the spherical part of each term.

Chapter 4

Lattice Dynamics from First Principles

4.1 Forces in the LAPW method

Although density functional theory and the LDA provide a route to the ground state total energy for a given configuration of atoms, the determination of the lowest energy structure often requires many such calculations for a variety of atomic positions. Mapping out the Born-Oppenheimer energy surface in this manner can become prohibitively difficult for systems containing a large number of atoms or many internal degrees of freedom. The ability to calculate forces on the constituent atoms can greatly speed up the search for the ground state structure. This idea is essential in the Car-Parrinello approach to molecular dynamics simulations.[16] Once the structural energy minimum is found, the lattice dynamical properties can also be determined using the forces which result from phonon-like distortions.

56

The force on atom a is evaluated by displacing the nucleus by a small amount $\delta \vec{r}_a$, and calculating the resulting first-order changes in the total energy:

$$\delta T[\rho] = \sum_i \delta n_i \, \varepsilon_i + \sum_i n_i \, \delta\varepsilon_i - \int \delta\rho(\vec{r}) \, v_{eff}(\vec{r}) \, d\vec{r} - \int \rho(\vec{r}) \, \delta v_{eff}(\vec{r}) \, d\vec{r}$$

$$\delta U[\rho] = \int \delta\rho(\vec{r}) \, v_c(\vec{r}) \, d\vec{r} - \vec{F}_{HF}^a \cdot \delta\vec{r}_a \tag{4.1}$$

$$\delta E_{xc}[\rho] = \int \delta\rho(\vec{r}) \, \mu_{xc}(\vec{r}) \, d\vec{r}.$$

The effective potential $v_{eff}(\vec{r})$ is the sum of the Coulomb potential $v_c(\vec{r})$,

$$v_c(\vec{r}) = \int \frac{\rho(\vec{r}')}{|\vec{r} - \vec{r}'|} \, d\vec{r}' - \sum_b \sum_{\vec{R}} \frac{Z_b}{|\vec{r} - \vec{r}_b - \vec{R}|}, \tag{4.2}$$

and the exchange-correlation potential $\mu_{xc}(\vec{r})$:

$$\mu_{xc}(\vec{r}) = \frac{\partial E_{xc}[\rho(\vec{r})]}{\partial \rho(\vec{r})}. \tag{4.3}$$

The Hellmann-Feynman (HF) force [49, 31] is the classical electrostatic force on the nucleus, given by

$$\vec{F}_{HF}^a = Z_a \nabla_{\vec{r}} \, v_c^a(\vec{r}), \, |_{\vec{r}=\vec{r}_a} \tag{4.4}$$

where $v_c^a(\vec{r})$ is the electrostatic potential due to all of the other charges in the system:

$$v_c^a(\vec{r}) = v_c(\vec{r}) + \frac{Z_a}{|\vec{r} - \vec{r}_a|}. \tag{4.5}$$

Due to conservation of electron number, the first term in the expression for δT disappears, yielding a total force of

$$\vec{F}^a = \vec{F}_{HF}^a - \frac{1}{\delta\vec{r}_a}\left[\sum_i n_i \delta\varepsilon_i - \int \rho(\vec{r}) \, \delta v_{eff}(\vec{r}) \, d\vec{r}\right]. \tag{4.6}$$

The existence of the term in brackets was demonstrated by Pulay and it depends upon the basis set used to solve the Kohn-Sham equations.[87] Normally, the single-particle wavefunctions $\psi_i(\vec{r})$ are expanded in a linear combination of basis functions $\phi(\vec{r})$:

$$\psi_i(\vec{r}) = \sum_q C_q \, \phi_q(\vec{r}), \tag{4.7}$$

leading to the matrix equations

$$\sum_{pq} C_p^*(H_{pq} - \varepsilon O_{pq}) \, C_q = 0. \tag{4.8}$$

When atom a is shifted by $\delta \vec{r}_a$, the matrix elements change by:

$$\delta H_{pq} = \langle \delta \phi_p | H | \phi_q \rangle + \langle \phi_p | H | \delta \phi_q \rangle + \langle \phi_p | \delta \hat{T} + \delta v_{eff} | \phi_q \rangle, \tag{4.9}$$

$$\delta O_{pq} = \langle \delta \phi_p | \phi_q \rangle + \langle \phi_p | \delta \phi_q \rangle, \tag{4.10}$$

so that to first order, Eq. (4.8) transforms into

$$\sum_{pq} C_p^* C_q (\delta H_{pq} - \varepsilon \, \delta O_{pq} - \delta \varepsilon O_{pq}) = 0. \tag{4.11}$$

Using the normalization of $\psi_i(\vec{r})$, this yields the first-order change in the eigenvalue,

$$\delta \varepsilon_i = \sum_{pq} C_p^* C_q (\delta H_{pq} - \varepsilon_i \delta O_{pq})$$

$$= \sum_{pq} C_p^* C_q \left[\langle \delta \phi_p | H - \varepsilon_i | \phi_q \rangle + \langle \phi_p | H - \varepsilon_i | \delta \phi_q \rangle + \langle \phi_p | \delta \hat{T} + \delta v_{eff}(\vec{r}) | \phi_q \rangle \right], \tag{4.12}$$

allowing us to write

$$\mathbf{F}^a = \vec{F}^a_{HF} - \sum_i \sum_{pq} n_i C_p^* C_q \left[\langle \frac{d\phi_p}{d\vec{r}_a} | H - \varepsilon_i | \phi_q \rangle + \langle \phi_p | H - \varepsilon_i | \frac{d\phi_p}{d\vec{r}_a} \rangle + \langle \phi_p | \frac{d\hat{T}}{d\vec{r}_a} | \phi_q \rangle \right]$$

$$= \vec{F}^a_{HF} - \sum_i n_i \left[\langle \frac{d\psi_i}{d\vec{r}_a} | H - \varepsilon_i | \psi_i \rangle + \langle \psi_i | H - \varepsilon_i | \frac{d\psi_i}{d\vec{r}_a} \rangle + \langle \psi_i | \frac{d\hat{T}}{d\vec{r}_a} | \psi_i \rangle \right]. \tag{4.13}$$

The second term on the right side of the equation is known as the *incomplete basis set* (IBS) correction to the HF force. [53, 87, 7] If the first-order changes in the basis functions can be expressed in terms of the original basis functions,

$$|\delta\phi_p\rangle = \sum_j |\phi_j\rangle\langle\phi_j|\delta\phi_p\rangle, \tag{4.14}$$

i.e. if the basis set is complete, then the IBS correction to the HF force will vanish. This can also occur if the basis functions are not dependent upon the positions of the atoms, as in the case of planewaves.

In the LAPW method, the basis functions *are* dependent upon the atomic positions, which complicates the formalism significantly. IBS forces [125] come from both the core states and the valence states. The first-order change in the sum of the core eigenvalues for atom a is

$$\sum_i n_i\,\delta\varepsilon_i = \int \rho_c^a(\vec{r})\left[\delta v_{eff}(\vec{r}) + \nabla v_{eff}(\vec{r})\cdot\delta\vec{r}_a\right]d\vec{r}, \tag{4.15}$$

where $\rho_c^a(\vec{r})$ is the spherical core charge density of atom a, resulting from core states which satisfy a spherical muffin-tin version of the Kohn-Sham equations. The core state correction to the HF force is thus

$$
\begin{aligned}
\vec{F}_{core}^a &= -\frac{1}{\delta\vec{r}_a}\left[\sum_i n_i\,\delta\varepsilon_i - \int \rho_c^a(\vec{r})\,\delta v_{eff}(\vec{r})\,d\vec{r}\right]\\
&= -\int \rho_c^a(\vec{r})\,\nabla v_{eff}(\vec{r})\,d\vec{r}.
\end{aligned}
\tag{4.16}
$$

The valence state correction is given by

$$
\begin{aligned}
\vec{F}_{val}^a &= -\frac{1}{\delta\vec{r}_a}\left[\sum_i n_i\,\delta\varepsilon_i - \int \rho_v^a(\vec{r})\,\delta v_{eff}(\vec{r})\,d\vec{r}\right]\\
&= -\sum_i n_i\left[\langle\frac{d\psi_i}{d\vec{r}_a}|H - \varepsilon_i|\psi_i\rangle + \langle\psi_i|H - \varepsilon_i|\frac{d\psi_i}{d\vec{r}_a}\rangle + \vec{D}_i^a\right],
\end{aligned}
\tag{4.17}
$$

where $\vec{\mathbf{D}}_i^a$ is

$$\vec{\mathbf{D}}_i^a = \oint \left[\left(\psi_i^*(\vec{\mathbf{r}}) \, \hat{T} \, \psi_i(\vec{\mathbf{r}}) \right)_{MT} - \left(\psi_i^*(\vec{\mathbf{r}}) \, \hat{T} \, \psi_i(\vec{\mathbf{r}}) \right)_I \right] d\vec{\mathbf{S}}_a, \tag{4.18}$$

and takes account of the fact that the second derivative of the basis functions is not continuous at the surface of the muffin-tin sphere. The integral is over the surface of the sphere, while the MT and I subscripts indicate that the proper representation of the basis function is used within the muffin-tin and interstitial regions. When the wavefunction is represented in the LAPW basis, moving atom a changes the basis function only inside of its muffin-tin sphere, giving

$$\frac{\delta \phi_{\vec{G}}(\vec{\mathbf{r}})}{\delta \vec{\tau}_a} = i \left(\vec{\mathbf{k}}_i + \vec{\mathbf{G}} \right) \phi_{\vec{G}} - \nabla \phi_{\vec{G}}, \tag{4.19}$$

where the subscript $\vec{\mathbf{G}}$ has been used to replace $\vec{\mathbf{k}}_i + \vec{\mathbf{G}}$. Here, the variations in $u_l(\vec{\mathbf{r}})$ and $\dot{u}_l(\vec{\mathbf{r}})$ are neglected, as their determination is computationally impractical. Typically, this omission does not introduce significant errors in the force.[125] Finally, the total IBS correction is given by

$$\vec{\mathbf{F}}_{IBS}^a = \vec{\mathbf{F}}_{core}^a - \sum_i n_i \left[i \sum_{\vec{G}, \vec{G}'} (\vec{\mathbf{G}} - \vec{\mathbf{G}}') C_i^*(\vec{\mathbf{G}}) C_i(\vec{\mathbf{G}}') \langle \phi_{\vec{G}} | (H - \varepsilon_i) | \phi_{\vec{G}'} \rangle_{MT} \right.$$
$$\left. - \langle \nabla \psi_i | (H - \varepsilon_i) | \psi_i \rangle_{MT} - \langle \psi_i | (H - \varepsilon_i) | \nabla \psi_i \rangle_{MT} + \vec{\mathbf{D}}_i^a \right]. \tag{4.20}$$

If pseudopotentials are being used, then above term is computed using the ionic charge instead of the nuclear charge in $v_{eff}(\vec{\mathbf{r}})$. Once this is done, an additional IBS correction must be added:

$$\vec{\mathbf{F}}_{ps}^a = - \sum_i n_i \left[i \sum_{\vec{G}, \vec{G}'} (\vec{\mathbf{G}} - \vec{\mathbf{G}}') C_i^*(\vec{\mathbf{G}}) C_i(\vec{\mathbf{G}}') \langle \phi_{\vec{G}} | \hat{V}_{ps} | \phi_{\vec{G}'} \rangle_{MT} \right.$$
$$\left. - \langle \nabla \psi_i | \widetilde{V} | \psi_i \rangle_{MT} - \langle \psi_i | \widetilde{V} | \nabla \psi_i \rangle_{MT} + \langle \psi_i | \frac{\delta \widetilde{V}}{\delta \vec{\tau}_a} | \psi_i \rangle \right], \tag{4.21}$$

with $\widetilde{V}$ and $\hat{V}_{ps}$ being given by Eqs. 3.38 and 3.39, respectively.

4.2 Linear response theory

Many physical quantities of interest are related to first and higher order derivatives of the total energy with respect to displacements of the atoms or variations in applied fields. Many of these are directly measurable from experiment, such as the macroscopic dielectric, piezoelectric, and elastic constants, lattice constants, bulk modulus and phonon frequencies. Knowledge of the phonon frequencies and energy dispersions give access to the dynamical properties of materials, such as electron-phonon interactions, thermal expansion coefficients, and heat capacity. Of key interest in this dissertation is the issue of dynamical stability. Transitions from one structural symmetry to another can occur due to phonon-related instabilities, which themselves occur when a particular phonon distortion lowers the total energy. These are especially important within the soft-mode description for ferroelectric transitions.[21]

Within the Born-Oppenheimer approximation, the dynamical properties may be determined by examining the system's response to real-space periodic displacements of the atoms, which can be specified by a reciprocal space wavevector $\vec{q}$. At the present, there are basically two approaches for the calculation of the system's response to these phonon-like perturbations. In the first approach, known as the *direct approach*, small but finite phonon displacements are frozen in time, and the subsequent geometry is calculated just as it was for the unperturbed system. This "frozen phonon" technique was initially carried out successfully within the Kohn-Sham formalism by Wendel and Martin,[120] and later by others,[124, 65] establishing that the LDA density functional theory could accurately describe the dynamics of materials from first principles. The advantage to the direct approach is that no

additional programming is required. However, it is limited by the fact that periodic phonon-like displacements of the atoms can only be simulated using supercells whose dimensions are commensurate with the phonon wavelength. Because of the unfavorable computational scaling which accompanies the increased number of atoms, the number of wavevectors, $\vec{q}$, for which calculations can be performed is small, being mostly restricted to high-symmetry points in the Brillouin zone. Determination of the dynamical matrix at an arbitrary wavevector is highly desirable, as this will yield the full phonon dispersions. This can be done within the direct method, via an interpolation technique that depends upon the calculation of interplanar force constants.[116, 117, 33] Because standard LDA methods depend upon the Born-von Karman periodic boundary conditions, the effect of macroscopic electric fields cannot be addressed directly, but rather must be determined by introducing a very long-wavelength ($\vec{q} \rightarrow 0$) phonon distortion,[76] which is again limited by the use of large supercells.

In the second approach, only the linear response of the system to the perturbation is considered. This is not a significant hinderance, however, since many experimental quantities can be exactly expressed in terms of the first-order changes in the system. Early work in this area focused on using the inverse dielectric matrix,[86, 5] which is calculated from the results for the unperturbed ground state only. Since the effort spent in calculating the dielectric matrix is not significantly increased as one moves from a high-symmetry $\vec{q}$-point to one of minimal symmetry, the restriction to special wavevectors is removed, permitting access to the entire Brillouin zone. However, the drawbacks inherent in this method are significant. Since this is a

planewave-based method, it is not well-suited for materials with localized orbitals, and it quickly becomes complicated if other types of basis functions are used. Also, modern planewave calculations require the use of non-local pseudopotentials in order to achieve accurate results, and the inclusion of these effects within the dielectric matrix formalism is difficult. Finally, the required inversion of the dielectric matrix, as well as the complicated summation over the conduction bands at $\vec{k}$ points in the Brillouin zone, have made this technique computationally inferior to other approaches.

Later, Baroni, Giannozzi, and Testa [6] introduced a new scheme for studying the linear response of crystals, modeled after the Sternheimer theory for atomic polarizabilities.[102, 73] The main feature of this technique is that the electronic response is obtained via perturbation theory through iteration to self-consistency, rather than through the dielectric matrix. Assuming that the ground state problem has been solved, and we know the single-particle eigenvalues and eigenfunctions along with the density, a perturbation, specified by the parameter λ, is introduced into the system. The task is then to solve the new set of single-particle equations,

$$\left[H(\lambda) - \varepsilon_i(\lambda)\right] |\psi_i(\lambda)\rangle = 0, \tag{4.22}$$

subject to the orthonormalization condition for occupied states

$$\langle \psi_i(\lambda) | \psi_j(\lambda) \rangle = \delta_{ij}, \ \forall \ i, j \in \{occ\}. \tag{4.23}$$

Expanding the quantities appearing in the above pair of equations in powers of λ, keeping only first-order terms, leads to the Sternheimer equation [102]

$$(H^{(0)} - \varepsilon_i^{(0)}) | \psi_i^{(1)} \rangle = -(H^{(1)} - \varepsilon_i^{(1)}) | \psi_i^{(0)} \rangle, \tag{4.24}$$

and the constraint

$$\langle \psi_i^{(0)} | \psi_j^{(1)} \rangle + \langle \psi_i^{(1)} | \psi_j^{(0)} \rangle = 0, \ \forall \ i,j \in \{occ\}. \tag{4.25}$$

The first-order wavefunctions can be expanded in terms of the complete set of ground-state wavefunctions, making a distinction between the occupied (valence) and unoccupied (conduction) single-particle states,

$$| \psi_i^{(1)} \rangle = P_v | \psi_i^{(1)} \rangle + P_c | \psi_i^{(1)} \rangle \tag{4.26}$$

$$= \sum_v^{val} | \psi_v^{(0)} \rangle \langle \psi_v^{(0)} | \psi_i^{(1)} \rangle + \sum_c^{cond} | \psi_c^{(0)} \rangle \langle \psi_c^{(0)} | \psi_i^{(1)} \rangle \tag{4.27}$$

$$= \sum_v^{val} \alpha_{iv}^{(1)} | \psi_v^{(0)} \rangle + \sum_c^{cond} \alpha_{ic}^{(1)} | \psi_c^{(0)} \rangle. \tag{4.28}$$

Here, P_v and P_c are the projection operators for the valence and conduction manifolds, respectively. Substitution of Eq. (4.28) into Eq. (4.24) with an additional multipication on the left by a valence ground state reveals that $\alpha_{ii}^{(1)}$ cannot be determined uniquely. However, we do arrive at

$$\varepsilon_i^{(1)} = \langle \psi_i^{(0)} | H^{(1)} | \psi_i^{(0)} \rangle, \tag{4.29}$$

and

$$\alpha_{iv}^{(1)} = \frac{\langle \psi_v^{(0)} | H^{(1)} | \psi_i^{(0)} \rangle}{\varepsilon_i^{(0)} - \varepsilon_v^{(0)}}. \tag{4.30}$$

The normalization condition for occupied ground states dictates that the real parts of the $\alpha_{iv}^{(1)}$ coefficients (and $\alpha_{ii}^{(1)}$) must be zero. As a consequence, the contribution to $\psi_i^{(1)}$ from the valence manifold does not affect the charge density or the total energy. Furthermore, we may impose a unitary transformation on this contribution such that

it is fully aligned with $|\psi_i^{(0)}\rangle$. Once this is done, the indeterminate quality of $\alpha_{ii}^{(1)}$ allows us to force it to be zero, replacing the normalization condition with the stronger constraint:

$$\langle \psi_i^{(0)} | \psi_j^{(1)} \rangle = \alpha_{ji}^{(1)} = 0, \ \forall i,j \in \{occ\},$$

(4.31)

which is also known as the *parallel transport gauge*.[41] The task now is to find the projection of the first-order wavefunction within the conduction manifold. Thus, Eq. (4.24) is replaced by a modified Sternheimer equation

$$P_c \left(H^{(0)} - \varepsilon_i^{(0)}\right) P_c \, | \, \psi_i^{(1)} \rangle = -P_c \, H^{(1)} | \, \psi_i^{(0)} \rangle,$$

(4.32)

where $H^{(1)}$ is the first-order Kohn-Sham effective potential

$$v_{eff}^{(1)}(\vec{r}) = v_{ext}^{(1)}(\vec{r}) + \int \frac{\rho^{(1)}(\vec{r}')}{|\vec{r} - \vec{r}'|} d\vec{r}' + \rho^{(1)}(\vec{r}) \left. \frac{d\mu_{xc}}{d\rho(\vec{r})} \right|_{\rho = \rho^{(0)}(\vec{r})},$$

(4.33)

which is constructed from the general first-order charge density

$$\rho^{(1)}(\vec{r}) = \sum_i^{occ} \left[\psi_i^{(0)*}(\vec{r}) \, \psi_i^{(1)}(\vec{r}) + \psi_i^{(1)*}(\vec{r}) \, \psi_i^{(0)}(\vec{r}) \right].$$

(4.34)

The iterative procedure begins with $H^{(1)} = v_{ext}^{(1)}$, and Eq. (4.32) is solved for the first-order wavefunctions $\psi_i^{(1)}$. These are used to create the first-order charge density, which allows construction of a new $H^{(1)}$. This is reinserted into Eq. (4.32) and the process is iterated to self-consistency. Because the self-consistent response is obtained via iteration, as in the unperturbed problem, incorporation of non-local pseudopotentials is possible, as is the use of wavefunctions other than plane waves. Also, since the projection operator, P_c, is easily created from knowledge of only the valence states, there is no explicit sum over the conduction bands as in the

dielectric matrix method, and the computational scaling of the problem is on the same order as that of the direct approach. In this way, phonon-like perturbations can be performed independent of the specific wavevector. Finally, the effects of macroscopic electric fields can be studied by taking the long wavelength limit.[6, 38, 37] Gonze and Vigneron have extended this idea so that even higher-order energy corrections may calculated.[39, 41] They demonstrate that for density functional perturbation theory, the first n-order changes in the wavefunctions and charge density are sufficient for determination of the first $2n+1$ order changes in the total energy. Within the linear response regime, total energy may be determined up to third order:

$$E_{tot}(\lambda) = E^{(0)}[\rho^{(0)}] + \lambda E^{(1)}[\rho^{(0)}] + \lambda^2 E^{(2)}[\rho^{(0)}, \rho^{(1)}] + \lambda^3 E^{(3)}[\rho^{(0)}, \rho^{(1)}] \qquad (4.35)$$

Consequently, as mentioned before, the dynamical properties of materials can also be determined from the linear electronic response; specifically from the first-order forces (second-order energy).

4.2.1 Incommensurate periodic perturbations

The external perturbation of most interest to the present work is one corresponding to a periodic lattice displacement wave, or "frozen" phonon, specified by a reciprocal space wavevector $\vec{q}$:

$$w_{i\alpha}(\vec{R}) = w_{i\alpha}\, e^{i\vec{q}\cdot\vec{R}}, \qquad (4.36)$$

where α is a Cartesian index, and w is the displacement from equilibrium of atom i in the unit cell specified by $\vec{R}$. The wavelength of this perturbation is, in general, not commensurate with the real-space lattice. The first-order change in the self-consistent

potential due to this perturbation is given by Eq. (4.33) with the first-order external potential

$$v_{ext}^{(1)}(\vec{r}) = -\sum_i \sum_{\vec{R}} \frac{Z_i \, \vec{w}_i(\vec{R}) \cdot (\vec{r} - \vec{r}_i - \vec{R})}{|\vec{r} - \vec{r}_i - \vec{R}|^3}. \qquad (4.37)$$

Notice that this potential is non-Hermitian unless $2\vec{q}$ is a reciprocal lattice vector. This does not present a fundamental problem in the linear response method, as the electronic response to the sum of this potential and it's Hermitian conjugate is equal to the sum of the response to each perturbation separately. Translation by a lattice vector results in

$$v_{ext}^{(1)}(\vec{r} + \vec{R}) = e^{i\vec{q}\cdot\vec{R}} \, v_{ext}^{(1)}(\vec{r}), \qquad (4.38)$$

which suggests that the perturbing external potential can be written as a product of a cell-dependent phase factor and a potential which has the periodicity of the original ground state problem. For the moment, let us suppose that this feature is a property of the entire first-order potential as it appears in Eq. (4.33). Since we have chosen to use the parallel transport gauge, the first-order valence wavefunction can be written entirely in terms of the conduction ground states,

$$\psi_{v,\vec{k}}^{(1)}(\vec{r}) = \sum_c^{cond} \alpha_{c,\vec{k}}^{(1)} \psi_{c,\vec{k}}^{(0)}(\vec{r}), \qquad (4.39)$$

with

$$\alpha_{c,\vec{k}}^{(1)} = \frac{\langle \psi_{c,\vec{k}}^{(0)} | \, v_{eff}^{(1)} \, | \psi_{v,\vec{k}}^{(0)} \rangle}{\varepsilon_{v,\vec{k}}^{(0)} - \varepsilon_{c,\vec{k}}^{(0)}}. \qquad (4.40)$$

Now, upon translation by a lattice vector, the Bloch form of the ground state wavefunctions implies that

$$\psi_{v,\vec{k}}^{(1)}(\vec{r} + \vec{R}) = e^{i(\vec{k}+\vec{q})\cdot\vec{R}} \, \psi_{v,\vec{k}}^{(1)}(\vec{r}). \qquad (4.41)$$

Constructing the charge density according to Eq. (4.34), utilizing time reversal symmetry, we arrive at

$$\rho^{(1)}(\vec{r}) = 2\sum_{v,\vec{k}}^{occ} \psi_{v,\vec{k}}^{(0)*}(\vec{r})\,\psi_{v,\vec{k}}^{(1)}(\vec{r}). \qquad (4.42)$$

where summation is over the occupied bands. Translation by a lattice vector reveals

$$\rho^{(1)}(\vec{r}+\vec{R}) = e^{i\vec{q}\cdot\vec{R}}\,\rho^{(1)}(\vec{r}), \qquad (4.43)$$

which is consistent with the assumption that the entire first-order potential has the same periodicity as the phonon perturbation. Since it has now been shown possible to write the potential and wavefunctions as a cell-dependent phase factor times lattice periodic functions, this phase can be factored out of Eq. (4.32), and the incommensurate problem is effectively mapped onto one which has the periodicity of the unperturbed system. This approach is known as the *traveling-wave representation*.

4.2.2 LAPW linear response method

The implementation of the linear response method within the LAPW formalism was first performed by Yu and Krakauer,[126] permitting direct calculation of dynamical matrices at a general wavevector. In the LAPW linear response method (LAPW-LR), the first-order wavefunction results from the changes in the LAPW basis functions and the variational coefficients of Eq. (2.86),

$$\psi_{n,\vec{k},\vec{q}}^{(1)}(\vec{r}) = \sum_{\vec{G}}\left[C_{n,\vec{k}}^{(1)}(\vec{G})\,\phi_{\vec{k}+\vec{q}+\vec{G}}^{(0)}(\vec{r}) + C_{n,\vec{k}}^{(0)}(\vec{G})\,\phi_{\vec{k}+\vec{G}}^{(1)}(\vec{r})\right]. \qquad (4.44)$$

Since the plane wave representation used outside of the muffin-tin spheres is not dependent upon the atomic positions, the only contribution to $\phi_{\vec{k}+\vec{G}}^{(1)}(\vec{r})$ comes from

the muffin-tin spheres. As mentioned previously in the description of the LAPW force formalism,[125] the relaxation of the muffin-tin orbitals $u_l(r)$ and $\dot{u}_l(r)$ is rather small and can be neglected. Thus, the linear-order change in the basis function is

$$\phi^{(1)}_{\vec{k}+\vec{G}}(\vec{r}) = \frac{\partial \phi^{(0)}_{\vec{k}+\vec{G}}}{\partial \vec{w}} \cdot \vec{w}$$

$$= [i(\vec{k} + \vec{G}) - \nabla] \, \phi^{(0)}_{\vec{k}+\vec{G}}(\vec{r}) \cdot \vec{w}. \tag{4.45}$$

Here, $\vec{w}$ is used to represent all of the ionic displacements without reference to individual ions. The first term in brackets results from the gradient of a term which appears when the basis functions and their first derivatives in Eq. (3.2) are matched at the muffin-tin boundary. This was encountered before in Eq. (4.19). The problem of determining the first-order wavefunctions now becomes one of finding the first-order variational coefficients. Yu and Krakauer showed that these could be determined from the following matrix equation:

$$\sum_{\vec{G}'} \left\langle \phi^{(0)}_{\vec{k}+\vec{q}+\vec{G}} \left| (\varepsilon^{(0)}_{n,\vec{k}} - H^{(0)}) \right| \phi^{(0)}_{\vec{k}+\vec{q}+\vec{G}'} \right\rangle C^{(1)}_{n,\vec{k}}(\vec{G}')$$

$$= \sum_{\vec{G}'} \left[\left\langle \phi^{(0)}_{\vec{k}+\vec{q}+\vec{G}} \left| V^{(1)}_{SCF} \right| \phi^{(0)}_{\vec{k}+\vec{G}'} \right\rangle + \left\langle \frac{\partial \phi^{(0)}_{\vec{k}+\vec{q}+\vec{G}}}{\partial \vec{w}} \cdot \vec{w}^* \left| (H^{(0)} - \varepsilon^{(0)}_{n,\vec{k}}) \right| \phi^{(0)}_{\vec{k}+\vec{G}'} \right\rangle_{MT} \right. \tag{4.46}$$

$$+ \left\langle \phi^{(0)}_{\vec{k}+\vec{q}+\vec{G}} \left| (H^{(0)} - \varepsilon^{(0)}_{n,\vec{k}}) \right| \frac{\partial \phi^{(0)}_{\vec{k}+\vec{G}'}}{\partial \vec{w}} \cdot \vec{w} \right\rangle_{MT}$$

$$\left. + \delta \left\langle \phi^{(0)}_{\vec{k}+\vec{q}+\vec{G}} \left| \hat{T} \right| \phi^{(0)}_{\vec{k}+\vec{G}'} \right\rangle \right] C_{n,\vec{k}}(\vec{G}'),$$

where the suffix MT indicates integration over the muffin-tin spheres only. The last term on the right-hand side represents the change in kinetic energy which comes about from the discontinuity in the second derivative of the basis functions:

$$\delta \left\langle \phi^{(0)}_{\vec{k}+\vec{q}+\vec{G}} \left| \hat{T} \right| \phi^{(0)}_{\vec{k}+\vec{G}'} \right\rangle$$

$$= \sum_i \oint \left[\phi^{(0)*}_{\vec{k}+\vec{q}+\vec{G}}(\vec{r}) \, \hat{T} \, \phi^{(0)}_{\vec{k}+\vec{G}'}(\vec{r})|_{MT} - \phi^{(0)*}_{\vec{k}+\vec{q}+\vec{G}}(\vec{r}) \, \hat{T} \, \phi^{(0)}_{\vec{k}+\vec{G}'}(\vec{r})|_I \right] dS_i \cdot \vec{w}_i, \tag{4.47}$$

where the subscript I denotes integration in the interstitial region, and the index i refers to individual muffin-tin spheres. The matrix equation above only deals with the occupied states of the unperturbed system, and one should take notice that each matrix element within this equation has the full periodicity of the lattice, despite the use of the traveling wave representation. The first-order potential only couples states that differ by the wavevector $\vec{q}$. An iterative diagonalization technique is used to solve Eq. (4.46), and the resulting first-order variational coefficients are used along with Eq. (4.45) to create the first-order wavefunctions as per Eq. (4.44). These, in turn, go into Eq. (4.42) to yield the first-order charge density, which is used to calculate a new first-order self-consistent potential. This new $V_{SCF}^{(1)}$ is inserted into Eq. (4.46), and the whole process is iterated to self-consistency.

4.3 The dynamical matrix

The dynamical properties of a specific material can be derived from knowledge of the *dynamical matrix*. The elements of the dynamical matrix are determined from the first-order forces, which result from atomic displacements about the equilibrium positions. In the harmonic approximation, the second-order energy is given by

$$E^{(2)} = \frac{1}{2} \sum_{\vec{R}, \vec{R}'} \sum_{i\alpha, j\beta} C_{i\alpha, j\beta}(\vec{R} - \vec{R}') \, w_{i\alpha}(\vec{R}) \, w_{i\beta}(\vec{R}'), \qquad (4.48)$$

where the *force constants*, $C_{i\alpha, j\beta}(\vec{R} - \vec{R}')$, are the second derivatives of the total energy with respect to the displacements specified by w. The subscripts i and j label individual atoms, while α and β denote the regular Cartesian directions. The

first-order forces are defined in terms of the force constants by

$$F_{i\alpha}^{(1)}(\vec{R}) = -\frac{\partial E^{(2)}}{\partial w_{i\alpha}(\vec{R})} = -\sum_{\vec{R}'}\sum_{j\beta} C_{i\alpha,j\beta}(\vec{R}-\vec{R}')\,w_{j\beta}(\vec{R}') \qquad (4.49)$$

$$= M_i\,\ddot{w}_{i\alpha}(\vec{R}). \qquad (4.50)$$

The instantaneous displacements and accelerations of the atoms in a general phonon are given by

$$w_{i\alpha}(\vec{R},t) = w_{i\alpha}\,e^{i(\vec{q}\cdot\vec{R}-\omega(\vec{q})\,t)}$$

$$\ddot{w}_{i\alpha}(\vec{R},t) = -\omega^2(\vec{q})\,w_{i\alpha}\,e^{i(\vec{q}\cdot\vec{R}-\omega(\vec{q})\,t)} \qquad (4.51)$$

$$= -\omega^2(\vec{q})\,w_{i\alpha}(\vec{R},t),$$

with $w_{i\alpha} = w_{i\alpha}(\vec{R}=0)$, such that for a "frozen" phonon of the form Eq. (4.36), the first-order force becomes

$$F_{i\alpha}^{(1)}(\vec{R}) = F_{i\alpha}^{(1)}\,e^{i\vec{q}\cdot\vec{R}}, \qquad (4.52)$$

where

$$F_{i\alpha}^{(1)} = -\omega^2(\vec{q})\,M_i\,w_{i\alpha} \qquad (4.53)$$

$$= -\sum_{\vec{R}'}\sum_{j\beta} C_{i\alpha,j\beta}(\vec{R}')\,w_{j\beta}\,e^{i\vec{q}\cdot\vec{R}'} \qquad (4.54)$$

$$= -\sum_{j\beta} \widetilde{C}_{i\alpha,j\beta}(\vec{q})\,w_{j\beta}. \qquad (4.55)$$

Here, the $\widetilde{C}_{i\alpha,j\beta}(\vec{q})$ are the Fourier transforms of the real-space force constants, and $\omega(\vec{q})$ is the wavevector-dependent phonon frequency. Now, the reciprocal-space force constants at a specific wavevector $\vec{q}$ may be determined by at most $3N_a$ lattice distortions, where N_a is the number of atoms in the primitive unit cell. For example, setting

all but one of the $w_{j\beta}$ to zero immediately yields the contents of a single column of the $3N_a \times 3N_a$ force constant matrix in reciprocal space, provided the first-order forces on the atoms are known.

As was mentioned earlier, the second-order energy can be determined from the unperturbed and first-order wavefunctions and charge densities.[39] It is therefore possible to find the first-order forces within the linear response approach. Yu and Krakauer [126] have described in extensive detail how the *first-order* forces may be calculated within the LAPW formalism, so it will not be repeated here. The sought-after dynamical matrix is simply related to the reciprocal-space force constants by the masses of the atoms:

$$\tilde{D}_{i\alpha,j\beta}(\vec{q}) = \frac{\tilde{C}_{i\alpha,j\beta}(\vec{q})}{\sqrt{M_i \, M_j}}. \tag{4.56}$$

Thus, the dynamical equation becomes

$$-\omega^2(\vec{q}) \sqrt{M_i} \, w_{i\alpha} = -\sum_{j\beta} \tilde{D}_{i\alpha,j\beta}(\vec{q}) \, w_{j\beta} \sqrt{M_j}, \tag{4.57}$$

leading to

$$\sum_{j\beta} \left[\tilde{D}_{i\alpha,j\beta}(\vec{q}) - \omega^2(\vec{q})\delta_{ij}\,\delta_{\alpha\beta} \right] \sqrt{M_j}\,w_{j\beta} = 0. \tag{4.58}$$

The $3N_a$ phonon frequencies, $\omega(\vec{q})$, and the $3N_a$ normal modes (eigenvectors) $\vec{\xi}(\vec{q})$, are then determined by solving the standard secular equation

$$\left| \tilde{D}_{i\alpha,j\beta}(\vec{q}) - \omega^2(\vec{q})\delta_{ij}\,\delta_{\alpha\beta} \right| = 0. \tag{4.59}$$

Any structural symmetries that exist in the reciprocal-space Brillouin zone, (i.e. a four-fold rotation axis) will appear in the structure of the dynamical matrices, and will

help reduce the actual number of distortions needed to compute the entire dynamical matrix.

4.4   Macroscopic fields

The linear response method as formulated above offers the advantage of determining the system's response to macroscopic electric fields. When placed between conductive plates having a potential difference across them, the external potential acting on the system is replaced by

$$v_{ext}^{(1)}(\vec{r}) = \lim_{q \to 0} \frac{E_{ext} \, e^{i\vec{q}\cdot\vec{r}}}{i|q|}. \tag{4.60}$$

This form results by Fourier transforming the external field and applying the traveling-wave representation in the long wavelength limit, allowing us to utilize the results for phonon-like perturbations. In practice however, a small but finite wavevector is used, which is typically only 1% of a reciprocal lattice vector, and the response is computed directly. The electronic contribution to the dielectric tensor is determined from

$$\hat{q} \cdot \epsilon_{\infty} \cdot \hat{q} = \lim_{q \to 0} \left[1 - \frac{v_{ind}(\vec{q})}{v_{total}(\vec{q})} \right], \tag{4.61}$$

where

$$v_{ind} = v_{total} - v_{ext} \tag{4.62}$$

is the induced potential in the system when the external field is present.[1] In the perovskites, the cubic paraelectric phase will have an isotropic (diagonal) dielectric tensor, while in the ferroelectric phases, there may be longitudinal, transverse, and possibly off-diagonal elements as well, depending upon the crystal symmetries of these phases.

4.4.1   The Born effective charge tensor

In their comprehensive work on the theory of lattice vibrations, Born and Huang [10] demonstrated that during a sublattice distortion, the normal ionic charges of the atoms may be replaced by effective charge values which describe the modifications to the macroscopic polarization generated by the distortion. Because these effective charges are dynamical in nature, their values are dependent upon the direction of the motion, and are thus represented by a tensor. The *Born effective charge* is defined as the linear change in the macroscopic polarization created by an atomic displacement:

$$Z^*_{i,\alpha\beta} = Z_i + \Omega \frac{\partial P_\alpha}{\partial w_{i,\beta}}\bigg|_{E=0}. \tag{4.63}$$

Here, Z_i is the normal ionic charge on atom i, and α and β are Cartesian directions. To determine the Born effective charges, the polarization is evaluated using the Fourier transform of Poisson's equation

$$\nabla \cdot \vec{P} = -\rho^{(1)}(\vec{r})$$

$$\nabla \cdot \left[\sum_{\vec{G}} \vec{P}(\vec{q} + \vec{G}) \, e^{i(\vec{q}+\vec{G})\cdot\vec{r}} \right] = -\sum_{\vec{G}} \rho^{(1)}(\vec{q} + \vec{G}) e^{i(\vec{q}+\vec{G})\cdot\vec{r}} \tag{4.64}$$

$$\vec{P} = \hat{q} \cdot \vec{P}(0) = \lim_{\vec{q}\to 0} \frac{i}{q} \rho^{(1)}(\vec{q}).$$

Ideally, the zone-center dynamical matrix should allow for pure translations as eigenvectors, requiring that the following "acoustic sum rule" [86] be obeyed:

$$\sum_j \widetilde{C}_{i\alpha,j\beta}(\vec{q} = 0) = 0, \tag{4.65}$$

from which charge neutrality is implied:

$$\sum_i Z^*_{i,\alpha\beta} = 0. \tag{4.66}$$

In practice, this requirement is not satisfied ideally due to the finite k-point sampling used for the Brillouin zone integrations. Increasing the density of the sampling typically improves the agreement.[126]

4.4.2 Zone center mode splitting

In non-metals and polar semiconductors, a long wavelength $\vec{q} \to 0$ optical phonon mode creates a non-vanishing macroscopic polarization $\vec{P}$ due to the oppositely directed motions of oppositely charged ions. In general, this will necessarily be accompanied by a macroscopic displacement field and a macroscopic electric field as well;

$$\vec{D} = \vec{E} + 4\pi\vec{P}. \tag{4.67}$$

In crystals having cubic symmetry, these fields are all parallel to one another, and the effective local electric field at the positions of the ions can be written as [54]

$$\vec{E}^{local}(\vec{r}) = \vec{E}(\vec{r}) + \frac{4\pi\vec{P}(\vec{r})}{3}. \tag{4.68}$$

In a purely longitudinal optic (LO) phonon, the polarization is parallel to the wavevector, requiring that

$$\vec{D} = 0, \quad \text{or} \quad \vec{E} = -4\pi\vec{P} \tag{4.69}$$

However, in the case of a purely transverse optic (TO) phonon, the polarization is perpendicular to the wavevector, requiring that $\vec{E} = 0$. Thus, for each type of phonon,

there are differing *local* fields at each atom:

$$\vec{E}^{local}(\vec{r}) = -\frac{8\pi\vec{P}(\vec{r})}{3} \qquad \text{longitudinal,} \qquad (4.70)$$

$$\vec{E}^{local}(\vec{r}) = \frac{4\pi\vec{P}(\vec{r})}{3} \qquad \text{transverse.} \qquad (4.71)$$

As a consequence, the local field enhances the short-range restoring forces in the longitudinal case by reducing the polarization, and weakens these same forces in the transverse case by boosting the polarization. If the increase in polarization is large enough, a ferroelectric instability can develop due to a sign change in the restoring force. In a TO mode, the presence of a non-vanishing polarization and a vanishing macroscopic electric field implies that $\epsilon \rightarrow \infty$, while a LO mode must necessarily have $\epsilon = 0$. The relationship between the dielectric constant and the phonon frequency is given by the dielectric dispersion formula [10] for the interaction of phonons with transverse electromagnetic waves:

$$\epsilon = \epsilon_\infty + \frac{\epsilon_0 - \epsilon_\infty}{1 - (\omega/\omega_0)^2}, \qquad (4.72)$$

where ϵ_0 is the static or low frequency $(\omega \ll \omega_0)$ dielectric constant, ϵ_∞ is the high-frequency $(\omega \gg \omega_0)$ or electronic dielectric constant. The ω_0 term is the infrared dispersion frequency at which the dielectric constant diverges. Thus, $\omega_0 = \omega_{TO}$, since we've already seen that $\epsilon \rightarrow \infty$ for a TO phonon. Similarly, for a LO phonon, Eq. (4.72) reduces to the Lyydane-Sachs-Teller relation

$$\omega_{LO}^2 = \omega_{TO}^2 \frac{\epsilon_0}{\epsilon_\infty}. \qquad (4.73)$$

So the effect of the long-range Coulomb interactions on zone-center phonons is to split the degenerate frequencies such that the LO mode is elevated relative to the TO mode.

To account for the macroscopic field effects in the linear response calculations at the zone-center, proper treatment of the dynamical matrix elements is necessary. Although such effects are automatically included in the linear response approach at finite wavevectors, calculations for zone-center phonons are performed by breaking the dynamical matrix into the sum of two terms:[10, 86]

$$\tilde{D}^{LO}_{i\alpha,j\beta} = \tilde{D}^{TO}_{i\alpha,j\beta} + \tilde{D}^{na}_{i\alpha,j\beta}.$$ (4.74)

The first term corresponds to the motions of the atoms with no macroscopic electric field, and is analytic in $\vec{q}$, so that the Fourier transform to real space is easily obtained. It corresponds to the zone center TO mode. The second term is nonanalytic, and has the general form

$$\tilde{D}^{na}_{i\alpha,j\beta}(\vec{q}) = \frac{4\pi e^2}{\Omega \sqrt{M_i M_j}} \frac{(\mathbf{Z}^*_i \cdot \vec{q})_\alpha (\mathbf{Z}^*_j \cdot \vec{q})_\beta}{\vec{q} \cdot \epsilon_\infty \cdot \vec{q}},$$ (4.75)

where $\mathbf{Z}^*_i$ and M_i are the Born effective charge tensor and mass for atom i, Ω is the volume of the unit cell, and α, β are Cartesian indices.

4.5 Phonon dispersion curves

Since the phonon frequencies are dependent upon the wavevector, it is tremendously useful to determine the phonon dispersions throughout the Brillouin zone. Once acquired, this information may be used to predict thermodynamic properties such as the specific heat and will reveal any structural instabilities responsible for phase transitions such as those seen in ferroelectrics. Fortunately, a full mapping of the dispersion curves can be obtained through linear response calculations. First, the

dynamical matrix is computed on a uniform grid of q-points within the Brilloiun zone. Crystal symmetries reduce the actual number of mesh points at which calculations must be performed. The phonon frequencies at any wavevector are then obtained using an interpolation technique which properly accounts for the LO-TO splitting at the zone center. The long-range dipolar forces responsible for the splitting are determined analytically, via the Born effective charges and the dielectric tensor, using

$$C_{i\alpha,j\beta}^{dipole}(\vec{\mathbf{R}} - \vec{\mathbf{R}}') = \frac{(\mathbf{Z}_i^* \cdot \vec{\mathbf{q}})_\alpha (\mathbf{Z}_j^* \cdot \vec{\mathbf{q}})_\beta}{\vec{\mathbf{q}} \cdot \boldsymbol{\epsilon}_\infty \cdot \vec{\mathbf{q}}} \left[\frac{\delta_{\alpha\beta}}{r^3} - \frac{3r_\alpha r_\beta}{r^5} \right], \qquad (4.76)$$

where $\vec{\mathbf{r}} = \vec{\mathbf{R}} - \vec{\mathbf{R}}' + \vec{\tau}_j + \vec{\tau}_i$. Utilizing Ewald summation methods,[29, 30, 10, 37] this expression is Fourier transformed and subtracted from each $\tilde{D}_{i\alpha,j\beta}(\vec{\mathbf{q}})$ to obtain the reciprocal-space short-range force constants on the grid.[40] These short-range force constants undergo Fourier transformation to generate interpolation coefficients, and the inverse transform is performed to give their reciprocal space values at any desired wavevector. Recombination with the analytic dipolar term results in a dynamical matrix at an arbitrary q-point.

Chapter 5

Structural instabilities in SrTiO$_3$

5.1   Introduction

The temperature-dependent phase transitions of SrTiO$_3$ differ from similar materials such as KNbO$_3$, BaTiO$_3$, and PbTiO$_3$, which undergo ferroelectric (FE) phase transitions as the temperature is lowered. Although SrTiO$_3$ possesses the same high-temperature cubic perovskite structure, its behavior is markedly different. Despite early indications of a FE transition close to 0 K, no such transition has yet been observed without the application of external stress [105] or an external electric field.[50] However, it does undergo an antiferrodistortive (AFD) transition from cubic to tetragonal symmetry as the temperature drops below about 105 K. The AFD distortion is associated with the softening of the lowest frequency triply degenerate Brillouin-zone corner (R-point) vibrational mode R$_{25}$.[104, 32, 94] The possibility of a transition of this type was first postulated by Cochran and Zia [20] during their extension of the original zone-center soft mode concept from diatomic cubic materials to the
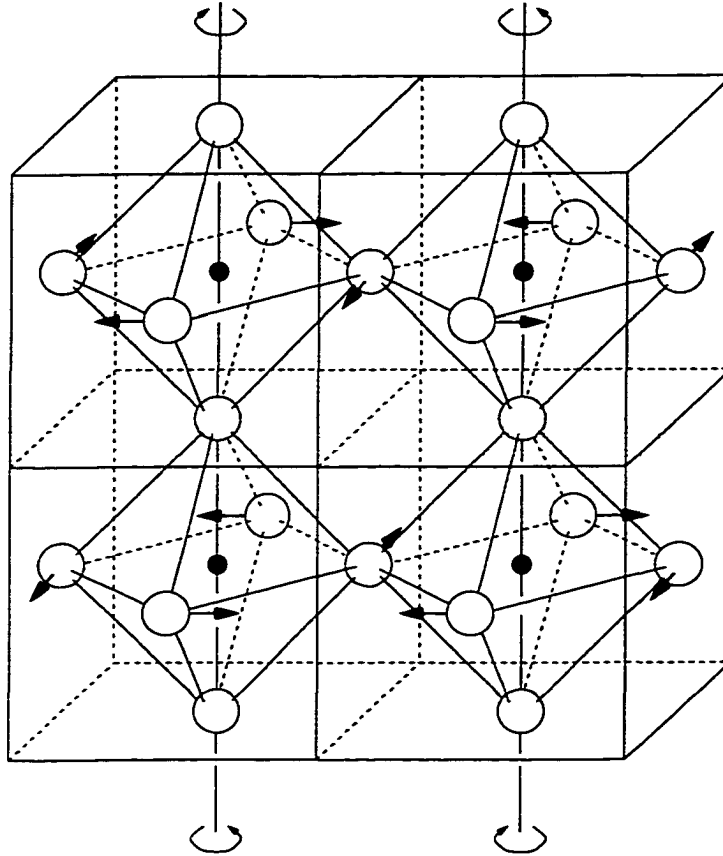
Figure 5.1: The R_{25} AFD vibration mode in perovskite $SrTiO_3$. The stationary titanium atoms rest inside the oxygen octahedra. Strontium atoms are not shown.

perovskites. $SrTiO_3$ was the first material to provide conclusive evidence for a *zone-boundary* soft mode transition. In the R_{25} mode, the TiO_6 octahedra rotate about parallel Cartesian axes, but in opposite directions in all neighboring cells as illustrated in Fig. 5.1 and Fig. 5.2. The new tetragonal structure does not differ too greatly from the paraelectric phase, assuming a small distortion c/a of 1.0005 after the oxygen octahedra move into their new rotated positions. This occurs without any apparent change in cell volume per formula unit.[72] The new symmetry is D_{4h}^{18} and

Figure 5.2: The rotation direction of the octahedra in the R_{25} mode alternates in adjacent cells. The axis of rotation is out of the page.

Figure 5.3: The new primitive cell in the tetragonal phase. Former lattice vectors are indicated by broken lines.

represents a cell-doubling to an unusual primitive cell shown in Fig. 5.3.

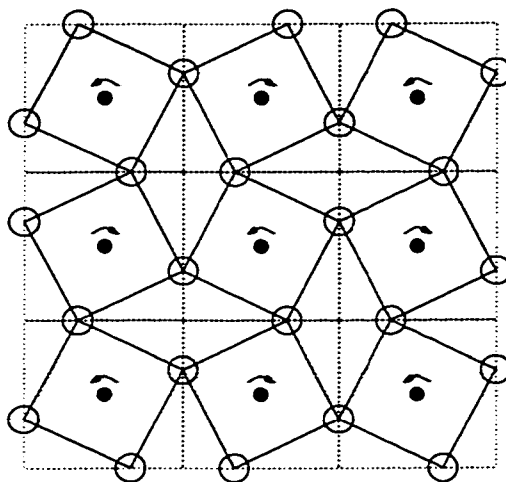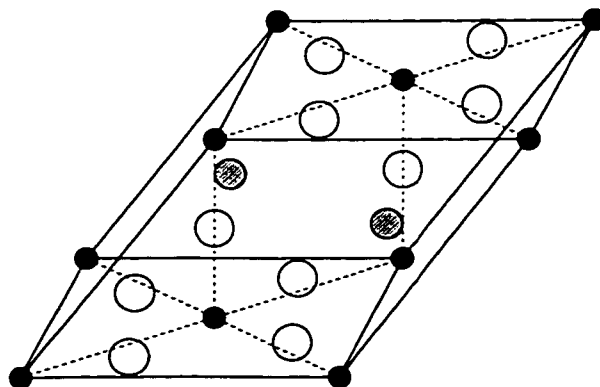With decreasing temperature, the static dielectric constant ϵ begins to rise with Curie-Weiss behavior, indicative of a FE transition at about 35 K (see Fig. 1.6). This transition never occurs however, and instead, ϵ saturates at values in excess of 10^4 at the lowest temperatures.[114, 90, 82, 106] This behavior was well known prior to the explanation of the AFD transition, and together with neutron scattering experiments,[28] was considered to be strong evidence in support of the soft mode concept.[21] Moreover, electric-field-induced Raman scattering shows a softening of the lowest frequency zone-center TO phonon mode as the temperature approaches absolute zero, [123] which is also a characteristic feature of an incipient displacive-type FE transition. However, the failure of $SrTiO_3$ to become ferroelectric has thus far eluded definitive explanation. It has been suggested that the long-range FE order in $SrTiO_3$ is suppressed by the zero-point motion of the Ti atoms, resulting in a quantum paraelectric phase.[82, 83] More recently, Zhong and Vanderbilt have performed path-integral Monte Carlo (MC) simulations, using an ab-initio effective Hamiltonian, in which the inclusion of quantum fluctuations is found to suppress the FE transition entirely.[130] In view of the two different types of soft phonons in $SrTiO_3$, the R-point mode associated with the AFD instability and the Γ-point mode associated with the incipient FE instability, it is of great interest to fully characterize the wavevector dependence of these instabilities, and to determine whether or not the tetragonal phase is energetically stable against FE-type distortions. In the present work, a full mapping of all of the structural instabilities of $SrTiO_3$ in the ideal cubic phase is presented. In addition, the interaction between the AFD and FE instabilities is

examined in the tetragonal phase, using frozen-phonon total energy calculations.

5.2 Method

The first-principles LAPW method described in Chapter 3 was used to determine the ground state properties of $SrTiO_3$, and the first-principles linear response method of Chapter 4 was carried out within the LAPW formalism to determine the phonons and wavevector dependence of structural instabilities.[126, 6] The linear response approach also permits the effects of external macroscopic electric fields to be studied. The electronic contribution to the dielectric constant, ϵ_∞, the macroscopic polarization, and the Born effective charges can also be found using the procedures detailed in Chapter 4. In order to map the phonon dispersion curves throughout the Brillouin zone (BZ), the dynamical matrix is first computed on a uniform grid of q-points. The interpolation procedure described in Chapter 4 is used to compute the dynamical matrix at an arbitrary wavevector $\vec{q}$. The signature of any structural instabilities at a given wavevector is that some of the eigenvalues of the dynamical matrix are negative (corresponding to imaginary phonon frequencies).

A uniform 6×6×6 special k-point mesh was used for the BZ integrations.[81] Pseudopotentials of the Kerker form [58] were used to exclude the tightly-bound core states. The higher-lying $Sr(4s,4p)$ and $Ti(3s,3p)$ semi-core states were also pseudized, and were included in a two-window variational calculation. About 570 LAPW basis functions were used at each k-point, corresponding to a kinetic energy cut-off of about 19 Ry.

The FE instability results from a delicate cancellation of long-range dipolar in-

teractions, favoring the instability, and short-range repulsive chemical interactions that oppose it. Covalent interactions, manifested by Ti(3d) and O(2p) hybridization, soften the short-range repulsion, permitting the FE transition to occur. Increasing the volume reduces the repulsion, enhancing the FE instability, while decreasing volume increases the short-range repulsion, and opposes the instability. The treatment of these different interactions requires accurate self-consistent calculations. The typically 1% LDA underestimates of equilibrium lattice parameter can eliminate or suppress the ferroelectric instabilities,[26, 75] and the use of generalized gradient approximations (GGA) may not improve the agreement, since the GGA can overcorrect by the same amount.[75] The weighted density approximation (WDA) is superior to both LDA and GGA in these materials,[75] but is much less computationally efficient even for planewave-based methods. However, good agreement with experiment is usually obtained when LDA calculations are performed at the experimental volume. Using the Wigner [121] formula for the LDA exchange-correlation potential is a simple alternative that has corrected the volume in $KNbO_3$ and yielded good agreement with other experimental measurements,[127, 112] and this form was used here.

To find the dynamical matrix over the entire BZ, a uniform $6\times6\times6$ grid of q-points was used, and the dynamical matrix was calculated at the 20 resulting irreducible q-points shown in Fig. 5.4. Due to the reduced symmetry of q-dependent perturbations induced by atomic distortions, the first-order change in the Kohn Sham wavefunctions must in general be calculated at many more k-points than in the smaller irreducible BZ-wedge of the high-symmetry cubic reference phase. Moreover, the determination of all elements of the dynamical matrix at a general q-point requires in
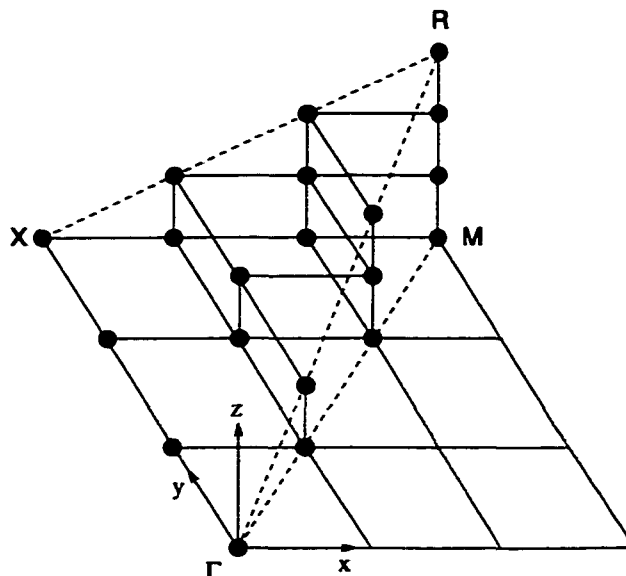
Figure 5.4: The wavevectors of the symmetry-reduced Brillouin zone wedge at which the dynamical matrix was calculated.

general 3N separate self-consistent calculations. As a result, the full calculation of the dynamical matrix is very time consuming, so it was prohibitive to do this at more than one volume. Instead, additional investigations of the volume dependence and the interaction between structural instabilities were carried out for selected combinations of representative eigenmodes using LAPW total energy frozen-phonon supercell calculations.

5.3 Results and Discussion

First, self-consistent total-energy calculations were performed in the cubic phase to determine the equilibrium bulk properties. The ground state wavefunctions produced were later used in the linear response calculations. As mentioned, the use of the

Table 5.1: Calculated equilibrium properties and Born effective charges for SrTiO$_3$.

	Lattice parameter a (a.u.)	Bulk modulus (GPa)
Present	7.412	190
PW^a	7.303	200
Exp.^{b}	7.380	179 ± 6

Born effective charge	Present	PW-BP^c
Z*(Sr)	2.55	2.54
Z*(Ti)	7.56	7.12
Z*(O)$_{\parallel}$	-5.92	-5.66
Z*(O)$_{\perp}$	-2.12	-2.00

^{a}Reference [59], planewave pseudopotential calculations.
^{b}Reference [80], 298 K.
^{c}Reference [128], planewave Berry phase calculation.

Wigner exchange-correlation potential yields an LDA equilibrium lattice parameter close to experiment as shown in Table 5.1, and this LDA value was used in the linear response calculations. Also shown in the upper part of Table 5.1 are the bulk moduli for the present work, along with the results of King-Smith and Vanderbilt,[59] as well as some experimental values. The bulk modulus was determined by fitting the total energy at differing volumes to Murnaghan's equation.[84] The electronic contribution to the dielectric constant, ϵ_∞, was found to be 6.63, which is about 28% larger than the value of 5.18 extrapolated from experiment. This overestimation of ϵ_∞ is typical

for the LDA.[38, 37] The lower portion of Table 5.1 compares the Born effective charge tensors (Z^*), which are diagonal in the cubic phase. The metal atoms have isotropic tensors, whereas the oxygen atoms have two distinct values, for displacements along and perpendicular to the Ti-O bonds, labeled $\parallel$ and $\perp$, respectively. The anomalously large values for $Z^*(\text{Ti})$ and $Z^*(\text{O})_\parallel$ created by dynamical modifications of the strong covalent interactions between these atoms have been noted before.[88, 128] Also listed are the results of planewave Berry phase calculations by Zhong, King-Smith, and Vanderbilt.[128] The respective theoretical lattice parameters were used in both calculations, and the results appear to be generally in good agreement. The acoustic sum rule for the linear response Born effective charges is satisfied to within 0.05, reflecting the adequacy of the 6×6×6 special BZ k-point sampling.

Table 5.2 compares theoretical and experimental zone-center and zone-corner phonon frequencies. Unstable transverse optic (TO) modes with imaginary frequencies are found at the Γ-point and at the R point. The Γ-point longitudinal optic (LO) mode frequencies are obtained from Eq. (4.74). The results of Zhong, King-Smith, and Vanderbilt were determined using $\epsilon_\infty = 5.18$ extracted from experiment,[128] whereas a larger calculated value of 6.63 was used for the present work. To determine what effect ϵ_∞ has on the LO mode eigenvalues, a second calculation was performed using $\epsilon_\infty = 5.18$. The highest LO mode is most sensitive, changing from 751 to 832 cm^{-1}.

The calculated phonon dispersion curves for the cubic phase have been plotted along high-symmetry directions in Fig. 5.5. The ΓX, ΓM, and ΓR lines correspond to the $\langle 100 \rangle$, $\langle 110 \rangle$, and $\langle 111 \rangle$ directions, respectively. The symmetry character of the modes at the zone-center and zone-boundaries has been labeled according to

Table 5.2: A comparison of linear response (LR) phonon frequencies (in cm^{-1}) with theoretical and experimental values at high-symmetry points.

Γ-point	Longitudinal modes				Transverse modes			
	Γ_{15}	Γ_{25}	Γ_{15}	Γ_{15}	Γ_{15}	Γ_{15}	Γ_{25}	Γ_{15}
LR	146	219	439	751	100i	151	219	522
LR[a]			[449]	[832]				
PW[b]	158		454	829	41i	165		546
90 K[c]	170	265			42	170	265	
296 K[d]	175		460	830	88	178		543
297 K[e]	169	265	457	823	92	169	265	547

R-point	R_{25}	R_{15}	R_{15}	$R_{25'}$	R_{25}	R_{15}	R_{15}	$R_{25'}$
LR	86i	122	417	426	86i	122	417	426
297 K[4]	52	145	450	473	52	145	450	473

[a]Computed using ϵ_∞=5.18, extracted from experiment; other LO frequencies did not change.
[b]Reference [128], planewave pseudopotential frozen-phonon calculations.
[c]Reference [28], neutron scattering results.
[d]Computed in Reference [28] from infrared reflectivity data of Reference [100].
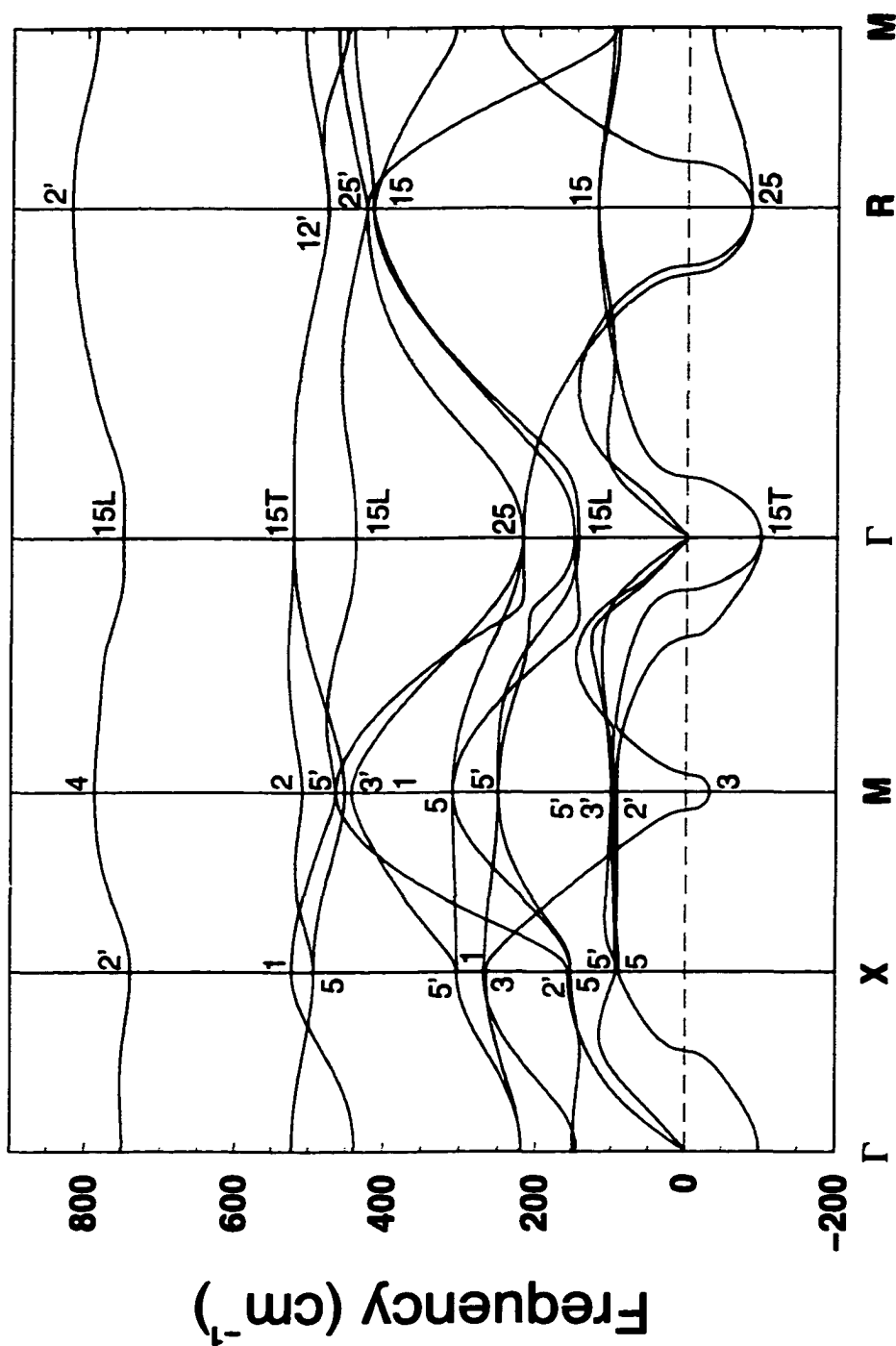[e]Reference [103], neutron scattering results.

Figure 5.5: Calculated phonon dispersion curves for cubic SrTiO₃ at the LDA lattice parameter a=7.412 Bohr. Imaginary frequencies lie below the dashed zero-frequency line.
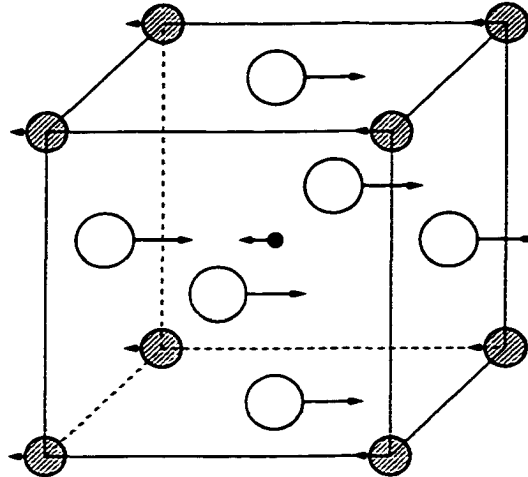
Figure 5.6: Motion of the atoms in the Γ_{15} TO mode.

the notation of Cowley.[28] Structural instabilities are indicated by portions of the dispersion curves lying below the $\omega=0$ dashed line in Fig. 5.5, indicating a large phase space for unstable modes. Instabilities appear at the R-point (R_{25}), zone-center (Γ_{15}), and M-point (M_3). The characters of these instabilities are of two types: FE (Γ_{15}), and AFD (R_{25} and M_3). The Γ_{15} mode is a TO mode where the Ti atoms and the oxygen octahedra roughly move antiparallel along the $\langle 100 \rangle$ direction. This is the displacive FE mode and is similar to that in $BaTiO_3$ and $KNbO_3$. The M_3 mode is nearly identical to the R_{25} mode, except that the octahedra are twisted in the same sense in adjacent cells along the c-axis.

The regions of instability in the BZ are better visualized in Fig. 5.8 as $\omega=0$ isosurfaces, for the lowest unstable phonon modes. The cubic BZ is also shown in outline, with the Γ point at the center, the R point at the cube corners, and the M point at the center of the cube edges. The inner isosurface, centered about the Γ-point, can be visualized as three interpenetrating disks, one perpendicular to
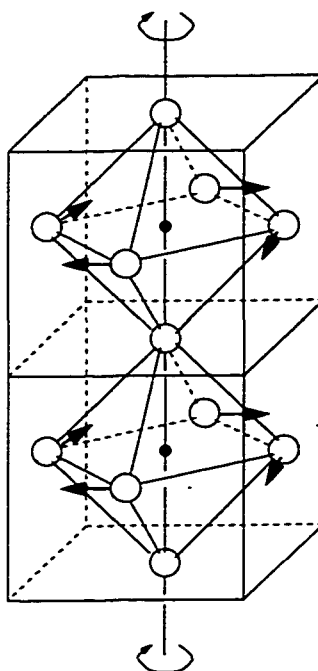
Figure 5.7: The rotations of the oxygen octahedra in the M_3 mode are in-phase along the rotation axis.

each Cartesian direction. The region of FE-type instability is interior to this first isosurface. Between the first and the second isosurface, which lies near the zone edges, all modes are stable. Unstable modes corresponding to the AFD instability are present in the regions adjacent to the R-M-R edge of the BZ. In a repeated-zone scheme, this region of instabilitiy would appear as cylindrical tubes centered on the R-M-R line. As discussed below, with decreasing volume the inner FE isosurface shrinks and the outer AFD surface grows, corresponding to the suppression of the FE and enhancement of the AFD instabilities. As pictured in Fig. 5.8, the phase space volume of the FE instability is greatly diminished compared to that in ferroelectric $KNbO_3$ [126] and $BaTiO_3$ [36] at their experimental volumes. In these materials, the region of FE instability consists of three interpenetrating planar slabs extending all the way out to the BZ boundary. Results discussed below demonstrate that the energy gain associated with the FE instability in $SrTiO_3$ is also much smaller than in these materials, showing that the instability is only marginal in $SrTiO_3$ compared to $BaTiO_3$ and $KNbO_3$.

As mentioned, the transformation to the tetragonal phase is due to the energetically favorable anti-phase rotations of the oxygen octahedra characteristic of the R-point instability. This is a second-order phase transition, and is not accompanied by a discrete volume change. Thus, both the equilibrium twist angle, ϕ, and the c/a ratio change smoothly from the transition point to experimentally observed maximum values of 2.1 degrees [104] and 1.0009 [48] at the lowest temperatures, respectively. The LDA equilibrium value of ϕ in the tetragonal phase (10 atom primitive cell) was determined by performing total energy calculations using the all-electron
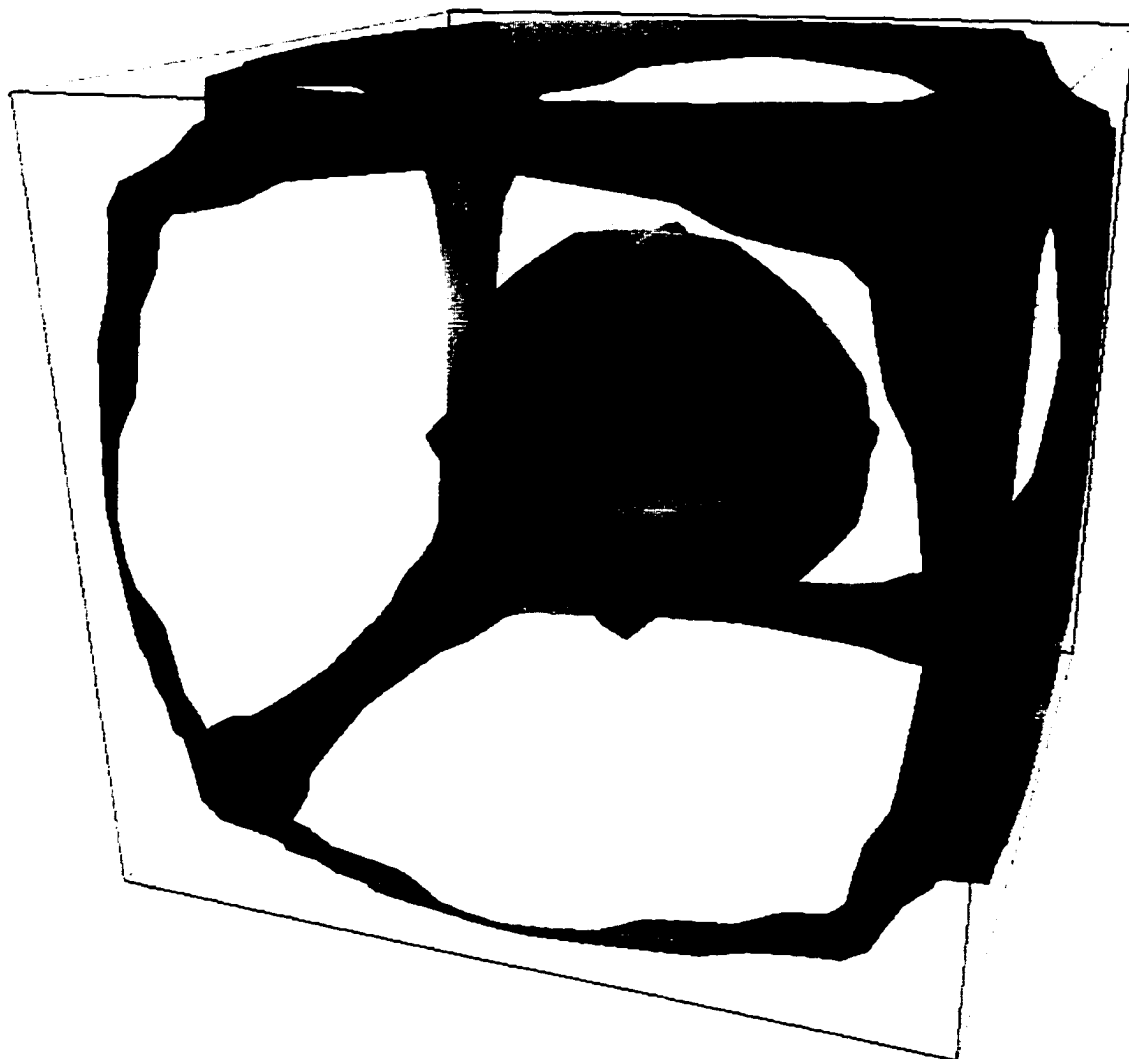
Figure 5.8: Zero-frequency isosurfaces of the lowest unstable phonon modes over the entire BZ. The Γ-point is located at the center of the cube. Unstable modes exist inside the central surface and along the full length of the zone-edges.
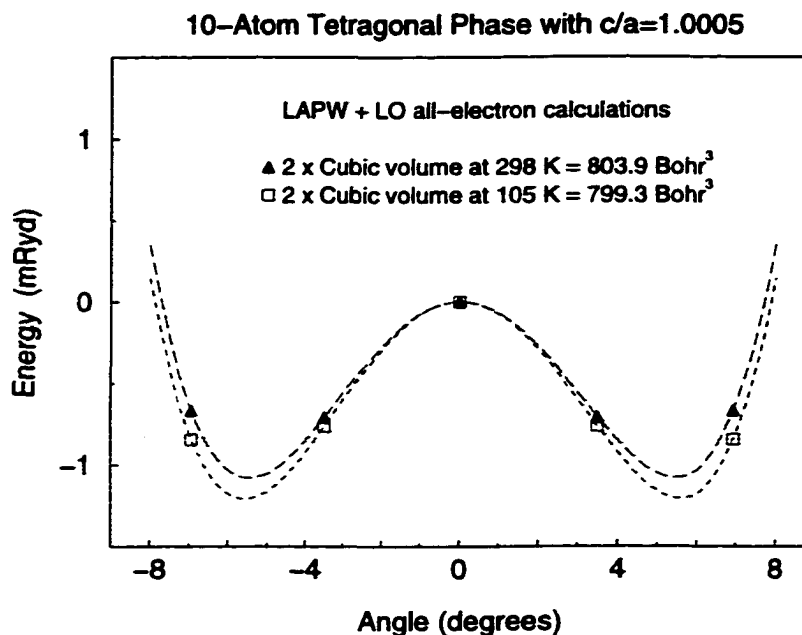
10–Atom Tetragonal Phase with c/a=1.0005

Figure 5.9: Total energy calculations for the AFD mode at two experimental volumes. A reduction in volume enhances the AFD instability.

LAPW+LO method,[96] where extra local orbitals are used in the basis. This was done using a fixed value of c/a=1.0005 at both the 298 K and 105 K experimental volumes, resulting in equilibrium angles of 5.4 and 5.6 degrees, respectively, as shown in Fig. 5.9. Work reported by Schwarz shows a similar overestimation of the equilibrium twist angle.[92] Calculations were also performed for M-point distortions and the well-depth was found to be much smaller than that for the antiphase rotations, in agreement with the smaller imaginary frequency of the cubic linear response results of Fig. 5.5. Schwarz, however, reports a deeper double well M-point distortion.

Using the curvature of the total energy double well minimum in Fig. 5.9, a value of 124 cm^{-1} was extracted for the A_{1g} AFD twist mode phonon frequency. This is somewhat larger than the measured value of only 48 cm^{-1}.[32] This overestimate is
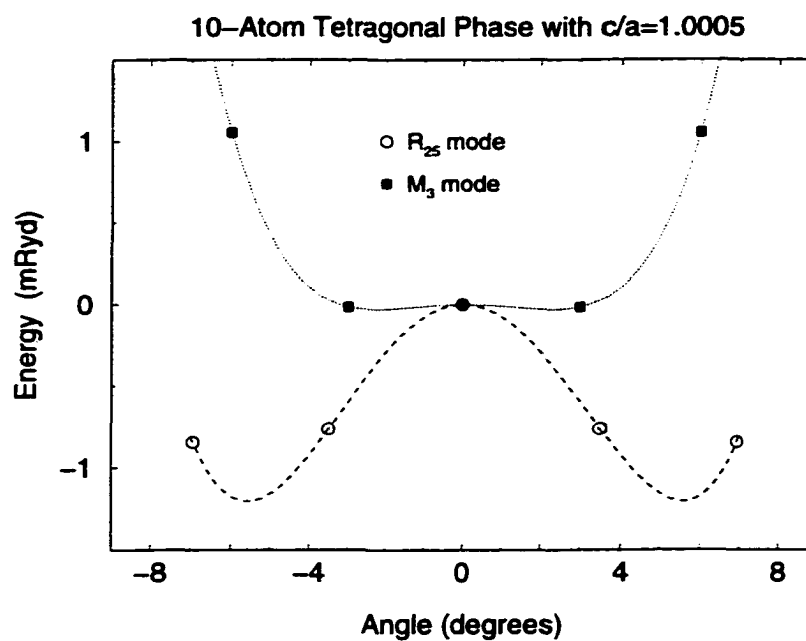
Figure 5.10: Total energy calculations for the M_3 mode reveal a very shallow well-depth compared to the R_{25} mode at the same 105 K volume.

probably related to the larger result for the LDA equilibrium twist angle. Reducing the volume in the calculations enhances the AFD instabilities at both the R and M points, as was also seen in the MC calculations of Zhong and Vanderbilt.[129] Changing the value of c/a did not significantly affect the results.

We now address the coupling of the ferroelectric and AFD instabilities. One important question is whether the FE instability, which is present in the cubic linear response calculations, is still present in the AFD tetragonal phase. In all of the remaining calculations discussed here, the all-electron LAPW+LO method was used. The first step in this analysis was to examine the effect of the smaller 105 K volume on the FE instability. Total energy calculations were performed in the 5-atom cubic phase to assess the effects of volume reduction. This was done by imposing the Γ_{15} eigenmode displacements in the cell for different mode amplitudes, and calculating the total energies. The results for three volumes are illustrated in Fig. 5.11, and show that a reduction in the volume from our LDA value to the 105 K value has the effect of eliminating the FE instability altogether. It should be noted, that compared to $BaTiO_3$ and $KNbO_3$, where the double-well depths are 1-2 mRy deep, the FE instability is very marginal in $SrTiO_3$, and its presence or absence is thus more sensitive to the choice of exchange-correlation potential. To test the effect of the exchange-correlation potential, the calculations were repeated using the Hedin-Lundqvist form. This yields a FE instability at the 298 K and 105 K volumes, but the energy scale is similar to that in Fig. 5.11 at the 7.41 a.u. lattice parameter. This indicates the marginal character of the FE instability in $SrTiO_3$.

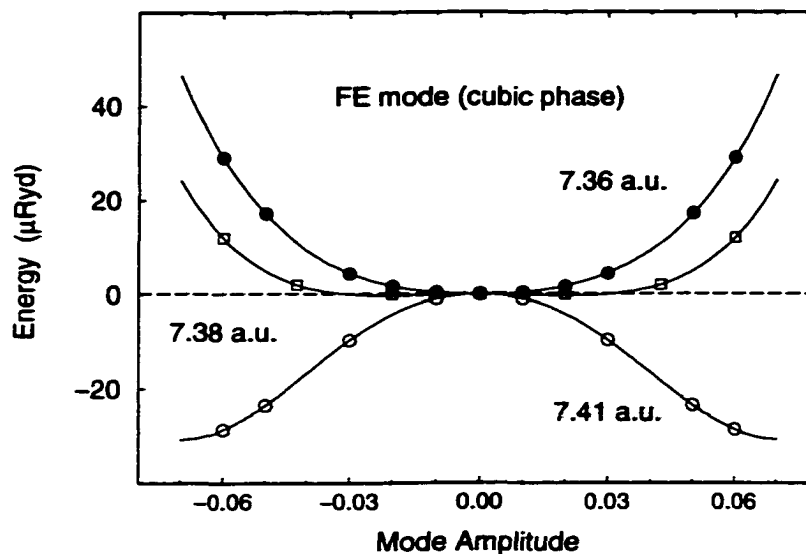To examine the direct coupling between the AFD and FE distortions, cubic

Figure 5.11: The TO soft phonon energy curves in the cubic phase at the LDA theoretical, 105 K experimental, and near 0 K lattice parameters.

eigenmode distortions were added to the slightly rotated octahedra. Although this is not the exact A_{2u} or E_u FE-type eigenmode in the tetragonal cell, the actual atomic distortions are very small (less than about 1% of the lattice parameter), so this should be a good approximation. This approximation also underlies the use of the effective Hamiltonian in Ref.[130] At the 105 K volume, fixing $c/a = 1.0005$, and setting $\phi=1.4$ degrees, total energy calculations as a function of FE distortion along the c- and a-axes were performed and are shown in Fig. 5.12. These curves are essentially similar, and the crystal is seen to be stable with respect to the FE distortion. 4th-order polynomial fits yielded phonon frequencies of 27 cm^{-1} and 30 cm^{-1} for the c- and a-axes respectively. Electric-field-induced Raman scattering results show the soft mode frequency to be about 11 cm^{-1} at 8 K for a 400 V/cm field.[123] At this field strength, splitting between soft modes parallel or perpendicular to the field was
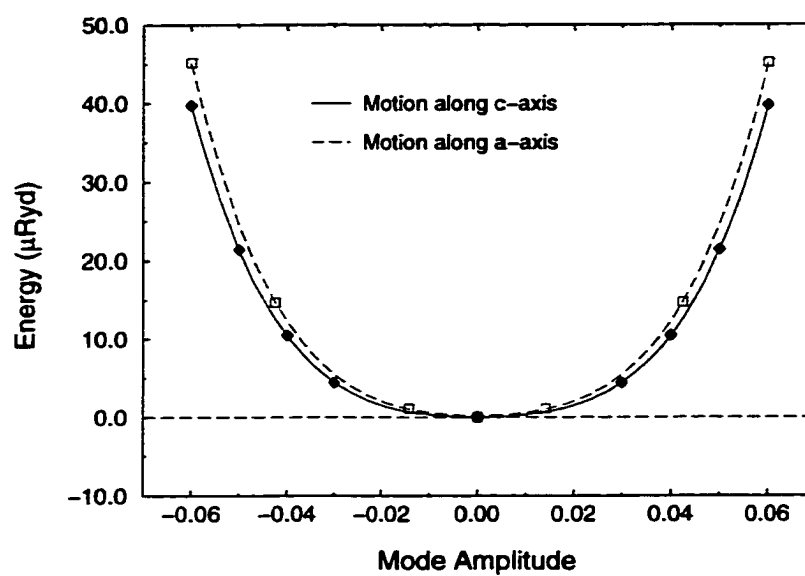
Figure 5.12: FE distortions along c-axis and a-axis in the tetragonal phase at the 105 K volume.

observed. The coexistent AFD and FE (a-axis) mode calculations were repeated at the 105 K volume for additional twist angles of 0, 2.1, and 5.6 degrees, spanning the range of experimental and LDA values. This resulted in a series of total energy curves which were virtually identical to the curves in Fig. 5.12, again showing essentially no interaction between the AFD and FE modes. These calculations were repeated at the 298 K volume, and again there was no appreciable interaction between the AFD and FE modes.

5.4 Conclusions

In summary, the direct coupling between the AFD and FE instabilities in $SrTiO_3$ is negligible. Their opposite volume dependence is the dominant factor determining the presence or absence of these instabilities, with volume reduction stabilizing the cell against a FE transition while enhancing the AFD transition. Moreover, the FE instability is much weaker compared to related materials like $BaTiO_3$ and $PbTiO_3$, making it difficult to determine whether the FE instability is even present as the volume is reduced upon cooling below room temperature. The absence of any detectible FE phase at low temperatures could thus be explained by the absence of the instability. On the other hand, if it were present, quantum fluctuations appear to be sufficient to suppress it. In either case, the shallow anharmonic character of the FE energetics may explain the Curie-Weiss behavior of the dielectric constant caused by the temperature-dependent anharmonic softening of the FE TO mode.

Bibliography

[1] S. L. Adler, Phys. Rev. **126**, 413 (1962).

[2] O. K. Andersen, Phys. Rev. B **12**, 3060 (1975).

[3] P. W. Anderson, in *Fizika dielectrikov*, ed. by G. I. Shansvi (Academy of Sciences of USSR, Moscow, 1960)

[4] G. B. Bachelet, D. R. Hamann, and M.Schlüter, Phys. Rev. B **26**, 4199 (1982).

[5] A. Baldereschi and R. Resta in *Ab-Initio Calculation of Phonon Spectra* ed. by J. T. Devreese, V. E. Van Doren and P. E. Van Camp (Plenum, New York, 1983), p.1.

[6] S. Baroni, P. Giannozzi, and A. Testa, Phys. Rev. Lett **58**, 1861 (1987).

[7] P. Bendt and A. Zunger, Phys. Rev. Lett. **50**, 1684 (1983).

[8] H. Bilz, G. Benedek, and A. Bussmann-Holder, Phys. Rev. B **35**, 4840 (1987).

[9] R. Blinc and B. Zeks, *Soft Modes in Ferroelectrics and Antiferroelectrics* (North Holland Pub. Co., Amsterdam, 1974), p.179.

[10] M. Born and K. Huang, *Dynamical Theory of Crystal Lattices*, (Oxford University Press, Oxford, England, 1954) p.166.

[11] C. G. Broyden, Math. Comp. **19**, 577 (1965).

[12] G. Busch and P. Scherrer, Naturwissenschaften **23**, 737 (1935).

[13] G. Busch, Helv. Phys. Acta **11**, 269 (1938).

[14] A. Bussmann-Holder and H. Büttner, Nature **360**, 541 (1992).

[15] J. Callaway, *Energy Band Theory*, (Academic Press, New York, 1964) p.99.

[16] R. Carr and M. Parrinello, Phys. Rev. Lett. **55**, 2471 (1985).

[17] D. M. Ceperley, Phys. Rev. B **18**, 3126 (1978).

[18] D. M. Ceperley and B. I. Alder, Phys. Rev. Lett. **45**, 566 (1980).

[19] D. J. Chadi and M. L. Cohen, Phys. Rev. B **8**, 5747 (1973).

[20] W. Cochran, Phys. Rev. Lett. **3**, 412 (1959).

[21] W. Cochran, Advan. Phys. **9**, 387 (1960).

[22] W. Cochran and R. A. Cowley, J. Phys. Chem. Solids **23**, 447 (1962).

[23] W. Cochran and A. Zia, Phys. Stat. Sol. **25**, 273 (1968).

[24] A. J. Coleman, Rev. Mod. Phys. **35**, 668 (1963).

[25] A. J. Coleman, in *The Force Concept in Chemistry*, edited by B. M. Deb (Van Nostrand Reinhold, New York, 1981) p.418.

[26] R. E. Cohen and H. Krakauer, Phys. Rev. B **42**, 6416 (1990).

[27] R. Comes, M. Lambert, and A. Guinier, Sol. St. Comm. **6**, 715 (1968).

[28] R. A. Cowley, Phys. Rev. **134**, A981 (1964).

[29] P. P. Ewald, Ann. d. Phys. **54**, 519 (1917).

[30] P. P. Ewald, Ann. d. Phys. **64**, 253 (1921).

[31] R. P. Feynman, Phys. Rev. **56**, 340 (1939).

[32] P. A. Fleury, J. F. Scott, and J. M. Worlock, Phys. Rev. Lett. **21**, 16 (1968).

[33] W. Frank, C. Elsasser, and M.Fahnle, Phys. Rev. Lett. **74**, 1791 (1995).

[34] A. Garcia, C. Elsasser, J. Zhu, S. G. Louie, and M. L. Cohen, Phys. Rev. B **46**, 9829 (1992).

[35] P. Ghosez, Doctoral Dissertation, Universite Catholique De Louvain, (1997)

[36] P. Ghosez, X. Gonze and J.-P. Michenaud, Ferroelectrics **206-207**, 205 (1998).

[37] P. Giannozzi, S. de Gironcoli, P. Pavone, and S. Baroni, Phys. Rev. B **43**, 7231 (1991).

[38] S. de Gironcoli, S. Baroni, and R. Resta, Phys. Rev. Lett. **62**, 2853 (1989).

[39] X. Gonze and J. -P. Vigneron, Phys. Rev. B **39**, 13120 (1989).

[40] X. Gonze, J.-C. Charlier, D. C. Allan and M. P. Teter, Phys. Rev. B **50**, 13035 (1994).

[41] X. Gonze, Phys. Rev. A **52**, 1096 (1995)

[42] O. Gunnarsson, M. Jonson, B. I. Lundquivst, Phys. Rev. B **20**, 3136 (1979).

[43] J. Haglung, Phys. Rev. B **47**, 566 (1993).

[44] D. R. Hamann, Phys. Rev. Lett. **42**, 662 (1979).

[45] D. R. Hamann, M. Schlüter, and C. Chiang, Phys. Rev. Lett. **43**, 1494 (1979).

[46] D. R. Hamann, L. F. Mattheiss, and H. S. Greenside, Phys. Rev. B **24**, 6151 (1981).

[47] L. Hedin and B. I. Lundqvist, J. Phys. C **4**, 2064 (1971).

[48] A. Heidemann and H. Wettengel, Z. Phys. **258**, 429 (1973).

[49] H. Hellmann, *Einführung in die Quantenchemie* (Deuticke & Co, Leipzig, 1937)

[50] J. Hemberger, P. Lunkenheimer, R. Viana, R. Böhmer, and A. Loidl, Phys. Rev. B **52**, 13159 (1995).

[51] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).

[52] S. Hoshino, K. Vedam, Y. Okaya, and R. Pepinsky, Phys. Rev. **112**, 405 (1958).

[53] A. C. Hurley, Proc. Roy. Soc. London, Ser. A **260**, 379 (1954).

[54] D. J. Jackson, *Classical Electrodynamics* (J. Wiley and Sons, New York, 1962), pp.152-4.

[55] D. J. Jackson, *op. cit.*, p.310.

[56] O. Jepsen, J. Madsen, and O. K. Andersen, Phys. Rev. B **18**, 605 (1978).

[57] R. O. Jones and O. Gunnarsson, Rev. Mod. Phys. **61**, 689 (1989).

[58] G. P. Kerker, J. Phys. C. **13**, L189 (1980).

[59] R. D. King-Smith and D. Vanderbilt, Phys. Rev. B **49**, 5828 (1994).

[60] D. D. Koelling and G. O. Arbman, J. Phys. F **5**, 2041 (1975).

[61] D. D. Koelling and B. N. Harmon, J. Phys. F **10**, 3107 (1977).

[62] M. Korling and J. Haglund, Phys. Rev. B **45**, 13293 (1992).

[63] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).

[64] H. Krakauer, M. Posternak, and A. J. Freeman, Phys. Rev. B **19**, 1706 (1979).

[65] K. Kunc and R. M. Martin, Phys. Rev. B **24**, 2311 (1981).

[66] L. D. Landau and E. M. Lifshits, *Statistical Physics*, (Pergamon Press, London, 1962).

[67] M. Levy, Proc, Natl. Acad. Sci. USA **76**, 6062 (1979).

[68] M. E. Lines and A. M. Glass, *Principles and Applications of Ferroelectrics and Related Materials* (Clarendon, Oxford, 1977).

[69] T. L. Loucks, *Augmented Plane Wave Method*, (Benjamin, New York, 1967).

[70] Z. W. Lu, Ph.D. thesis, The College of William and Mary, 1989.

[71] R. H. Lyddane, R. G. Sachs, and E. Teller, Phys. Rev. **59**, 673 (1941).

[72] F. W. Lytle, J. Appl. Phys. **35**, 2212 (1964).

[73] G. D. Mahan, Phys. Rev. A **22**, 1780 (1980).

[74] L. F. Mattheiss and D. R. Hamann, Phys. Rev. B **33**, 823 (1986).

[75] I. I. Mazin and D. J. Singh, *Fifth Williamsburg Workshop on First-Principles Calculations for Ferroelectics*, AIP Conference Proceedings **436**, 251 (1998).

[76] J. B. McKitterick, Phys. Rev. B **28**, 7384 (1983).

[77] W. J. Merz, Phys. Rev. **76**, 1221 (1949).

[78] R. Migoni, H. Bilz, and D. Bäuerle, Phys. Rev. Lett. **37**, 1155 (1976).

[79] T. Mitsui and W. B. Westphal, Phys. Rev. **124**, 354 (1961).

[80] T. Mitsui *et al.*, Landolt-Bornstein Series, Group III, vol. 3 (1969).

[81] H. J. Monkhorst and J. D. Pack, Phys. Rev. B **13**, 5188 (1976).

[82] K. A. Müller and H. Burkard, Phys. Rev. B **19**, 3593 (1979).

[83] K. A. Müller, W. Berlinger, and E. Tosatti, Z. Phys. B **84**, 277 (1991).

[84] F. D. Murnaghan, Proc. Natn. Acad. Sci. **30**, 244 (1944).

[85] J. P. Perdew, J. A. Chevary, S. H. Vosko, K. A. Jackson, M. R. Pederson, D. J. Singh, and C. Fiolhais, Phys. Rev. B **46**, 6671 (1992)

[86] R. M. Pick, M. H. Cohen, and R. M. Martin, Phys. Rev. B **1**, 910 (1970)

[87] P. Pulay, Molec. Phys. **17**, 197 (1969).

[88] R. Resta, M. Posternak, and A. Baldereschi, Phys. Rev. Lett. **70**, 1010 (1993).

[89] G. Rupprecht, R. O. Bell, Phys. Rev. **135**, A748 (1964).

[90] T. Sakudo and H. Unoki, Phys. Rev. Lett. **26**, 851 (1971).

[91] L. Salem and E. B. Wilson, J. Chem. Phys. **36**,3421 (1962).

[92] K. Schwarz, Phase Transitions, **52**, 109 (1994)

[93] L. J. Sham and W. Kohn, Phys. Rev. **145**, 561 (1966).

[94] G. Shirane and Y. Yamada, Phys. Rev. **177**, 858 (1969).

[95] D. J. Singh and H. Krakauer, Phys. Rev. B **43**, 1441 (1991).

[96] D. J. Singh, Phys. Rev. B **43**, 6388 (1991).

[97] D. J. Singh and J. Ashkenazi, Phys. Rev. B **46**, 11570 (1992).

[98] D. J. Singh, *Planewaves, Pseudopotentials, and the LAPW Method*, (Kluwer Academic Publishers, Massachusetts, 1994)

[99] J. C. Slater, Phys. Rev. **51**, 846 (1937).

[100] W. G. Spitzer, R. C. Miller, D. A. Kleinman, and L. E. Howarth, Phys. Rev. **126**, 1710 (1962).

[101] C. C. Stephenson, J. M. Corbella, and L. A. Russel, J. Chem. Phys. **21**, 1110 (1953).

[102] R. M. Sternheimer, Phys. Rev. **96**, 951 (1954); see also R. M. Sternheimer, *ibid.* **84**, 244 (1951); and R. M. Sternheimer and H. M. Foley, *ibid.* **92**, 1460 (1953).

[103] W. G. Stirling, J. Phys. C: Solid State Phys. **5**, 2711 (1972); W. G. Stirling, and R. Currat, J. Phys. C: Solid State Phys. **9**, L519 (1976).

[104] H. Unoki and T. Sakudo, J. Phys. Soc. Japan **23**, 546 (1967).

[105] H. Uwe and T. Sakudo, Phys. Rev. B **13**, 271 (1976).

[106] R. Viana, P. Lunkenheimer, J. Hemberger, R. Böhmer, and A. Loidl, Phys. Rev. B **50**, 601 (1994).

[107] J. Valasek, Phys. Rev. **15**, 537 (1920); J. Valasek, Phys. Rev. **17**, 475 (1921).

[108] U. von Barth and L. Hedin, J. Phys. C **5**, 1629 (1972).

[109] S. H. Vosko, L. Wilk, and M. Nusair, Can. J. Phys. **58**, 1200 (1980); S. H. Vosko and L. Wilk, Phys. Rev. B **22**, 3812 (1980).

[110] C. -Z. Wang, R. Yu, and H. Krakauer, Phys. Rev. Lett. **72**, 368 (1994).

[111] C. -Z. Wang, Ph. D. thesis, The College of William and Mary, 1995.

[112] C.-Z. Wang, R. Yu, and H. Krakauer, Phys. Rev. B **54**, 4467 (1996).

[113] C. -Z. Wang, R. Yu, and H. Krakauer, Phys. Rev. B **54**, 11161 (1996).

[114] H. E. Weaver, J. Phys. Chem. Solids **11**, 274 (1959).

[115] S.-H. Wei, Ph.D. thesis, The College of William and Mary, 1985.

[116] S. Wei and M. Y. Chou, Phys. Rev. Lett. **69**, 2799 (1992).

[117] S. Wei and M. Y. Chou, Phys. Rev. B **50**, 2221 (1994).

[118] M. Weinert, J. Math. Phys. **22**, 2433 (1981).

[119] M. Weinert, E. Wimmer, and A. J. Freeman, Phys. Rev. B **26**, 4571 (1982).

[120] H. Wendel and R. Martin, Phys. Rev. B **19**, 5251 (1979).

[121] E. Wigner, Phys. Rev. **46**, 1002 (1934).

[122] E. Wimmer, H. Krakauer, M. Weinert, and A. J. Freeman, Phys. Rev. B **24**, 864 (1981).

[123] J. M. Worlock and P. A. Fleury, Phys. Rev. Lett. **19**, 1176 (1967).

[124] M. T. Yin and M. L. Cohen, Phys. Rev. Lett **45**, 1004 (1980).

[125] R. Yu, D. Singh, and H. Krakauer, Phys. Rev. B **43**, 6411 (1991).

[126] R. Yu and H. Krakauer, Phys. Rev. B **49**, 4467 (1994).

[127] R. Yu and H. Krakauer, Phys. Rev. Lett. **74**, 4067 (1995).

[128] W. Zhong, R. D. King-Smith, and D. Vanderbilt, Phys. Rev. Lett. **72**, 3618 (1994).

[129] W. Zhong and D. Vanderbilt, Phys. Rev. Lett. **74**, 2587 (1995)

[130] W. Zhong and D. Vanderbilt, Phys. Rev. B **53**, 5047 (1996).

Vita

Christopher Andrew LaSota was born on March 2, 1970 in northern Pennsylvania to young parents dedicated to making their children's lives better than their own. He spent numerous years teaching himself about science and building rockets (from scratch), and eventually graduated high-school and enrolled in Kutztown University of Pennsylvania in the Fall of 1987. After successfuly completing his Bachelors Degree in physics in 1991, he was accepted to attend graduate school at the College of William and Mary in Virginia, where, by 1994, he earned a Master of Science degree in physics. Later, he passed his "professional hazing" exam, proving him worthy to perform doctoral research. Add on five more years of eating pasta, overdosing on carbonated beverages, and reading off-thesis-topic physics books instead of working, and he finally completes the requirements for a Ph.D in physics in August of 1999. Twelve years of higher education has taught him as much about life and people as it has about physics.