Ab initio lattice dynamics and structural phase transitions

Michele Carol Warren

A thesis submitted in fulfilment of the requirements
for the degree of Doctor of Philosophy
to the
University of Edinburgh
1997

Abstract

Prediction of the conditions required for the transformation of one phase of a mineral into another has long been a goal of condensed matter physics. This is especially desirable for phase transitions which are believed be involved in geological processes, but for which the conditions of temperature or pressure are hard to reproduce experimentally. This thesis examines a number of structural phase transitions including those of $MgSiO_3$ perovskite, which is thought to form the largest part of the Earth's mantle and of ZrO_2 which plays an important role in inhibiting crack formation in ceramics. These phase transitions, in which an alternative phase may be reached by continuous distortions of the structure on an atomic level, are examined primarily through the use of first principles electronic structure calculations. Existing first principles techniques were extended to facilitate determination of the equilibrium structure by relaxation of the unit cell and the calculation of the lattice dynamics of complex phases.

The distortion involved in most of the phase transitions studied is found to reflect the normal vibrational modes of one or both phases. The phase transitions of $MgSiO_3$ are found to be well described by only a few normal modes of the highest-symmetry cubic phase, dominated by two modes involving tilting of the SiO_6 octahedra. These modes resemble rigid unit modes, in which SiO_6 octahedra are assumed to remain perfectly rigid but may rotate with respect to other octahedra, whilst preserving linkages between them. The extent to which such simple modes are an accurate description of the dynamics of $MgSiO_3$, $BaZrO_3$ and SiO_2 is investigated by way of structural analysis and lattice dynamics of both stable and metastable phases. Both simple models

deduced from the lattice dynamical analysis and molecular dynamics using forces calculated from first principles are used to estimate transition temperatures for thermally induced phase transitions in $MgSiO_3$.

Declaration

This thesis has been composed by myself and it has not been submitted in any previous application for a degree. The work reported within was executed by me, unless otherwise stated.

April 1997

Acknowledgements

There are many people who have helped and encouraged me along the way, especially over the last few months.

Thanks must of course go to my supervisor, Graeme Ackland, for generally putting up with me; my second supervisor, Stuart Pawley, also had helpful things to say at crucial times, and kept us all smiling.

To Stewart Clark must go enormous thanks for not only endless technical assistance, advice and valuable opinion, but also for acting as a sounding board for all sorts of problems and ideas, supplying coffee and cake when necessary, and some heated arguments! Without his patient tolerance of many rants and paranoid ramblings this thesis might not have been written. My long-suffering colleagues and office-mates also endured an increasingly fractious fellow inmate: they know who they are and no doubt will still be receiving counselling. David Greig and David Allan amongst others were particularly encouraging in the final stages.

The many deeply kind and generous people at St Peter's Episcopal Church also listened wisely when I let off steam, and soon came to the conclusion that the best way to improve my state of wellbeing was to supply food: to all those that made me feel so welcome and valued go many thanks. My flatmates K and N also contributed to relaxation and reminded me that there is life beyond computational physics.

Throughout the entirety of my degree and PhD, my parents have been a constant source of encouragement, support and love, as well as more mundane forms of assistance. I could not have done any of it without them. My grandparents, too, have become good friends over that time, and have offered help and support in many different ways. My brother has been an amazing source of inspiration, achieving far more than anyone ever imagined — even if he does think I'm really rather strange.

Finally, to Martin Dove and Simon Redfern must go thanks for having faith in me, and offering me such an excellent incentive to finish this thesis and start on new projects and fresh horizons.

Contents

Chapter 1

Introduction

> *The human mind is seldom satisfied, and is certainly never exercising its highest function, when it is doing the work of a calculating machine. What the man of science, whether he is a mathematician or a physical enquirer, aims at is, to acquire and develop clear ideas of the things he deals with. For this purpose he is willing is enter on long calculations and be for a season a calculating machine, if he can only at last make his ideas clearer.*
>
> James Clerk Maxwell, *Scientific Papers* 1890/1891

A compound with a given chemical composition may exist in many different crystal structures, known as polymorphs, each stable over some range of temperature and pressure. Transitions between them can occur via a range of direct or indirect mechanisms. Prediction of the structure and properties of the most stable phase of a given compound under specified conditions has long been a goal of condensed matter theory. Furthermore, understanding the mechanisms by which one structure transforms into another can not only help explain why a transition occurs, but can also suggest ways of detecting the phase transition, and possible applications of the processes involved. The atomic rearrangements involved in transition mechanisms may not be directly observed experimentally, but instead can sometimes by deduced, for example by way of optical probes into vibrational properties.

In contrast to experimental techniques, computer simulations have the capacity for modelling the energetics and dynamics of individual atoms and thus allow investigation

of possible transition pathways. They also allow investigation of configurations which are unstable and are therefore never observed experimentally. However, there are a large number of computational techniques for modelling solid-state systems, which vary according to the computational effort required, the range of systems to which they are applicable and the reliability and accuracy of the results obtained. Much of the work presented in this thesis was performed using a first principles scheme to obtain the electronic charge density, from which all properties may be calculated. However, it is also intended to show how such calculations may be interfaced with simpler but less computationally demanding approaches.

1.1 Types of phase transition

There are many types of phase transition; only a subset will be considered in this thesis. *Displacive* phase transitions involve only small changes in the crystal structure, which do not require alterations to the bonding topology; in such transitions the symmetry of the low temperature phase is usually a subgroup of that at high temperature. In contrast, *reconstructive* phase transitions involve more drastic changes to the structure, necessitating the breaking and re-forming of bonds and possibly local migration of atoms; the space groups of the two phases may be entirely unrelated. A third type of transition occurs where equivalent sites in the high-temperature phase may be occupied by one of two (or more) different species, but have a spatially averaged occupancy of 50% of each species. On lowering the temperature the different species form an ordered distribution over these sites, which is likely to reduce the symmetry. Such *order-disorder* transitions also involve bond-breaking and diffusion.

Only the displacive class of transitions are discussed in detail in this thesis. Such transitions, involving only small changes in the structure, are often related to the vibrational properties of the phases. It is hoped to investigate in this thesis this relation in a number of crystal structures.

It should be noted that the terms 'displacive' and 'order-disorder' are also used to describe extremes of behaviour in a simple model which may represent both these types of transition [1]. The intermediate stages of behaviour have been found to describe a great range of phase transitions, and the applicability and parametrisation of this model are discussed in this thesis.

1.2 *Ab initio* techniques

Traditionally, atomistic calculations have used analytic interatomic potentials, often parametrised from experimental data, to model the interactions between atoms in a crystal. Usually interactions between only pairs or triplets of atoms are considered, thereby including only bond-stretching and bond-bending terms. Extremely simple approaches, like the Lennard-Jones potential for noble gas elements, or the Madelung energies of Coulomb interactions in ionic crystals (which although simple in concept still require careful evaluation) have produced invaluable insights into the behaviour of fundamental crystal structures. Increasingly sophisticated potentials have explained many features of more complicated systems [2]. Such potentials have the great advantages of being analytically tractable and usually quick to evaluate computationally. Simulations of millions of atoms are therefore perfectly possible, allowing investigation of features such as grain and crack formation, dislocations and surfaces. However, they are limited to the extent to which they can describe complex processes, often because the parameters which would be needed to describe the necessary potentials may not be obtained experimentally. In many cases it is impossible to predict in advance which are the salient features of a process, which would be needed to determine the principal variables in an appropriate potential. In other cases, the bond-breaking and re-forming involved in a process may simply not be described with empirical potentials.

In contrast, *ab initio* calculations make no *a priori* assumptions about a system, such as the form of the solution or which details, if any, may be neglected. In principle, such

an approach requires the solution of the full Schrödinger equation. This should give completely accurate results, but for a macroscopic system is too complex to be solved. The advent of density functional theory [3, 4] presented an alternative framework for finding the total energy, and by making only a very few approximations, a feasible scheme for solving the Schrödinger equation for the electrons was determined (described in Chapter 2). Over the last twenty years, such schemes have been used with great success to calculate various static and dynamic crystal properties with high accuracy [5].

However, density functional theory calculations are extremely computationally intensive, and thorough studies are possible only for systems with modest numbers of atoms in the simulation cell: typically, cells contain tens of atoms, although a few hundred atoms have been treated. With current levels of computational power, molecular dynamics simulations of tens of thousands of atoms are still infeasible, and may only be performed with empirical potentials. It would therefore be advantageous to develop intermediate approaches whereby the results of first principles calculations are used to develop reliable simple potentials, which may then be used to perform much larger simulations. *Ab initio* studies may also be used to identify the dominant features of complex processes in solids, and hence determine the variables in terms of which such potentials should be parametrised.

In order to compare related phases, and to calculate the properties of each, the equilibrium structure of each phase must be found at the required conditions of temperature and pressure. Traditionally this has required evaluation of the total energy over a large range of structural parameters, followed by minimisation via a fitted form for the energy as a function of these parameters. However, if the forces on atoms and the internal stresses on the unit cell are also calculated from first principles, relaxation of the structure towards the minimum of energy may be possible. We describe in this thesis developments to this technique, and the sources of inaccuracies in practical schemes.

1.3 Soft and unstable phonon modes

The relation of some low-frequency vibrational modes to the onset of structural phase transitions has been long known [6, 7] and used as an experimental handle to detect and follow such transitions [8]. A low-frequency phonon indicates a distortion of the structure which requires only low energy. If the mode is coupled to the strain, a change in lattice vectors may be sufficient to render the mode unstable. The slightest thermal activity, as must always be present, is then sufficient to introduce a non-zero amplitude of the phonon eigenvector. This will reduce the total energy of the system, and will be manifested as a phase transition to a new structure, which usually has a space group which is a subgroup of that of the initial structure.

The original, higher-symmetry phase may be recovered in two ways. Firstly, at a sufficiently high temperature, the thermal energy may overcome the energy differences between the phases. Secondly, a change in pressure may favour the higher-symmetry phase, in which the phonon must then be stable. Such pressure-induced phase transitions may be detected in a straightforward way from first principles simulations, by finding the phonon frequency at a range of volumes. However, *ab initio* simulations at finite temperature, which ideally should consider long-range disorder, require such large amounts of computer time that proper simulations of phase transitions, although possible [9], are not yet routine. Other methods of estimating transition temperatures must thus be considered.

The amplitude of any soft mode involved in such a transition is clearly strongly related to the progress of the transition, and may be treated as an order parameter. This simplifies the description of the atomic degrees of freedom, by concentrating only on those which play a critical role in the transition.

Not all phase transitions occur by means of an unstable vibrational mode. Displacive transformations, resulting in only small changes in structure, are the most commonly described in this way; reconstructive transformations typically involve such drastic rearrangement of atoms that a phonon description would be stretched beyond

the limit of applicability. Any transition requiring extensive diffusion could not be described in this way, since under no conditions would regular oscillatory motion be observed.

The relative displacements of individual atoms in a phonon may not be deduced experimentally, although the symmetry and and polarisation may be found. However, they may be found via theoretical determinations of the vibrational modes, so computational studies may complement and extend any results available from experiment. A method of calculating phonon frequencies and eigenvectors has been developed as part of the work described here, using the first-principles forces available from an established density-functional scheme.

Phonon eigenvectors may yield useful information on the nature of the normal modes, and those of low frequency or unstable modes show the initial distortion involved in any corresponding phase transition, so that lower energy structures may thus be proposed. In some cases, the eigenvectors may be interpreted in terms of groups of atoms within the structure, and the origin of any strain coupling may be deduced. Such interpretations suggest ways of defining the extent of coupling between the strain and a vibrational mode, which will be developed in this thesis.

1.4 Work submitted for publication

Much of the work described in this thesis has been submitted for publication; references are given here and for completeness also in the Bibliography.

[10] "Ab initio studies of structural instabilities in magnesium silicate perovskite", M.C. Warren and G.J. Ackland *Phys. Chem. Minerals*, 23:107–118, 1996.

[11] "Vibrational properties of the layered semiconductor germanium sulfide under hydrostatic pressure: theory and experiment", H.C. Hsueh, M.C. Warren, H. Vass, G.J. Ackland, S.J. Clark and J. Crain. *Phys. Rev. B*, 53:14806–14817, 1996.

[12] "Structure and elasticity of MgO at high pressure", B.B. Karki, L. Stixrude, S.J. Clark, M.C. Warren, G.J. Ackland and J. Crain. *Am. Mineral.*, 82:51–60, 1997.

[13] "Ab initio elasticity and lattice dynamics of $AgGaSe_2$", B.B. Karki, S.J. Clark, M.C. Warren, H.C. Hsueh, G.J. Ackland and J. Crain. *J. Phys.: Cond. Mat.*, 9:375–380, 1997.

[14] "Origin of the negative thermal expansion in ZrW_2O_8 and ZrV_2O_7", A.K.A. Pryde, K.D. Hammonds, M.T. Dove, V. Heine, J.D. Gale and M.C. Warren. *J. Phys.: Cond. Mat.*, 8:10973–10982, 1996.

[15] "Ab initio studies of high-pressure structural transformations in silica", B.B.Karki, L. Stixrude, M.C. Warren, G.J. Ackland and J. Crain. *Phys. Rev. B*, 55:3465–3471, 1997.

[16] "Rigid unit modes and the negative thermal expansion in ZrW_2O_8", A.K.A. Pryde, K.D. Hammonds, M.T. Dove, V. Heine, J.D. Gale and M.C. Warren. *Phase Transitions.* In press.

[17] "Soft-mode phase transitions from first principles". G.J. Ackland and M.C. Warren. *Phase Transitions.* In press.

[18] "Elastic properties of orthorhombic $MgSiO_3$ perovskite at lower mantle pressures", B.B. Karki, L. Stixrude, S.J. Clark, M.C. Warren, G.J. Ackland and J. Crain. *Am. Mineral.* In press.

1.5 Overview of thesis

In Chapter 2 the main features of density functional theory are presented, including a comparison with earlier attempts to solve the Schrödinger equation in a many-electron system. The various approximations and numerical techniques which are needed in a practical scheme are also described.

Chapter 3 explains how a total-energy scheme, and specifically the pseudopotential plane-wave codes used in this work, may be used to find the equilibrium structure of a crystal; Chapter 4 describes how the vibrational modes may then be determined. Both these chapters include developments to existing code which were made by the author, and which are used in subsequent chapters.

Chapter 5 presents the results of calculations on three different phases of $MgSiO_3$ perovskite, and implications for the mechanisms of transition between them. First-principles molecular dynamics techniques are used to further investigate one of the transitions. The particular vibrational modes of $MgSiO_3$ which dominate any phase transitions have been described in terms of 'rigid unit modes', in which the SiO_6 octahedra behave as rigid entities. The extent to which these simple models are an accurate description of this and other systems is investigated in Chapter 6. A simple model for these transitions is deduced from the vibrational information, and the transition temperature estimated, in Chapter 7.

The relation between an unstable vibrational mode and a structural phase transition is also evident in ZrO_2, *ab initio* studies of which are presented in Chapter 8. A phase transition from a cubic to a tetragonal phase is explained by an unstable phonon, but a further transition to a monoclinic phase is not predicted, adding to previous speculation that the phase present in experiments must have a more complicated structure than assumed.

Chapter 2

Total energy calculations

> *"Do you see, Piglet? Look at their tracks! Three, as it were, Woozles, and one, as it was, Wizzle.* Another Woozle has joined them!*"*
>
> *And so it seemed to be. There were the tracks; crossing over each other here, getting muddled up with each other there; but, quite plainly every now and then, the tracks of four sets of paws.*
>
> A.A. Milne, *Winnie the Pooh*

The equilibrium properties of a system, such as cohesive energy, structure or thermal vibration, are determined by the interactions between the nuclei and electrons. At such small scales, quantum mechanics must be used to calculate these properties. In principle, all properties can be obtained from the many-body wavefunction, which is the solution of the many-body Schrödinger equation. However, this cannot be solved for the enormous number of electrons in a solid which interact very strongly with each other and with the nuclei. The various approximations which have been applied to try to solve this problem are presented in this chapter.

This is an overview of established theory, of which there are thorough presentations elsewhere [5, 19, 20]. This chapter aims to establish the principles on which the rest of the thesis is founded, and to form a basis for later discussion.

The state of a system includes both the nuclei and the electrons, but since the electrons are much lighter than the nuclei, they respond much faster to any change in

the system. The Born-Oppenheimer approximation assumes that they move instantaneously when the nuclear configuration is changed, so that the many-body problem is reduced to that of finding the electronic configuration of a fixed arrangement of atomic nuclei. This approximation is made throughout the work described here.

2.1 Many-body problems in quantum-mechanics

The electrons in a solid interact strongly with each other via the Coulomb interaction, so the state of the system must depend on all the electrons. The Schrödinger equation is therefore applied to a many-body wavefunction $\Psi(\mathbf{r}_1, \mathbf{r}_2, ...\mathbf{r}_N)$ describing all the N electrons, rather than to each one separately. For a static external potential V_{ext} (e.g. the Coulomb interaction with the nuclei in the Born-Oppenheimer approximation) the Hamiltonian takes the form

$$\mathcal{H} = -\frac{\hbar^2}{2m}\nabla^2 + V_{\text{ext}} + \frac{1}{2}\sum_i \sum_{j \neq i} \frac{e^2}{4\pi\epsilon_0 \mid \mathbf{r}_i - \mathbf{r}_j \mid} \tag{2.1}$$

The third term in the Hamiltonian is the Coulomb interaction between pairs of electrons. However, it is not known in what functional form Ψ should be expressed, and thus this equation cannot be solved. We therefore turn to various approximations that have been made, which each make assumptions for the form used to write Ψ.

2.1.1 Electron interaction

The Coulomb repulsion between electrons favours a state in which they are as far apart as possible. However, for the electrons to be each highly localised and well separated from one another, the many-electron wavefunction would require high kinetic energy components (via $\nabla^2\Psi$), so a compromise must be reached. In addition, electrons are fermions and must be antisymmetric under exchange. This helps to keep electrons apart, via the Pauli exclusion principle, since no two electrons of the same spin may be localised in the same place. The overall problem is not a simple one: in a system

with more than two electrons, many-body effects contribute to the full electron correlation. For example, pairs of electrons with opposite spins are indirectly affected by exchange via their individual interactions with other electrons, some of which must include exchange effects.

2.1.2 The Hartree approximation

The Hartree approximation treats electrons as independent non-interacting particles, with wavefunctions which are solutions of a one-electron Schrödinger equation. The interaction with the other electrons is taken into account only by calculating each electron's Coulomb interaction with the overall electronic charge density. The Schrödinger equation for the ith electron, represented by $\psi_i(\mathbf{r})$, is then:

$$-\frac{\hbar^2}{2m}\nabla^2\psi_i(\mathbf{r}) + (V_N + V_H)\psi_i(\mathbf{r}) = \varepsilon_i\psi_i(\mathbf{r}) \tag{2.2}$$

where the ion-electron interaction forms the nuclear potential, which is the 'external' potential in (2.1):

$$V_N(\mathbf{r}) = \sum_I \frac{Z_I e^2}{4\pi\epsilon_o \mid \mathbf{r} - \mathbf{R}_I \mid} \tag{2.3}$$

The interaction with the overall charge density $\rho(\mathbf{r})$ is given by the Hartree potential:

$$V_H(\mathbf{r}) = e\int \frac{\rho(\mathbf{r}')}{4\pi\epsilon_0 \mid \mathbf{r} - \mathbf{r}' \mid} d\mathbf{r}' \tag{2.4}$$

If the charge density is known, so that (2.2) can be solved, the eigensolutions may be populated by electrons, enforcing orthogonality to ensure Pauli exclusion. The overall charge density could then be obtained by summing the contributions from all the wavefunctions:

$$\rho(\mathbf{r}) = e\sum_i \psi_i^*(\mathbf{r})\psi_i(\mathbf{r}) \tag{2.5}$$

Equations (2.2–2.5) can thus be solved iteratively. However, not only does (2.4) include each electron's interaction with itself, which must be incorrect, it also takes no account

of individual interactions. The Hartree approximation that electrons are independent, apart from interaction via the average charge, is equivalent to representing the overall wavefunction Ψ as a product of individual one-electron wavefunctions:

$$\Psi_H(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N) = \psi_1(\mathbf{r}_1)\psi_2(\mathbf{r}_2)\ldots\psi_N(\mathbf{r}_N) \tag{2.6}$$

so that the Schrödinger equation can be solved for each one. However, Ψ_H is symmetric under the exchange of any two electrons, violating the Fermi spin statistics. The Hartree approximation thus takes no account of the exchange energy of the electrons.

The Hartree-Fock approximation to the many-body wavefunction takes the form of a Slater determinant, which contains all possible products of the individual wavefunctions with appropriate signs, so as to be antisymmetric under electron exchange. This includes the exchange interaction between electrons with the same spin but still neglects the interactions between electrons with opposite spins; correlation is therefore ignored. Even so, the Hartree-Fock approximation is much more complicated to solve than the Hartree approximation, and produces some unfortunate results such as a zero density of states at the Fermi level [21].

Density-functional theory provides a more general framework for taking account of exchange and correlation between electrons, without increasing the complexity of the problem beyond that of the Hartree approximation.

2.1.3 Density functional theory

Density functional theory (DFT) was first proposed by Hohenberg and Kohn [3] in 1964, and states that all the properties of the system, including the total energy, are unique functionals of only the electron density $n(\mathbf{r})$. Furthermore, that electron density which minimises the total energy is the ground state density and thus gives the correct total energy. The electron density thus acquires a central variational role in the problem.

This is in contrast to the usual status of the charge density[1] as a quantity found after solving the Schrödinger equation with a given external potential.

An exact expression for the energy within the formalism of density-functional theory was given by Hohenberg and Kohn for an external potential V_{ext} as a functional of the electron density:

$$E[n(\mathbf{r})] = \int V_{\text{ext}}(\mathbf{r})n(\mathbf{r})d\mathbf{r} + \frac{e^2}{2}\iint \frac{n(\mathbf{r})n(\mathbf{r}')}{4\pi\epsilon_0 \mid \mathbf{r}-\mathbf{r}' \mid} d\mathbf{r}'\, d\mathbf{r} + G[n(\mathbf{r})] \tag{2.7}$$

where $G[n(\mathbf{r})]$ is a universal functional of the electron density $n(\mathbf{r})$, acting in addition to the Hartree interaction energy. This functional contains all the exchange and correlation due to electron interaction, and the electronic kinetic energy, but is not known analytically.

Kohn and Sham [4] broke up the $G[n(\mathbf{r})]$ functional by identifying the kinetic energy of an equal number of *non-interacting* 'pseudo-electrons'. This mapped the many-body problem of minimising the total energy given in (2.7) with respect to $n(\mathbf{r})$ to a number of one-particle equations. These 'Kohn-Sham equations' represent pseudo-electrons moving in an effective potential V_{eff}:

$$-\frac{\hbar^2}{2m}\nabla^2\psi_i(\mathbf{r}) + V_{\text{eff}}\psi_i(\mathbf{r}) = \varepsilon_i\psi_i(\mathbf{r}) \tag{2.8}$$

with

$$V_{\text{eff}} = V_{\text{ext}}(\mathbf{r}) + \int \frac{n(\mathbf{r}')e^2}{4\pi\epsilon_0 \mid \mathbf{r}-\mathbf{r}' \mid} d\mathbf{r}' + \mu_{XC}(n(\mathbf{r})). \tag{2.9}$$

The potential $\mu_{XC}(n(\mathbf{r}))$ contains the exchange and correlation term, and is the functional derivative of the exchange and correlation energy with respect to electron density

[1] The charge density $\rho(\mathbf{r})$ is taken in this thesis to include a factor of e, so has units C/m^3, but the electron density $n(\mathbf{r})$ is dimensionless.

(see Section 2.1.4). The electron density may now be found from the individual pseudo-wavefunctions:

$$n(\mathbf{r}) = \sum_i \psi_i^*(\mathbf{r})\psi_i(\mathbf{r}) \tag{2.10}$$

The Kohn-Sham system of equations (2.8–2.10) can be solved iteratively, from an initial guess for $\{\psi_i\}$, to find the single-electron pseudo-wavefunctions. These represent pseudo-electrons moving in an effective potential, with energy eigenvalues ε_i which are strictly no longer the true electron band energies, but are known as Kohn-Sham eigenvalues. The electron density given can then be used to find the correct total energy using (2.7).

Density functional theory would in principle be exact, if a correct form for the exchange and correlation potential were available; if this were the case, equations (2.8–2.10) could be solved iteratively. All the exchange and correlation physics is contained in μ_{XC} and the orthogonality between wavefunctions, but an exact form for μ_{XC} is still unknown. Approximation to this are thus still required, and the simplest of these is the *local density approximation* [4].

2.1.4 Local density approximation

The local density approximation takes the exchange-correlation energy per electron at a point $\mathbf{r}$ to be that of a homogeneous electron gas (jellium) with the same electron density $n(\mathbf{r})$, so that

$$E_{XC}[n(\mathbf{r})] = \int \epsilon_{XC}^{\text{hom}}(n)\, n(\mathbf{r}) d\mathbf{r} \tag{2.11}$$

The exchange-correlation potential is then given by [4]

$$\mu_{XC}(n(\mathbf{r})) = \frac{\delta E_{XC}[n(\mathbf{r})]}{\delta n(\mathbf{r})} = \frac{\partial[n\,\epsilon_{XC}^{\text{hom}}(n)]}{\partial n} \tag{2.12}$$

The Pauli exclusion principle, which underlies the many-body effects of exchange and correlation, states that two electrons with like spin may not have the same state.

When full exchange and correlation are considered, exactly one electron is excluded from the immediate vicinity of another. This should be reproduced by the LDA or other approximations if they are to give satisfactory results.

Figure 2.1 illustrates the effects of different exchange correlation schemes on the electron probability at a distance r from an electron of some spin. In the Hartree approximation, the electrons affect each other only via the Coulomb interaction with the total charge density, effectively moving independently of one another. The Hartree-Fock approximation includes the exchange hole, satisfying the Pauli exclusion principle, but has no effect on the electron of opposite spin. The local density approximation includes both exchange and correlation, and acts on electrons of both spins; it may be thought of as the spin average of the local spin density approximation (LSDA) which is also shown.

Although the LDA is in principle very simple, it was found to work surprisingly well. This is because it results in each electron excluding the right amount of charge around it, even if the shape of the exchange hole (or depletion hole) is not exactly right [22]. The depletion hole is also identifiable in Figure 2.1 for all but the Hartree approximation, where is it neglected. Early attempts to produce a more sophisticated approximation failed, because they did not obey this one-electron hole rule. However, more sophisticated schemes are now possible, including the generalised gradient correction (GGC).

The LDA is implemented by means of a parametrisation for $\mu_{XC}(n)$ [22], based on exact quantum Monte-Carlo results, such as those of Ceperley and Alder [23]. In these calculations, the positions of a system of N electrons are represented by a point in $3N$-dimensional space; because the positions of the electrons in each system are explicitly known, the many-body Hamiltonian can be exactly evaluated for each system. Many such systems (about 100 in [23]) are selected such that the probability of finding a point in $d\mathbf{R}$ is initially proportional to $|\Psi_T(\mathbf{R})|^2$ for some trial wavefunction Ψ. The initial distribution of points is then allowed to propagate via a diffusion equation

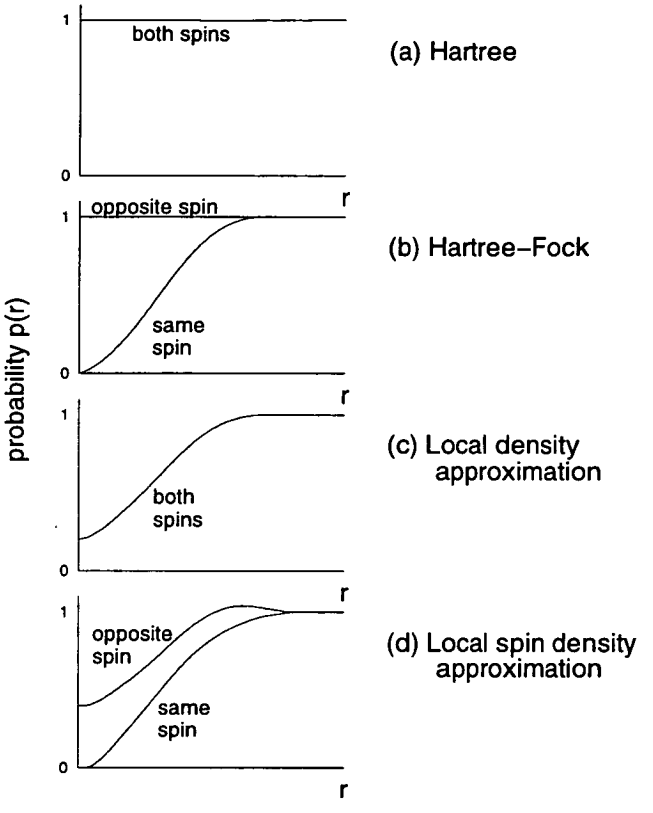


Figure 2.1. The treatment of exchange and correlation in different levels of approximation in a N-electron system, represented by the probability distribution of $N-1$ electrons at a distance r from an electron of fixed spin. Probabilities are normalised such that $p(r) \to 1$ as $r \to \infty$ The generalised gradient approximation is not illustrated but would not differ qualitatively from the local density approximation. (After Schlüter and Sham [21]).

which includes the effects of random diffusion, a quantum force $\nabla \ln |\Psi_T(\mathbf{R})|^2$, and a 'branching' term which eliminates systems with high energies but replicates systems with low energies. The steady-state final population can be used to give the correct eigenvalue, from which the energy $\epsilon_{XC}^{\text{hom}}(n)$ may be calculated. The function $\mu_{XC}(n)$, which is parametrised with only five variables, can then be stored very easily and is used in almost all the simulations described in this work.

2.2 Properties of periodic systems

Although density-functional theory enables exchange and correlation to be included, the Kohn-Sham equations (2.8–2.10) still have to be solved for each electron in a solid. Since the electron density in a typical solid is of the order of $10^{27}\mathrm{m}^{-3}$, this would be a somewhat impractical task. However, the translational symmetry and other properties of perfect crystals can be used in an extremely powerful way to reduce the number of electrons which need to be considered.

The use of *supercells* with Born-von Karman periodic boundary conditions [24] allows simulation of an infinite periodic crystal by consideration of only one unit cell. The 'supercell' used for simulation may contain multiples of the crystallographic primitive unit cell, in order to study properties other than the total energy, such as phonon modes with wavelengths longer than the unit cell dimension. The formalism is the same, however, and allows the application of Bloch's theorem, plane wave basis sets and crystal symmetry to the problem of finding the electron distribution. These techniques are described below.

2.2.1 Bloch's theorem

In an infinite perfect crystal, the external potential V_{ext} of (2.1) and (2.7) which acts on the electrons is due to a periodic array of nuclei, so the resulting charge density $\rho(\mathbf{r})$ must also have the periodicity of the lattice, that is:

$$\rho(\mathbf{r}+\mathbf{R}_L) = \rho(\mathbf{r}) \tag{2.13}$$

where $\mathbf{R}_L$ is a lattice vector. However, since $\rho(\mathbf{r}) = e\sum_i \psi_i^*(\mathbf{r})\psi_i(\mathbf{r})$, this can be satisfied even with a phase difference $e^{i\theta}$ between $\psi(\mathbf{r})$ and $\psi(\mathbf{r}+\mathbf{R}_L)$:

$$\psi(\mathbf{r}+\mathbf{R}_L) = e^{i\theta}\psi(\mathbf{r}) \tag{2.14}$$

Bloch's theorem states that the phase difference θ can be written as $\theta = \mathbf{k}\cdot\mathbf{R}_L$, introducing a wavevector $\mathbf{k}$. By writing each wavefunction $\psi(\mathbf{r})$ as the product of a wavelike part $e^{i\mathbf{k}\cdot\mathbf{r}}$ and a function $f(\mathbf{r})$,

$$\psi(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}} f(\mathbf{r}), \tag{2.15}$$

we see that it will then satisfy (2.13) only if $f(\mathbf{r}+\mathbf{R}) = f(\mathbf{r})$, i.e. if $f(\mathbf{r})$ has the periodicity of the lattice. This function need however only be defined over one unit cell.

The wavevectors $\mathbf{k}$ now label various sets of valid wavefunctions in a solid: in an infinite crystal they may take any value, but are restricted by convention to the first Brillouin Zone (i.e. the Wigner-Seitz cell of reciprocal space) without loss of generality. If each unit cell contains $2m$ electrons, there are m doubly occupied bands having different $f(\mathbf{r})$ (labelled by j) at each wavevector; $\mathbf{k}$ and j now label the electron states. Bloch's theorem thus provides the basic formalism of bandstructure which underlies much of condensed matter physics.

2.2.2 Sampling of k-space

The problem of a potentially infinite number of electrons has now been mapped to that of a finite number of electron wavefunctions at an infinite number of wavevectors $\mathbf{k}$. The electron density should be found by summing contributions from all $\mathbf{k}$ and j, and the total energy found from this density. If wavefunctions for doubly occupied orbitals are normalised such that

$$\int \psi^*_{j,\mathbf{k}}(\mathbf{r})\psi_{j,\mathbf{k}}(\mathbf{r})d\mathbf{r} = 2 \tag{2.16}$$

then the total electron density is given by

$$n(\mathbf{r}) = \frac{\nu}{(2\pi)^3}\int_{BZ}\sum_j \psi^*_{j,\mathbf{k}}(\mathbf{r})\psi_{j,\mathbf{k}}(\mathbf{r})d\mathbf{k} \tag{2.17}$$

where ν is the unit cell volume, and BZ denotes integration over all $\mathbf{k}$ in the Brillouin Zone, thus maintaining

$$\int n(\mathbf{r})d\mathbf{r} = 2m. \tag{2.18}$$

However, since the wavefunctions in semiconductors usually only change smoothly and slowly with $\mathbf{k}$, integration over the whole Brillouin Zone may be approximated by sampling the bandstructure at only a finite number of $\mathbf{k}$-points. Each $\mathbf{k}$-point is assigned a weight α_i such that $\sum_i \alpha_i = 1$. The electron density averaged over these is then taken as representative for the whole BZ:

$$n(\mathbf{r}) = \sum_i \alpha_i \sum_{j=1}^{m} \psi^*_{j,\mathbf{k}_i}(\mathbf{r})\, \psi_{j,\mathbf{k}_i}(\mathbf{r}) \tag{2.19}$$

This maintains the normalisation of (2.18). The energy may then be calculated from (2.7). The concept of the 'unit cell' used here may be extended to the Born von Karman 'supercell' which is used for the purposes of simulation, which may contain an integer number (L, say) of crystallographic unit cells; the Brillouin Zone of the supercell must then be used, with normalisation to $2Lm$ electrons.

Various schemes exist for choosing a suitable set of points [25, 26], which should be distributed over a regular grid throughout the Brillouin Zone, but avoid extremal values of the electron energy $\varepsilon_{j,\mathbf{k}}$. In metals, the abrupt changes in bandstructure at the Fermi surface necessitate the use of a much finer grid [19]. The work described here used the the Monkhorst-Pack [26] set of points, reduced by symmetry to give the minimum number (i.e. the number in the irreducible Brillouin Zone) with appropriate weightings. The use of symmetry is described further in Section 2.6.

2.2.3 Basis sets and the plane wave representation of $f(\mathbf{r})$

We need now to calculate a finite number of wavefunctions at a finite number of reciprocal lattice wavevectors. However, the wavefunctions still contain a cell-periodic function $f(\mathbf{r})$ of apparently arbitrary form, although satisfying (2.13). A plane wave

basis set gives one of the most general forms possible, and can be made lattice-periodic simply by restricting the wavevectors of the plane waves to be only reciprocal lattice vectors **G**.

A wavefunction of (2.15) now takes the form

$$\begin{aligned} \psi_{j,\mathbf{k}}(\mathbf{r}) &= e^{i\mathbf{k}\cdot\mathbf{r}} \sum_{\mathbf{G}} c_{j,\mathbf{k}}(\mathbf{G})\, e^{i\mathbf{G}\cdot\mathbf{r}} \\ &= \sum_{\mathbf{G}} c_{j,\mathbf{k}}(\mathbf{G})\, e^{i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}} \end{aligned} \tag{2.20}$$

The wavefunctions are thus represented as a sum of plane waves with wavevectors $\mathbf{k}+\mathbf{G}$. Each pseudo-electron is still labelled by j, the band, and $\mathbf{k}$, the wavevector.

In theory there should be an infinite number of plane waves contributing to each wavefunction, but in practice the series is truncated once the kinetic energy of the plane waves becomes very high. A *cutoff energy* $\mathcal{E}_c$ is thus set, such that only plane waves with $\hbar^2|\mathbf{k}+\mathbf{G}|^2/2m_e < \mathcal{E}_c$ are used. In practical schemes which use pseudopotentials for the electron-ion interaction (Section 2.3) this will include at least a few thousand plane waves, and often many more, so although the problem is tractable, extremely efficient computational schemes are still required. The choice of $\mathcal{E}_c$ is discussed further in Section 2.2.5.

2.2.4 Representation of Kohn-Sham equations using plane waves

The use of a plane-wave basis set for the wavefunctions has several advantages for the implementation of density-functional theory. It is ideally suited for use with periodic boundary conditions, and many quantities needed in simulations may be easily calculated. For example, the Fourier components of the charge density as given in (2.17) may be shown [27] to be

$$\tilde{\rho}(\mathbf{G}) = e \sum_{j} \sum_{\mathbf{G}'} c^*_{j,\mathbf{k}}(\mathbf{G}') c_{j,\mathbf{k}}(\mathbf{G}+\mathbf{G}') \tag{2.21}$$

One of the main advantages of using such a basis set is that the Kohn-Sham equations take a particularly simple form [27]. Equations (2.8) and (2.9) become the matrix eigenproblem

$$\sum_{\mathbf{G}'} \left[\frac{\hbar^2}{2m} |\mathbf{k}+\mathbf{G}|^2 \delta_{\mathbf{G}\mathbf{G}'} + \tilde{V}_{\text{eff}}(\mathbf{k}+\mathbf{G}, \mathbf{k}+\mathbf{G}') \right] c_{j,\mathbf{k}}(\mathbf{G}') = \varepsilon_i c_{j,\mathbf{k}}(\mathbf{G}) \tag{2.22}$$

where

$$\tilde{V}_{\text{eff}}(\mathbf{k}+\mathbf{G}, \mathbf{k}+\mathbf{G}') = \tilde{V}_N(\mathbf{k}+\mathbf{G}, \mathbf{k}+\mathbf{G}') + \tilde{V}_H(\mathbf{G}-\mathbf{G}') + \tilde{V}_{XC}(\mathbf{G}-\mathbf{G}') \tag{2.23}$$

The electronic kinetic energy takes a diagonal form, since it is simply the weighted sum of the plane wave kinetic energies; $\tilde{V}_H$ and $\tilde{V}_{XC}$ are local potentials so depend on $\mathbf{G}-\mathbf{G}'$ only. Equation (2.22) and (2.23) may be solved by matrix diagonalisation, followed by iteration with equation (2.10) until the eigenstates converge to the ground state. However, more efficient schemes exist and are discussed in Section 2.5.

In practice, the evaluation of the charge density in real space requires a Fourier transform of the contributions from each plane wave. However, fast Fourier transform algorithms all operate over an array of points lying within a parallelopiped, rather than the spherical region of reciprocal space containing the $(\mathbf{k}+\mathbf{G})$ within the energy cutoff. Furthermore, (2.21) shows that the charge density may have Fourier components with wavevectors up to twice the maximum ($\mathbf{G}_c$, say) for the electronic wavefunctions, since $(\mathbf{G}+\mathbf{G}')$ must range from $-\mathbf{G}_c$ to $\mathbf{G}_c$. Hence the charge density Fourier transform uses points in a parallelopiped of reciprocal space which extends at least twice as far in every direction as the original sphere.

Plane-wave basis sets have been used for many years in DFT calculation of periodic, extended systems. However, they have recently been shown [28] to also give good results for isolated molecules, for which more localised sets have been traditionally used. The use of plane waves also has advantages for calculating other properties of a system, such as the forces exerted on the ions (discussed in Section 3.1.1).

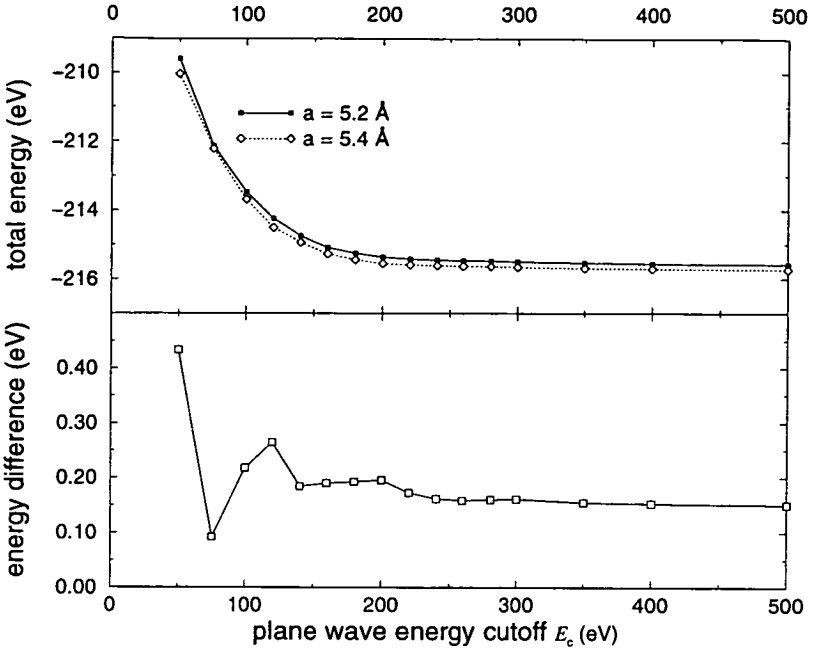


Figure 2.2. Variation of total energy of two atoms of silicon, for two lattice parameters, and the energy difference between them, as the plane wave energy cutoff $\mathcal{E}_c$ is increased.

2.2.5 Determination of the basis set size

The effect on system properties of the energy cutoff $\mathcal{E}_c$ applied to the basis set will be important in much of the work discussed in this thesis, so an introduction is included here as a basis for further discussion.

Since additional plane waves provide further variational freedom for the wavefunctions, the total energy drops monotonically as the energy cutoff increases. Strictly, the energy cutoff to be used in a simulation should be found by increasing its value until the total energy E_{tot} converges to a desired tolerance [29, 30]; the level of convergence may be increased to an arbitrary level by using a higher cutoff. The variation of total energy with energy cutoff is illustrated for silicon in Figure 2.2: in this case a cutoff of at least 300 eV is required for good convergence, but the total energy continues to decrease thereafter.

However, plane waves with high kinetic energy represent rapid fluctuations in the

charge density, which usually occur only in response to an extreme external potential. In the systems considered here, these occur only in the deep potential close to the ions, occupied mainly by the core electrons, and the charge density thus represented plays very little part in any bonding between atoms. The difference in total energy between two slightly different structures (e.g. different cell parameters) should be due only to the charge involved in bonding, so should not be affected by changes in the fine electronic structure close to the nuclei. [19, 31]. Thus a cutoff which allows convergence of the energy *difference* between structures should be sufficient for accurate calculation of most properties. The energy difference between primitive unit cells of silicon having two different lattice parameters is also shown in Figure 2.2, and is seen to converge beyond 220 eV; in practice a cutoff of 250 eV is routinely used. However, cutoffs such as these, which give a tractable number of plane waves, are only obtained by the use of pseudopotentials, which will be described in Section 2.3.

The degree of convergence which has been reached can be quantified as the remaining energy by which the total energy differs from the fully converged value [30]. Another important quantity which measures the degree of basis-set convergence is

$$\frac{\partial E_{\mathrm{tot}}}{\partial \ln(\mathcal{E}_{\mathrm{c}})} \tag{2.24}$$

which will be important for the discussion of basis-set corrections to the stress and total energy in Section 3.2.2.

2.3 Ion-electron interaction: pseudopotentials

The interaction between the nuclei and electrons must now be discussed. Near to the nuclei, the valence electrons not only experience an extremely strong Coulomb potential, but must also be orthogonal to the tightly-bound core electron states. This means that, near to a nucleus, valence wavefunctions oscillate very rapidly with $\mathbf{r}$, and thus an extremely large plane-wave basis set is required.

Pseudopotential theory, developed over the last thirty years or so, is an attempt to model the ion-electron interaction, by screening the nucleus with the 'core' electrons to give a much weaker potential which acts on the valence electrons [32]. The core electrons play a negligible role in the binding properties of solids, so their states are not explicitly required. The total energy calculated in this way is no longer the binding energy of all the electrons to the nuclei, but instead relates only to the valence electrons and the screened nuclei.

In replacing the Coulomb potential with a much weaker potential inside some core region, usually a sphere of radius r_c, the pseudopotential approximation replaces a valence electron wavefunction $\psi(r)$ with a much smoother pseudo-wavefunction $\psi_{ps}(r)$. Ideally, the pseudo-wavefunction should have no nodes inside r_c, so that it is likely to be the ground state of the pseudopotential. However, it should be identical to the true wavefunction, thus generating the correct charge density, beyond r_c. The pseudo-wavefunctions (and hence the corresponding pseudopotential) should match the true functions smoothly at r_c: usually not only $\psi(r)$ and $\psi_{ps}(r)$ but also their first and often second derivatives with respect to r are constrained to be equal at r_c.

Generation of pseudopotentials is based on the results of full-potential calculations of isolated atoms, which solve the Schrödinger equation for all the electrons. For some electronic configuration of the ion, the full wavefunction of a valence electron is smoothed inside inside r_c to give a nodeless form which matches the true wavefunction at and beyond r_c as described above. By inverting the Schrödinger equation, the pseudopotential to which this is a solution is deduced, and then used in solid-state calculations. This process is illustrated in Figure 2.3. The problem of generating pseudopotentials is thus to choose a form for the wavefunction inside r_c which gives as good a pseudopotential as possible.

To correctly reproduce the properties of the full potential, the pseudopotential should give the same scattering behaviour between electrons and the core, imitating the full interaction. This should hold for as large a range of electron energies as possible

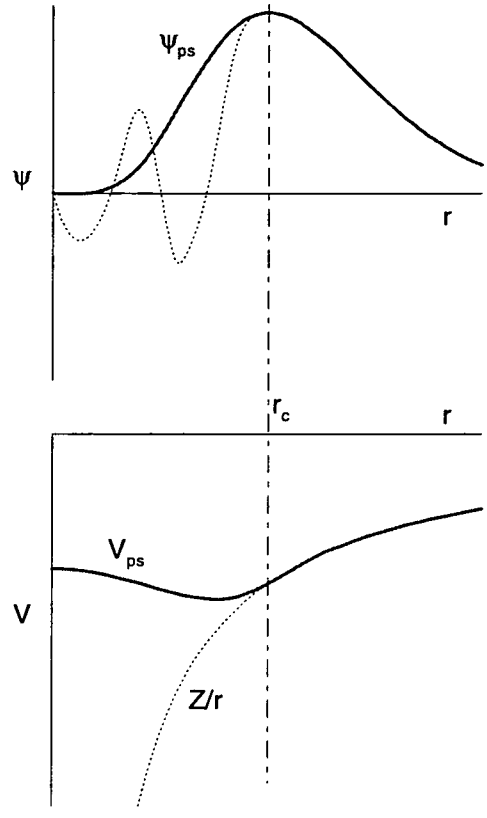


Figure 2.3. Schematic illustration of the process of generating pseudopotentials. The Coulomb potential and all-electron valence wavefunction are shown as dotted lines; the smoothed pseudo-wavefunction and corresponding pseudopotential are shown in solid lines. Both sets of functions must agree beyond r_c.

if the pseudopotential is to be applicable to a variety of solid-state applications. The scattering behaviour can be expressed as the logarithmic derivative of the wavefunction with respect to electron energy [32], and its derivative with respect to energy is related to the amount of charge located in the core region; guaranteeing the right amount of charge inside the core, and thus the right Coulomb potential outside, thus gives the minimum error in the scattering behaviour [33]. If the wavefunctions match exactly beyond r_c, the charge distribution must obey this constraint. Such pseudopotentials are known as *norm-conserving.*

2.3.1 Non-local pseudopotentials

The interaction of an electron with a nucleus screened by core electrons, that is, with a pseudopotential, depends on the angular momentum of the electron. This can be seen by considering, for example, that the orthogonalisation between valence and core electrons depends on whether there exist core electrons with the same angular momentum. A pseudopotential should therefore have a different effect on each angular momentum component of the pseudo-wavefunction, although for some species it has been found that all components may be treated identically with little error. The general form of a pseudopotential can be written

$$V_{NL} = \sum_{lm} \mid Y_{lm}\rangle V_l \langle Y_{lm} \mid \tag{2.25}$$

where the spherical harmonics Y_{lm} project out each individual angular momentum component. However, when ψ is represented by a basis set of plane waves $e^{i(\mathbf{k}+\mathbf{G})}$, evaluation of $\langle\psi^*|V_{NL}|\psi\rangle$ requires that the projection operators be used for each pair of plane waves $\{\mathbf{G}, \mathbf{G}'\}$ in the simulation, so that each pair requires a separate integral. The number of plane waves, N_{PW}, used in realistic calculations is usually of order 10^3–10^4, and can be much higher, so the $N_{PW}(N_{PW}+1)/2$ integrals required at each $\mathbf{k}$-point make this approach computationally infeasible.

The Kleinman-Bylander form for non-local pseudopotentials [34] uses the difference between each component of the pseudopotential and an arbitrary local part which is the same for all components:

$$\delta V_l = V_l - V_{\text{local}} \tag{2.26}$$

The general form (2.25) is then approximated by

$$V_{KB} = V_{\text{local}} + \sum_{lm} \frac{\mid \Psi^0_{lm}\delta V_l\rangle\langle\Psi^0_{lm}\delta V_l \mid}{\langle\Psi^0_{lm} \mid \delta V_l \mid \Psi^0_{lm}\rangle} \tag{2.27}$$

where Ψ^0_{lm} is the eigenstate from which the pseudopotential was generated. The non-local part of (2.27) involves only separate integrals over $\mathbf{G}$ and $\mathbf{G}'$, each of the form $\langle \Psi^0_{lm} \delta V_l \mid e^{i\mathbf{G}\cdot\mathbf{r}} \rangle$. The number of plane-wave integrals that need to be performed is thus greatly reduced, from $\mathcal{O}({N_{PW}}^2)$ to $\mathcal{O}(N_{PW})$: this is the great advantage of using this construction. However, (2.27) is exact only for the atomic eigenstate Ψ^0_{lm}. To be accurate in other situations, the electronic configuration in the solid state must be very close to this eigenstate, further demonstrating that pseudopotentials may not be perfectly transferable from one environment to another.

The Kleinman-Bylander form requires a choice of the local potential V_{local}. This is made to give the best accuracy and computational efficiency. Often one angular-momentum component of the pseudopotential can be selected as the local form, reducing the number of spherical harmonic projectors which need to be used in calculations.

2.3.2 Optimisation with Q_c tuning

The lower the maximum curvature of the pseudo-wavefunction, the fewer the number of plane waves required to expand it satisfactorily, and so the lower the cutoff energy $\mathcal{E}_c$ required. This forms part of the motivation for finding the kinetic energy of the pseudo-wavefunction $\Psi_l(\mathbf{r})$ due to Fourier components above some wavevector Q_c. It was proposed [35, 36] that if $\Psi_l(\mathbf{r})$ is expanded as a series of Bessel functions, the coefficients of each can then be determined by minimising this energy. This should minimise the error when an equivalent cutoff energy is used in solid-state calculations, thus producing the fastest convergence.

This method was developed [30, 37] so that the Q_c parameter is used in a systematic manner to optimise the accuracy of a pseudopotential. The Bessel coefficients are still set as above for each value of Q_c, but Q_c is then tuned to optimise the logarithmic derivative by adjusting the shape of the pseudo-wavefunction. As discussed above, a transferable pseudopotential will have a logarithmic derivative which agrees with the all-electron results as closely as possible. These 'Q_c tuned' pseudopotentials are used

for most of the applications in this work.

2.3.3 Transferability and testing of pseudopotentials

The specific electronic configuration used in the construction of pseudopotentials may be that of the isolated neutral atom or an excited state (which may have fractional occupation numbers) which is ionised to ensure that the highest occupied states are still bound. Different angular-momentum pseudopotential components may be derived from different configurations; this is necessary when not all components are found in the ground state of the atom. Those suggested in the comprehensive work of Bachelet, Hamann and Schlüter (BHS) [38] are widely used.

However, solid-state applications may require simulations of electronic configurations which are rather different to those of the atomic reference states, and pseudopotentials are thus needed that will adequately represent a range of electronic environments. This property is known as transferability, and is related by many workers [33, 36, 37, 38] to how well the logarithmic derivative of the pseudo-wavefunction reproduces that of the full-electron wavefunction, and over what range of energies.

Use of pseudopotentials in the solid state produces one of the most stringent tests; transferable pseudopotentials should reproduce well the properties of a range of materials. It would be expected that pseudopotentials should give values of lattice constant, bulk modulus etc. close to all-electron calculations, since these must also use an approximation for exchange and correlation, such as the LDA [39], even if both differ from experiment. However, the more solid-state testing which needs to be done against known results, the weaker the predictive power of such simulations.

2.4 Ion-ion and long-range Coulomb interactions

Although the Born-Oppenheimer approximation considers the ions to be static while the electronic charge density is found, a comparison of the energies of different structures is only valid if the Coulomb interaction between the ions is included. Calculation of this energy is not straightforward as the Coulomb interaction is extremely long-ranged. The summation is instead split over real and reciprocal space, using the Ewald method [40]; this renders each part rapidly convergent. However, the $\mathbf{G}=0$ contribution is still divergent, as might be expected in an infinite crystal such as that represented by periodic boundary conditions.

The total ionic potential acting on an electronic wavefunction will also be made up of contributions from all the ions in the crystal, and the Coulomb part of the pseudopotential will also give a divergent $\mathbf{G}=0$ term in an infinite crystal. However, this term is matched exactly by the $\mathbf{G}=0$ terms of the ion-ion energy sum and the electron-electron sum. These three terms cancel one another and are thus removed from the computation [41, 42]. The difference between the Coulomb $\mathbf{G}=0$ term and that of the pseudopotential can be evaluated, giving rise to a 'core energy' contribution from each ion [19].

However, neither a constant potential nor a non-periodic one (such as $V(\mathbf{r}) = V_0\mathbf{r}\cdot\hat{\mathbf{a}}$ for some V_0 and direction $\hat{\mathbf{a}}$) may be represented without $\mathbf{G}=0$ terms. The effect of excluding the constant potential is the desired effect, but excluding $V = V_0\mathbf{r}\cdot\hat{\mathbf{a}}$ excludes the resulting constant field $\mathbf{E}(\mathbf{r}) = \mathbf{E}_0 = V_0\hat{\mathbf{a}}$. This is entirely consistent with the use of periodic boundary conditions, which prohibit a non-periodic potential such as that which produces a constant electric field.

Such a field should be produced, for example, by an atomic configuration which leaves each unit cell with the same net dipole: there would then be a macroscopic electric field throughout the infinite crystal. The formalism of periodic boundary conditions and exclusion of $\mathbf{G}=0$ terms thus cannot properly describe such configurations [43, 44]. It may be noted that this method has recently been applied to isolated molecules [28],

which often have a dipole moment. However, if the supercell is large enough for the molecule to be considered isolated, there will be negligible charge density at the 'edge' of the cell, between molecules. This allows the electric potential to change in order to recover periodicity (creating a smoothed 'saw-tooth' profile for $V(\mathbf{r})$ rather than one which increases monotonically) with negligible effect on the energy; the dipole moment is thus accurately simulated. The same possibility for compensation for macroscopic fields exists in simulations of surfaces separated by a large distance.

However, in periodic crystals a macroscopic electric field is still outwith these simulations. Such an electric field is also produced in longitudinal optic (LO) phonons [44], close to Γ ($\mathbf{q} = 0$), in which atoms are displaced parallel to the wavevector, setting up a macroscopic dipole. In a cubic crystal, this effect splits the three-fold degeneracy[2] which might be expected at Γ: the TO modes are unaffected and the Lyddane-Sachs-Teller relation [48] relates the two frequencies, if there are only two atoms in the basis:

$$\frac{\omega_{LO}^2}{\omega_{TO}^2} = \frac{\epsilon_0}{\epsilon_\infty} \tag{2.28}$$

However, since this field may not be represented with periodic boundary conditions, the LO-TO splitting is not observed, although the TO modes are calculated correctly. This is one of the very few restrictions of this approach; the problem is discussed further in Chapter 4.

2.5 Molecular dynamics

The Kohn-Sham equations (2.22) may be solved by matrix diagonalisation at each k-point, after which the charge density and resulting potentials are recalculated, and the process iterated until the ground state charge density is found. However, this requires

[2]Strictly, there should still be full degeneracy [45, 46] at $\mathbf{q} = 0$, but even in Raman scattering experiments, which probe the phonon dispersion very close to $\mathbf{q} = 0$, a phonon with strictly zero wavevector would scatter light directly along the incident beam and would not be detected. At this level of detail, however, there has been discussion as to whether normal modes really are periodic [47].

diagonalisation of very large matrices ($N_{PW} \times N_{PW}$), which is extremely computationally intensive.

The correct coefficients in the plane-wave expansion of the ground state wavefunctions minimise the total energy E as given by (2.7). Car and Parrinello [49] converted the problem to one of classical dynamics by considering each coefficient as a classical particle, with a fictitious mass μ. A Lagrangian was then defined

$$\mathcal{L} = \sum_i \mu \langle \dot{\psi}_i \mid \dot{\psi}_i \rangle - E[\{\psi_i\}, \{\mathbf{R}_I\}, h] \tag{2.29}$$

where $\{\mathbf{R}_I\}$ are the nuclear positions and h is a matrix giving the size and shape of the unit cell (see Section 3.2.1). From this Lagrangian, equations of motion for the coefficients $c_{j,\mathbf{k}+\mathbf{G}}$ of the wavefunctions can be derived. If a damping term is introduced, the minimum of E can be found by a simulated annealing process. After each molecular dynamics step, the wavefunctions must be re-orthonormalised. This can be done using the Gram-Schmidt orthogonalisation scheme.

The Car-Parrinello molecular dynamics method is, however, susceptible to instabilities and fluctuations. An alternative approach is to directly minimise the energy as a function of the wavefunction coefficients, since there should be only one well-defined minimum[3]. One of the most efficient methods is the conjugate-gradients algorithm, which takes successive steps 'downhill' in multidimensional space, utilising information from previous steps to make the best choice of search direction. The efficiency can be increased still further if preconditioning is used, which effectively maps the energy surface $E\{c_{j,\mathbf{k}}(\mathbf{G})\}$ to a more isotropic form [51].

[3] In some circumstances, the node-free pseudo-wavefunction is not in fact the lowest-energy state; so-called 'ghost' states with lower energy may arise as an artefact of the pseudopotential generating procedure. However, it has been shown [50] that such states may be analytically detected and avoided in Kleinman-Bylander pseudopotentials.

2.6 Symmetry

It is natural to consider symmetry when calculating the structure and properties of crystals from first principles. Bloch's Theorem exploits the translational symmetry that occurs in all perfect crystals, but any additional symmetry in a structure may also be used to reduce the amount of computational effort required and also ensure that the results are physical.

The symmetry of a crystal structure is represented by a space group R containing N_R symmetry elements, each denoted $\{\mathbf{S}|\mathbf{v}\}$, where $\mathbf{S}$ is a 3×3 matrix corresponding to a rotation, reflection or inversion, and $\mathbf{v}$ is a translation by an amount other than a primitive lattice translation, set for simplicity to a translation within the unit cell. The effect of symmetry elements may be expressed in terms of the mapping of one ion to another under each symmetry element; this will be important for force and phonon calculations.

The use of symmetry with plane-wave basis sets is well described elsewhere [27]: the main results are summarised here. The band structures at two symmetry-related $\mathbf{k}$-points must be equal, and so only one calculation is required to sample both. The Monkhorst-Pack special $\mathbf{k}$-points may thus be reduced in number: those $\mathbf{k}$-points in the irreducible part of the Brillouin Zone are weighted according to the number of symmetry-related points in the entire BZ; the resulting charge densities are then added in a weighted sum as shown in (2.19). If the charge density is symmetrised, so is the resulting potential; the reduced $\mathbf{k}$-point set must then give the same results as summing over all the original $\mathbf{k}$-points.

It must be ensured that the charge density obeys the requirement of each symmetry element $\{\mathbf{S}|\mathbf{v}\}$ that is:

$$\rho(\mathbf{r}) = \rho(\{\mathbf{S}|\mathbf{v}\}\mathbf{r}) \tag{2.30}$$

However, to use special $\mathbf{k}$-point sampling as described above, each special $\mathbf{k}$-point (say

$\mathbf{k}'$) should relate to a charge density contribution $\langle\rho_{\mathbf{k}'}(\mathbf{r})\rangle_R$ averaged over all symmetry-related k-points:

$$\langle\rho_{\mathbf{k}'}(\mathbf{r})\rangle_R = \frac{1}{N_R}\sum_{\mathbf{S}\in R}\rho_{\mathbf{Sk}}(\mathbf{r}) \tag{2.31}$$

so enforcing (2.30) requires calculation of the charge density from the symmetry-related points $\mathbf{Sk}'$ as well as the special k-points, seemingly defeating the purpose of using a reduced set. However, if the charge density is symmetrised in reciprocal space, by relating Fourier transformed elements of charge density $\tilde{\rho}(\mathbf{G})$ and $\tilde{\rho}(\mathbf{G}')$, it may be shown [27] that only the charge density from the special k-points is required. The requirement for symmetry (2.30), when applied to the Fourier transform of the charge density, with a plane-wave basis set, becomes

$$\tilde{\rho}_{\mathbf{k}'}(\mathbf{q}) = \tilde{\rho}_{\mathbf{k}'}(\mathbf{S}^{-1}\mathbf{q})\, e^{i\mathbf{q}\cdot\mathbf{v}(\mathbf{S})} \tag{2.32}$$

The plane wave representation of wavefunctions makes a Fourier treatment straightforward. In practical schemes, (2.32) may be implemented by averaging the total charge density from all special $\mathbf{k}'$, over all the effects of all N_R symmetry elements:

$$\langle\tilde{\rho}(\mathbf{G})\rangle_R = \frac{1}{N_R}\sum_{\mathbf{S}\in R} e^{i\mathbf{G}\cdot\mathbf{v}(\mathbf{S})}\,\tilde{\rho}(\mathbf{S}_J^{-1}\mathbf{G}) \tag{2.33}$$

This symmetrisation is performed after initialising and updating the charge density.

2.7 Computational details

The CASTEP (Cambridge Serial Total Energy Package) and CETEP (Cambridge-Edinburgh Total Energy Package) codes were used throughout this work [19]. They are widely used by members of the United Kingdom Car-Parrinello consortium, on a range of serial and parallel machines. Developments to the code that were made as part of the work described in this thesis are described in the following Chapters.

CASTEP uses a plane-wave basis set for the wavefunctions, as described above, together with conjugate-gradients minimisation [51] to find ground state energy. The local density approximation is commonly used, although the general gradients correction scheme is also available. The minimisation of the energy is carried out band by band and for each k-point in turn, from an initial random assignment of the plane-wave coefficients. CETEP is a parallel implementation of this code, in which the plane waves and real space grid points are distributed over different processors [52].

The total energy, charge distribution, forces on the ions (see Section 3.1.1) and stresses on the unit cell (Section 3.2.1) are all calculated, and the symmetrisation scheme described above is also implemented. The various ways in which these codes may be used to investigate various solid-state properties is developed in the following Chapters.

Chapter 3

Determination of equilibrium structures

All this brings us to the question of stresses and strains, words which the layman is apt to regard as alarming, distressing and confusing. This is perhaps partly because the words may conjure up the idea of a wilderness of mathematics but probably more because the words have been borrowed or stolen by non-scientists to describe the mental state of human beings.

J.E. Gordon, *The New Science of Strong Materials*

Introduction

All observable properties of a crystal structure are in some way related to the total energy. The total energy in the pseudopotential approximation is neither the cohesive energy nor the full electronic binding energy, and so is of little consequence in itself. However, a great deal of useful information may be obtained by comparison between energies of alternative structures containing the same number of each species of atom. The use of pseudopotentials is actually advantageous in this respect because small energy differences become a much greater fraction of the total energy, and thus have greater computational accuracy. This general approach leads not only to determination of equilibrium structures, but also to calculation of forces on ions, stresses on the unit

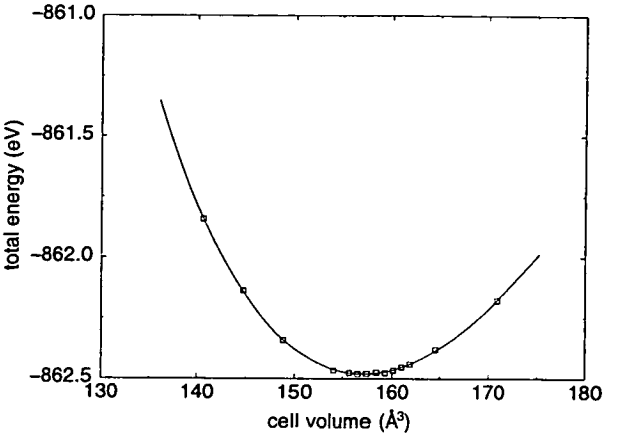


Figure 3.1. Fit to total energy of diamond-structured silicon at various lattice parameters, with a Birch-Murnaghan equation of state [53], giving the equilibrium structure at $a_0 = 5.395$ Å.

cell, and the vibrational properties of crystals.

The simplest case of the variation of total energy with structure is that of a cubic crystal with no internal parameters, such as silicon in the face-centred cubic diamond structure. The total energy is then a function of a single lattice parameter, and that value of the lattice parameter which minimises the total energy gives the equilibrium structure (Figure 3.1). The pressure required to sustain any other lattice parameter is then the negative of the gradient of E with respect to the volume Ω, and the bulk modulus may be found from the curvature of the graph; this is usually done via a polynomial or more a sophisticated fit to a series of energies. Hence macroscopic properties may be calculated from an *ab initio* simulation of just one unit cell if periodic boundary conditions are employed. The results may be compared to experiment if confirmation of the method is required, or used to predict the properties of materials which may not be easily synthesised or investigated experimentally.

The method of finding the equilibrium lattice parameter of a cubic crystal illustrates the underlying requirement for equilibrium structures to minimise energy. However, in general a crystal structure will have up to six independent lattice parameters (i.e. three lengths and three angles in the monoclinic case) and any number of internal parameters

describing the ionic basis. The problem of finding the equilibrium structure, i.e. that which minimises the total energy, is no longer so simple. A tetragonal unit cell with at most one internal parameter x may still be optimised by fitting a multidimensional energy surface $E(a, c, x)$ with respect to all three parameters, but even a slightly more complicated structure renders this approach computationally impractical. Alternative approaches must thus be found.

The methods for finding the equilibrium structures of crystals from total energy calculations are described in this Chapter, whereas the following Chapter presents methods of finding the vibrational behaviour of a given structure. Both these techniques are used in subsequent Chapters.

3.1 Optimisation of ionic positions

The set of equilbrium positions for a number of ions within a given cell is only a subset of the full set of structural parameters. However, if they can be determined this will give at least a smaller number of variables for which a multidimensional energy function must be fitted. Since at the equilibrium positions there must be no net forces on any of the ions, if the forces for a given ionic configuration can be calculated, the ions may be relaxed towards their equilibrium positions. It will be shown in the following Chapter that ionic force calculations can also be used to find the vibrational properties of an equilibrium structure.

3.1.1 *Ab initio* force calculation

The force on the Ith ion, at position $\mathbf{R}_I$, is in general given by

$$\mathbf{F}_I = -\frac{dE}{d\mathbf{R}_I} \tag{3.1}$$

but the energy must depend on the many-electron wavefunction Ψ. The Born-Oppenheimer approximation means that Ψ is a function only of the instantaneous ionic positions, (rather than, say, their momenta also) i.e. $E = E[\Psi(\{\mathbf{R}_I\}), \{\mathbf{R}_I\}]$; the solutions under this approximation are said to lie on the Born-Oppenheimer surface in the multidimensional space of $(E, \Psi, \mathbf{R}_I)$. The wavefunction must therefore change in response to a change in ionic positions, in order to remain on this surface, i.e. for the total energy to be minimised for the new ionic configuration. The change in wavefunction with $\mathbf{R}_I$ must therefore be taken into account, by expanding (3.1)

$$\mathbf{F}_I = -\frac{\partial E}{\partial \mathbf{R}_I} - \frac{\partial E}{\partial \Psi}\frac{d\Psi}{d\mathbf{R}_I} - \frac{\partial E}{\partial \Psi^*}\frac{d\Psi^*}{d\mathbf{R}_I} \tag{3.2}$$

where $E = \langle \Psi^* | H | \Psi \rangle$. However, the Hellmann-Feynman theorem [54] argues that since the wavefunction is an eigenstate of the Hamiltonian, then $\partial E/\partial \Psi^* = H\Psi = E\Psi$, and the second two terms in (3.2), which involve the wavefunctions, become

$$-\left(\frac{\partial E}{\partial \Psi}\frac{d\Psi}{d\mathbf{R}_I} + \frac{\partial E}{\partial \Psi^*}\frac{d\Psi^*}{d\mathbf{R}_I}\right) = -E\frac{\partial}{\partial \mathbf{R}_I}\langle \Psi^* \Psi \rangle \tag{3.3}$$

Since $\langle \Psi^* \Psi \rangle$ is the normalisation constant, its derivative must be zero, so the last two terms in (3.2) must cancel each other out. The physical force on each ion may then be obtained simply by taking the derivative of the total energy with respect to ionic position, using the many-body wavefunction which is an eigenstate for the instantaneous ionic configuration:

$$\mathbf{F}_I = -\frac{\partial E}{\partial \mathbf{R}_I}. \tag{3.4}$$

Forces calculated thus are known as Hellmann-Feynman forces. In the density-functional formalism, the energy is a function only of the charge density, which is itself calculated from individual electron wavefunctions ψ_i. Calculation of the forces, therefore, also requires only the charge density, not the individual wavefuctions, within the

DFT formalism for $E[n(\mathbf{r})]$.

However, the Hellmann-Feynman theorem only holds when the wavefunctions are exact eigenstates of the Hamiltonian. It can be shown [55] that the total energy has only a second order error with respect to deviations in the electronic wavefunction from the ground state, which will occur during the process of solution by iterative convergence. However, the forces have a first order error; the forces thus converge rather more slowly than the total energy in a self-consistent calculation. If the forces are to be used to optimise ionic positions, it must be ensured that the wavefunctions have converged sufficiently towards the ground state for each ionic configuration before the ions are moved, otherwise instabilities may result.

However, in all practical simulations, it is impossible to make the wavefunctions *exact* eigenfunctions of the Hamiltonian, if only because of the use of a finite basis set. If the wavefunctions are represented using a basis set such that $\psi_i = \sum_j c_{ij}\phi_j$, then the energy can be written $E(\{\mathbf{R}_I\}, \{c_{ij}\}, \{\phi_j\})$ [56]. The expansion of (3.1) should then include the effect of moving the ions on the basis set, by further expanding the terms involving Ψ:

$$\frac{\partial E}{\partial \Psi}\frac{d\Psi}{d\mathbf{R}_I} = \sum_{ij}\frac{\partial E}{\partial c_{ij}}\frac{dc_{ij}}{d\mathbf{R}_I} + \sum_{j}\frac{\partial E}{\partial \phi_j}\frac{d\phi_j}{d\mathbf{R}_I} \tag{3.5}$$

If the wavefunctions have completely converged within a calculation, then $\partial E/\partial c_{ij}$ is zero due to the variational nature of the Kohn-Sham theorem, and therefore the first term is zero [56]; this is a slight relaxation of the conditions under which the Hellmann-Feynman theorem holds. The second term gives rise to an effect known as the Pulay force [57] if the wavefunctions are not exact (i.e. if the Hellmann-Feynman theorem does not hold). This can occur if the basis set is not sufficiently complete to accurately describe the exact eigensolutions. However, a plane wave basis set is totally independent of ionic positions, so $d\phi_j/d\mathbf{R}_I = 0$ and the Pulay force is zero. Thus if a plane wave basis set is used, the error in using the Hellmann-Feynman force is directly related to the extent to which the calculation has converged, and may therefore be reduced to

an arbitrarily small value. Plane wave basis sets thus have a great advantage over less general sets for the calculation of forces, relaxation of ions and calculation of vibrational properties. However, a plane wave basis set is usually defined by the reciprocal vectors of the supercell, and does not remove the analogous Pulay stress; this is discussed in Section 3.2.2.

If other basis sets are used, the forces can only be found by complicated calculation of the Pulay force or by finding the total energy as a function of ionic position, requiring much more computational effort. Basis sets which are localised around the ions, designed to mimic atomic orbitals, must be moved with the ions if they are to adequately describe wavefunctions, and hence the second term in (3.5) will be large, giving a Pulay force which may not be ignored.

3.1.2 Relaxation of ions

If the forces on the ions are known, they may be relaxed to their equilibrium configuration (for a given unit cell) by moving them until the forces are all zero. The simplest 'steepest descents' scheme simply moves the ions according to the calculated forces at a given ionic configuration. Each each ion is moved in turn in the direction of the force, the charge density and hence forces are recalculated, and the process iterated until the ionic positions have converged.

Since the calculated Hellmann-Feynman forces are correct only when the wavefunctions have completely converged, the wavefunctions must be brought back onto the Born-Oppenheimer surface for each new ionic configuration during relaxation, as discussed in Section 3.1.1. The forces will inevitably have some error due to incomplete convergence of the wavefunctions; near to the ionic minimum, therefore, the wavefunctions have to be relaxed to a fine tolerance otherwise this 'noise' in the forces can cause fluctuation in ionic positions.

The steepest descents algorithm for finding the equilibrium ionic configuration may be replaced by the more sophisticated conjugate gradients scheme [19]. This mirrors

that used to minimise the total energy as a function of the plane wave coefficients, and for most systems is more efficient than steepest descents. It uses information from previous ionic displacement steps to determine the optimum search direction, and estimates the minimum energy along each direction from a number of energy evaluations. However, this efficiency of this method is greatly reduced if the energy surface is very anharmonic.

3.1.3 Symmetrisation of forces

The symmetry of the crystal may be further exploited to ensure that forces conform to the space group of the structure. This is particularly helpful if the forces are to be used to relax the ions to equilibrium, since the space group will then be preserved during relaxation. If an element $\{\mathbf{S}|\mathbf{v}\}$ maps atom I to atom J, and $\mathbf{F}_I$ is the force on the Ith atom, then the symmetry requires that

$$\mathbf{F}_J = \mathbf{S}\mathbf{F}_I \tag{3.6}$$

This set of requirements is enforced by averaging over pairs of forces which should be equal. Symmetry analysis of the structure generates a mapping list F_0 such that atom I is mapped by the Rth symmetry element to $J = F_0(R, I)$. The symmetrisation may be performed by averaging over forces which should be equal, using $\mathbf{S}^{-1} = \mathbf{S}^T$:

$$\mathbf{F}_I = \frac{1}{N_R} \sum_{R=1}^{N_R} \mathbf{S}_R^T \mathbf{F}_{F_0(R,I)} \tag{3.7}$$

This symmetrisation is performed in CASTEP before either writing out the forces or using them to move the ions.

3.2 Finding equilibrium structures and phases

If for each unit cell, the ions are relaxed to their equilibrium positions as described above, the resulting total energy can then be assigned to the corresponding lattice parameters **a**, **b** and **c**. An energy surface $E(\mathbf{a}, \mathbf{b}, \mathbf{c})$ can then be fitted to the results of total energy calculations for a variety of cells; the lattice parameters which minimise the energy are then predicted to be those of the equilibrium structure. The response of the structure to an applied pressure can then also be deduced, by finding the minimum energy for a given volume and hence $E(V)$. However, this minimisation may be carried out directly, as described in Section 3.2.

Such procedures are, however, restricted to optimising a structure only within the constraints of any initial symmetry or lattice periodicity; only the nearest local minimum will be found, not a global minimum. In many cases the energies of different local minima, i.e. different structures, will have to be found in order to determine the equilibrium structure of a crystal. If there are a number of possible structures, for example, cubic, tetragonal and orthorhombic, they may be compared by plotting energy versus volume curves for equal numbers and species of atoms. That phase having the overall lowest energy is then favoured over the others at zero pressure. Since the tangent of an energy-volume curve gives the pressure at which that phase is stable (at which its enthalpy is minimised), a tangent which touches two phases will give pressures at which both are in equilibrium, and hence the transition pressure; this was first performed for phases of silicon [58].

Since such calculations can never investigate every possible structure, it cannot ever be guaranteed that the global minimum has been found. Nor can the mechanism, if any is available, of phase transitions between such phases be unambiguously deduced by these simulations alone. We discuss in Section 4.4 some ways in which indicators of lower-energy solutions may be obtained, and possible mechanisms which may be elucidated from first principles.

In general, it would be preferable to find the equilibrium unit cell and atomic positions via one structural relaxation scheme, rather than by partial structural refinements and parametrised energy surfaces. Structures with a complicated ionic basis may require large amounts of computing resources to relax the ions; to do this for each set of lattice parameters was found to be impractical for anything less symmetric than a tetragonal cell.

The equilibrium unit cell is that which minimises the energy with respect to the lattice parameters, and therefore has zero stress; the ions must also be in their equilibrium positions and thus all forces must be zero. This is the basic condition which underlies strategies to relax the cell to equilibrium.

3.2.1 *Ab initio* stresses in plane-wave calculations

To relax the unit cell, the stress on the unit cell for a given ionic configuration must be found; the cell parameters may then be adjusted in response to the stress, in order to find the cell which gives zero stress. If we write $h_0 = \{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$ as a 3×3 matrix giving the elements of the unit cell, the cell volume is given by $\Omega_0 = |h_0|$. The strain matrix ϵ gives $h \rightarrow (1 + \epsilon)h_0$. The stress may then be formally given as the derivative of the total energy per unit volume with respect to strain ϵ:

$$\sigma_{\alpha\beta} = \frac{1}{\Omega}\frac{dE}{d\epsilon_{\alpha\beta}} \tag{3.8}$$

If the strain and strain are isotropic, i.e. $\epsilon_{\alpha\beta} = \epsilon\delta_{\alpha\beta}$, then the volume Ω is a function only of the strain ϵ, and $\partial\Omega/\partial\epsilon_{\alpha\beta} = \Omega_0\delta_{\alpha\beta}$. The isotropic stress is then $\sigma_{\alpha\beta} = \sigma\delta_{\alpha\beta}$, where

$$\begin{aligned} \sigma &= \frac{1}{\Omega}\frac{dE}{d\epsilon} \\ &= \frac{1}{\Omega}\frac{\partial E}{\partial\Omega}\frac{\partial\Omega}{\partial\epsilon} \\ &= \frac{\partial E}{\partial\Omega} \end{aligned} \tag{3.9}$$

With this sign convention, the stress is positive when the system is held at an expanded volume. In the Kohn-Sham formulation, E must mean the total energy when the electronic and ionic degrees of freedom have converged to equilibrium. However, it is also a function of the basis set used, or in this work, the set of plane waves in the simulation.

An expression for the quantum-mechanical (QM) stress was given by Nielsen and Martin [59, 60, 61], which requires the stress to be calculated using a constant number of plane waves N_{PW}. In contrast, conventional structural determination (via energy calculation at different lattice parameters) has always used a constant plane-wave energy cutoff over all the simulations. Keeping N_{PW} constant will alter the maximum plane wave kinetic energy, and thus the effective cutoff, as the cell changes, so a choice must be made between keeping a constant cutoff or a constant number of plane waves.

The minimum of energy for constant N_{PW} (or equivalently, zero stress) does not occur at the same lattice parameters as that for constant cutoff, as may be seen in Figure 3.2 (in which a correction is shown which is described below). The discrepancy should reduce as the size of the basis set is increased, since the effect of adding a single extra plane wave becomes proportionally less, and higher energy plane waves will be less populated with electronic charge, until the effect becomes negligible. It has been found by many workers (reviewed in [62]) that the use of a constant cutoff gives more satisfactory results, i.e. the equilibrium lattice parameters thus predicted are closer to the results in the limiting case of a very large basis set, where the calculation is almost completely converged. This can be understood in terms of the fact that a constant cutoff describes a limit of resolution in real space which is independent of the size or shape of the cell. If the unit cell becomes smaller, say, and the number of plane waves is kept constant, the wavelength of each decreases, and the kinetic energy of each increases. This increases the resolution in real space, and decreases the total energy by allowing a better description of short wavelength components of the wavefunction. Keeping a constant number of plane waves therefore biases the minimum of energy

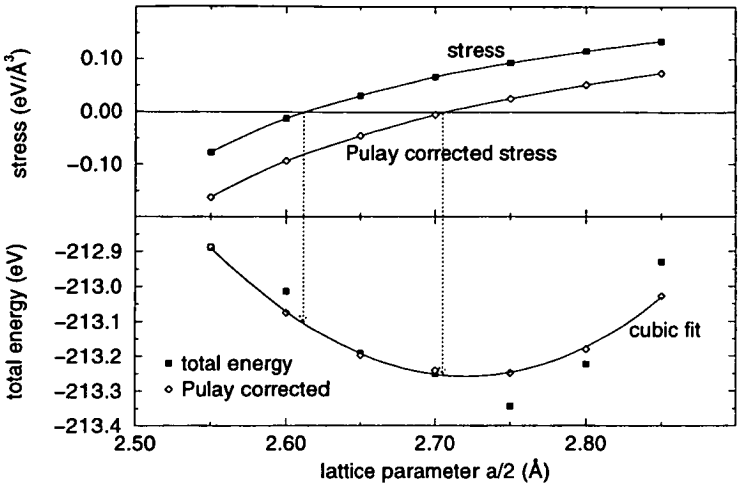


Figure 3.2. Energy and stress versus volume for a primitive unit cell of silicon (2 atoms), without (filled squares) and with (open diamonds) the Pulay correction. An abnormally low cutoff (100 eV) was used to show clearly the effect of the finite basis set, and $2 \times 2 \times 2$ k-point mesh.

towards a smaller volume. It is thus desired to find the lattice parameters obtained by minimising the energy with a constant cutoff.

However, this is awkward to implement in practice. Firstly, it would require adding or removing plane waves as the cell changes during the simulation, and the subsequent redistribution of charge, although this has recently been done [9]. Secondly, it is impossible to enforce exactly due to the discretisation of $|\mathbf{k}+\mathbf{G}|$: there must be an integer number of plane waves in a practical simulation. If instead N_{PW} is kept constant while the supercell is changed, the occupation $c_{j,\mathbf{k}+\mathbf{G}}$ of each plane wave may initially be transferred to the corresponding wave in the new supercell; the wavefunctions must in any case be recalculated and converged before proceeding further, so any departure from the converged wavefunctions made by doing this will be rectified. The total electronic charge is thus preserved throughout the simulation. The stress may then be calculated according to [59], but will not agree with $dE/d\epsilon_{\alpha\beta}$ at constant cutoff. The stress as it stands may not therefore be used for relaxation of the unit cell.

3.2.2 Pulay stresses

The effect on the stress of changing the basis set is analogous to the effect of the basis set on calculated forces, and is thus known as the Pulay stress. An estimate of this stress was derived by Froyen and Cohen [63] and developed by Francis and Payne [31]; a summary of the assumptions and results is given here. The aim of this analysis is to find a correction which gives the stress as if a constant cutoff were used, starting from the calculated stress using constant N_{PW}:

$$\sigma_N = \left.\frac{\partial E}{\partial \epsilon_{\alpha\beta}}\right|_N \tag{3.10}$$

If the plane waves are considered to be distributed homogeneously through reciprocal space, the non-integer number N_c having kinetic energy less than the cutoff energy $\mathcal{E}_c$, for a cell volume Ω, is given by

$$N_c = \frac{\Omega}{6\pi^2}\left(\frac{2m\mathcal{E}_c}{\hbar^2}\right)^{\frac{3}{2}} \tag{3.11}$$

The actual number of plane waves used in a simulation, N'_{PW}, will differ from this theoretical value, but N_c can serve as an approximation. If the approximation is made that $N'_{PW} = N_c = N$, then

$$\frac{\partial E}{\partial N} = \frac{\partial E}{\partial \mathcal{E}_c}\frac{\partial \mathcal{E}_c}{\partial N} \tag{3.12}$$

and, furthermore, $\partial\mathcal{E}_c/\partial N = 2\mathcal{E}_c/3N_c$ from (3.11). The stress should be given by the total derivative of energy $E(h, N)$ at constant cutoff with respect to changes in the boxmatrix h or, equivalently, the strain ϵ. This derivative may be expanded as

$$\sigma_{\alpha\beta} = \frac{dE}{d\epsilon_{\alpha\beta}} = \left.\frac{\partial E}{\partial \epsilon_{\alpha\beta}}\right|_N + \left.\frac{\partial E}{\partial N}\right|_\epsilon \frac{dN}{d\epsilon_{\alpha\beta}}. \tag{3.13}$$

The first term on the right hand side is simply σ_N, the quantum-mechanical stress calculated according to (3.10) [59]; the second term is thus the correction required to

obtain the stress for constant cutoff. However, according to (3.11), only changes in the volume have any effect on the number of plane waves, and so the effect on the resulting stress must be isotropic, since purely shear distortions do not change the volume to first order. Only the derivative of N with respect to the volume is thus required, in analogy with (3.9), and from (3.11) it is found that $\partial N/\partial\Omega = N_c/\Omega$. Using (3.12) it is found that the second term of the expansion in (3.13) gives the isotropic Pulay correction to the stress:

$$\begin{aligned} \sigma_P &= \left.\frac{\partial E}{\partial N}\right|_\Omega \frac{dN}{d\Omega} \\ &= \frac{2}{3\Omega}\frac{\partial E}{\partial(\ln \mathcal{E}_c)} \end{aligned} \tag{3.14}$$

An approximation to the stress, including Pulay terms, is thus given by

$$\sigma = \sigma_N + \sigma_P \tag{3.15}$$

To implement this correction, the 'convergence gradient' $\partial E/\partial(\ln \mathcal{E}_c)$ is required, as introduced in Section 2.2.5. This can be obtained from a linear or polynomial fit to the total energies from a series of simulations using different cutoff energies for the same volume, in the same way that the cutoff energy is determined. It should then hold for that system at any reasonable volume, because the changes in energy stem mainly from changes around the atomic cores (as described in Section 2.2.5); it is of course proportional to the number of unit cells simulated. Because increasing the number of plane waves (by increasing $\mathcal{E}_c$) always decreases the total energy E by providing additional variational freedom, the convergence gradient is always negative. This corresponds to a Pulay stress which, on relaxation, expands the crystal relative to the uncorrected structure. The magnitude of the convergence gradient decreases monotonically as the basis set increases in size, and is zero when the energy has converged.

An example of the effect of using the Pulay correction is shown in Figure 3.2, where

a very low plane-wave energy cutoff is used. This shows clearly the translation of the stress so that zero stress corresponds closely with the minimum of energy. However, the use of a low cutoff results in not only a large value of the convergence gradient, but also a significant error in its measurement and the linear assumption made above, so the correction is not perfect. At more realistic cutoffs these errors are much smaller but must still be considered: this is discussed further in Section 3.2.4.

The Francis-Payne approximation to the Pulay stress given above relies on the approximation of treating the number of plane waves as if they were homogeneously distributed in reciprocal space, as given in (3.11). The deviation from the homogenous reciprocal space described in (3.11) also directly affects the calculated total energy: $E(V)$ curves will not be smooth on a fine scale since at the point at which changes in volume allow an extra plane wave to be accepted, the energy decreases discontinuously due to the additional variational freedom. The is usually only a minor effect, if a large enough cutoff is used, but can produce errors in the fit which result in inaccurate equilibrium lattice parameters [31, 64]. A similar procedure to that of (3.13) may be followed, which is equivalent to integrating the corrected stress, to give a Pulay correction to the energy [31]:

$$E_{\text{tot}}(\text{cutoff}) = E_{\text{tot}}(N) - \frac{2}{3}\frac{\partial E}{\partial(\ln \mathcal{E}_c)} \ln\left(\frac{N'_{PW}}{N_c}\right) \tag{3.16}$$

This involves calculation of the ratio of the theoretical number of plane waves to be used, N_c, and the actual number used N'_{PW}, to obtain $\Delta \ln N$. The same convergence gradient is used as in (3.14) for the stress correction. The effect of this correction is also shown in Figure 3.2 where the energy-volume curves is smoothed considerably. If more than one **k**-point is used, with different numbers of plane waves N'_i and weightings α_i at each (Section 2.2.2), then the geometric weighted average is used [31], i.e. $N'_{PW} = \prod_i (N'_i)^{\alpha_i}$, since (3.16) involves taking the logarithm of N'_{PW}. In practice, the numbers of plane waves at different **k**-points are very close to each other, and using the weighted

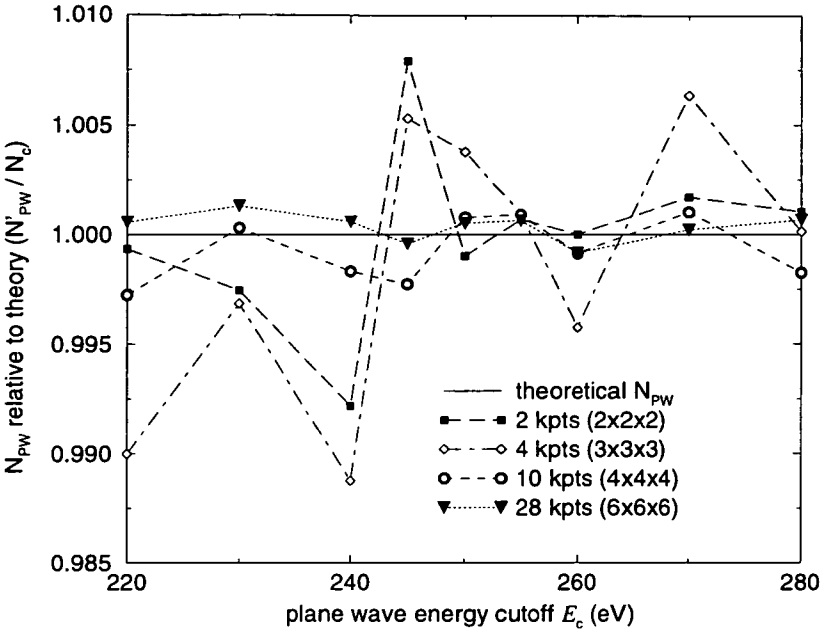


Figure 3.3. Ratio of the average number of plane waves used at each cutoff to that given by assuming that plane waves are distributed homogeneously and continuously through reciprocal space, for different densities of the **k**-point mesh. Finer **k**-point grids spread the discontinuities over different values of $\mathcal{E}_c$ and so conform more closely to the theoretical value.

mean makes an almost negligible difference. The difference between N_c and N'_{PW} for different numbers of **k**-points is illustrated in Figure 3.3: as a finer **k**-point mesh is used, the variation of the offset in reciprocal space introduced by each has the effect of dispersing the critical points at which extra plane waves are added, so the average number of plane waves conforms more closely to the homogeneous ideal.

As was noted above, only distortions which change the volume are affected by the Pulay stress (3.14), so the stress correction is isotropic. An earlier but broadly equivalent derivation [63] simply considered the derivative of energy with respect to volume. For example, in the case of an orthorhombic cell, changing the lattice parameter in one dimension (a, say) affects the number of plane-waves only via the change in volume, and so will be the same for a, b or c. Furthermore, if plane waves are distributed homogeneously, a volume-conserving shear distortion should not change the number

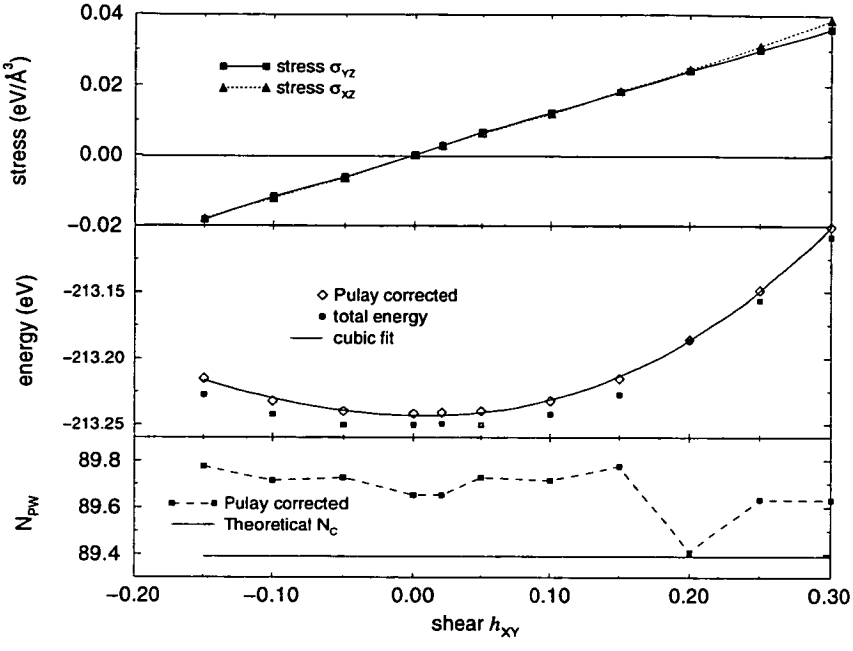


Figure 3.4. Stress, energy and average number of plane waves versus shear for a primitive unit cell of silicon (2 atoms), with (filled squares) and without (open diamonds) Pulay correction to the energy. An abnormally low cutoff (100 eV) was used with a $3 \times 3 \times 3$ **k**-point mesh.

of plane waves required by (3.11). The energy, stresses and average number of plane waves under a shear distortion of the unit cell are shown in Figure 3.4: as expected, the shear (off-diagonal) stresses do not require Pulay correction to coincide with the minimum of energy. There are still changes in the number of plane waves used, so the total energy still benefits from Pulay correction. A higher **k**-point density was found to be necessary in order for the energy to take the correct form, i.e. to have a minimum at zero shear.

3.2.3 Changing cell in response to stress

If the stresses are to be used only to find the equilibrium structure which reduces them to zero, then only the lattice parameters which produce zero stress are important, not the path taken from an initial guess. It is only with reference to the speed of

convergence that the choice of algorithm is important.

The method of Andersen [65] for a cubic system was extended by Parrinello and Rahman [66], writing a Lagrangian which assigns some kinetic energy to motion of the cell. The actual position of an ion is written in terms of the boxmatrix $h = \{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$ giving the lattice vectors, and fractional coordinates $\mathbf{s}_i$, so that $\mathbf{r}_i = h\mathbf{s}_i$. A metric tensor $g = h^T h$ may also be constructed, and the volume is, as before, $\Omega = |h| = \mathbf{a}.\mathbf{b}\times\mathbf{c}$. If the unit cell is assigned some fictitious 'boxmass' W, then the Parrinello-Rahman Lagrangian is

$$\mathcal{L} = \frac{1}{2}\sum m_i \dot{\mathbf{s}}_i^T g \dot{\mathbf{s}}_i - U(\{\mathbf{r}_i\}) + \frac{1}{2}W\mathrm{Tr}(\dot{h}^T\dot{h}) - P\Omega \tag{3.17}$$

where the first term gives the ionic kinetic energy, the second the potential energy (such as found in the DFT formalism), the third the kinetic energy of the unit cell, and the final term allows the inclusion of an external isotropic pressure. The electronic degrees of freedom are assumed to have converged before structural relaxation has taken place, as with relaxation of the ions.

From this Lagrangian, equations of motion for the lattice parameters contained in h may be derived [66]. Stresses must act over areas normal to the faces of the cell, represented by $\Omega\mathbf{B} = \{\mathbf{b}\times\mathbf{c}, \mathbf{c}\times\mathbf{a}, \mathbf{a}\times\mathbf{b}\}$. The lattice parameters must then change according to

$$\ddot{h} = \frac{1}{W}(\Pi - P)\Omega\mathbf{B} \tag{3.18}$$

using a stress Π given by

$$\Pi = \sigma + \frac{1}{\Omega}\sum_i m_i \mathbf{v}_i \mathbf{v}_i^T \tag{3.19}$$

where σ is the corrected stress given in (3.15) and the second term is the stress resulting from motion of the ions within the cell. However, if the cell and the ions are relaxed independently, assuming that $\mathbf{v}_i = 0$, and so $\Pi = \sigma$, then the lattice parameters may simply be relaxed in the direction of the stress according to (3.18). Since we

only require the equilibrium structure, where the velocities must be zero, the correct structure will still be found. If full molecular dynamics is required, for example to follow a phase transition, then the equations of motion for the ions should be integrated with a constant timestep, and the velocities calculated [9].

However, $\mathcal{L}$ is not invariant under arbitrary equivalent choices of unit cell vectors [67]. For example, a supercell made up of two adjacent cubic unit cells will feel a uniform isotropic compressive stress if the lattice parameter is too large. However, due to the use of supercell faces in (3.18), $\ddot{h}$ will be twice as large for the edges of length a then for those of $2a$; although this should also reach equilibrium (if the simulation is damped or quenched) it is inefficient. It was thus suggested by Wentzcovitch in [67] that the *strain* instead of the lattice vectors should be used as the fundamental variable, which modifies the Car-Parrinello Lagrangian to

$$\mathcal{L} = \frac{1}{2}\sum m_i \dot{\mathbf{s}}_i^T g \dot{\mathbf{s}}_i - U(\{\mathbf{r}_i\}) + \frac{1}{2}W\mathrm{Tr}(\dot{h}^T f_0 \dot{h}) - P\Omega \tag{3.20}$$

where the tensor $f_0 = \Omega^2 \mathbf{B}^T\mathbf{B}$ scales h to the strain ϵ, using the initial face vectors. The equation of motion for the lattice vectors now becomes

$$\ddot{h} = \frac{1}{W}(\Pi - P)\Omega\mathbf{B}f_0^{-1} \tag{3.21}$$

where W has been rescaled. It was demonstrated [67] that this produces better dynamics, and the advantages of using symmetrised stresses were again discussed, not least in preventing spurious rotation of the supercell.

The stresses apply to the whole system, not to individual atoms, so only the elements of the point group are required to symmetrise the stresses. The forces $\mathcal{F}$ on the faces of the unit cell are related to the stress $\boldsymbol{\sigma}$ and reciprocal lattice vectors $\mathbf{B}$ by $\mathcal{F} = \Omega\boldsymbol{\sigma}\mathbf{B}$ (Ω is the volume of the crystal and scales reciprocal lattice vectors up to the normal vectors of the faces). Both $\mathcal{F}$ and $\mathbf{B}$ must be invariant under point group

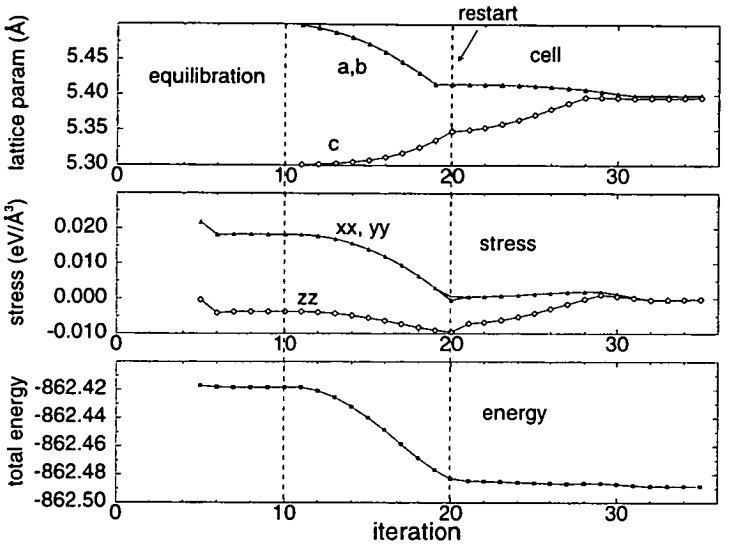


Figure 3.5. Relaxation of a tetragonal cell containing eight atoms of diamond-structured silicon, under stress with Pulay correction. The stress is equilibrated in the first 10 iterations, and the simulation is restarted after another 10 iterations, which resets the 'velocity' to zero. The equilibrium lattice parameter is 5.395 Å as in Figure 3.1, and can be seen to be reached in this relaxation.

operations, so

$$\begin{aligned} \mathbf{S}\mathcal{F} &= \mathbf{S}(\Omega\sigma\mathbf{B}) = \Omega\sigma\mathbf{S}\mathbf{B} \\ \Rightarrow \sigma &= \mathbf{S}\sigma\mathbf{S}^{-1} \end{aligned} \tag{3.22}$$

This is again implemented by averaging over all N_R rotations which form the point group of the space group:

$$\sigma = \frac{1}{N_R}\sum_{R=1}^{N_R} \mathbf{S}_R\sigma\mathbf{S}_R^T \tag{3.23}$$

Symmetrisation is performed separately for each contribution to the stress (e.g. electron-electron, electron-ion etc.). The symmetry of the stresses is vital for retaining the symmetry of the unit cell when relaxing the box; for example, in orthorhombic cells it prevents noise in the off-diagonal elements of the stress tensor from rotating or skewing the unit cell.

Full implementation of either of these Lagrangians will allow molecular dynamics, but in the present work only the configuration which produces zero stress and zero forces, thus minimising the energy, is required. The conjugate-gradients approach applied to ionic motion was discussed above, but the unit cell is relaxed to equilibrium by performing quenched molecular dynamics using the Verlet algorithm. The boxmass is chosen to produce an acceptable speed of relaxation, but does not affect the final equilibrium lattice parameters [65]. An example of the relaxation of an initially tetragonal cell back to the cubic equilibrium structure is shown in Figure 3.5.

3.2.4 Examples and practicalities

The Pulay stress correction described above gives only an approximation to the true derivative of energy with respect to strain. Equilibrium lattice parameters calculated using a Pulay correction will still, therefore, disagree slightly with those obtainable (in principle) from a series of simulations with a fully converged basis set, and at a fine mesh of lattice parameters. In many cases, the uncertainty in the estimation of the convergence gradient $\partial E/\partial \ln(\mathcal{E}_c)$ forms the predominant error in the method, and must be considered when determining a suitable cutoff energy for structual optimisation. Some of the effects of the imperfect nature of the Pulay correction can be seen in Figure 3.2.

The convergence gradient decreases rapidly while the cutoff is below its optimum value, as shown in Figure 3.6. The calculations illustrated there used a central differences method from the total energies calculated at relatively coarse intervals, and the deviation from a smooth function of $\mathcal{E}_c$ may be clearly seen. A simple linear fit to the total energy as a function of cutoff over a smaller range around the desired value will improve on this method, but may still yield an uncertainty in the convergence gradient of as much as 10%. This may be converted into a Pulay stress correction and compared with the bulk modulus to see if it will have an unacceptably high effect on the predicted lattice parameters. If so, a higher cutoff must be used, which will decrease

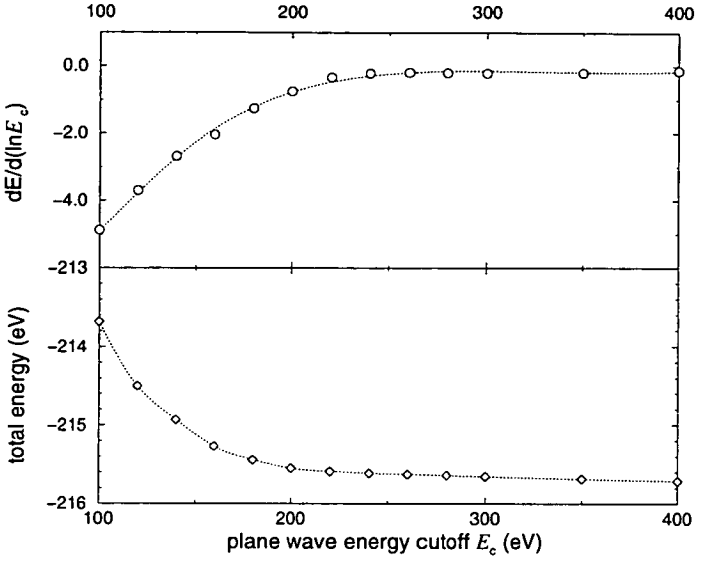


Figure 3.6. Variation of total energy and convergence gradient ($\partial E/\partial \ln \mathcal{E}_c$) with cutoff energy $\mathcal{E}_c$ for two atoms of Si; convergence gradient is calculated by central differences from total energy data of Figure 2.2; lines are guides to the eye only.

the magnitude of the convergence gradient and hence its error. Unexpectedly large errors may arise due to the common practice of using a cutoff energy which is large enough to converge energy differences between structures (see Section 2.2.5), but not large enough to converge the absolute energy from which the convergence gradient is calculated. This will lead to a large value for the convergence gradient and hence for its uncertainty.

A quadratic or higher fit to $E(\mathcal{E}_c)$ usually gives a more accurate value of the convergence gradient, but will obviously only be valid near to the exact cutoff energy for which the gradient is determined. As a structural optimisation progresses, the effective cutoff changes, which should, strictly, require the convergence gradient to change also. A polynomial fit may be used to determine the cutoff at which the discrepancy between the convergence gradient in use and that predicted for the new cutoff becomes unacceptable; the optimisation should then be restarted from the latest structural parameters found, with the original cutoff. In response to this potential source of inaccuracy, any

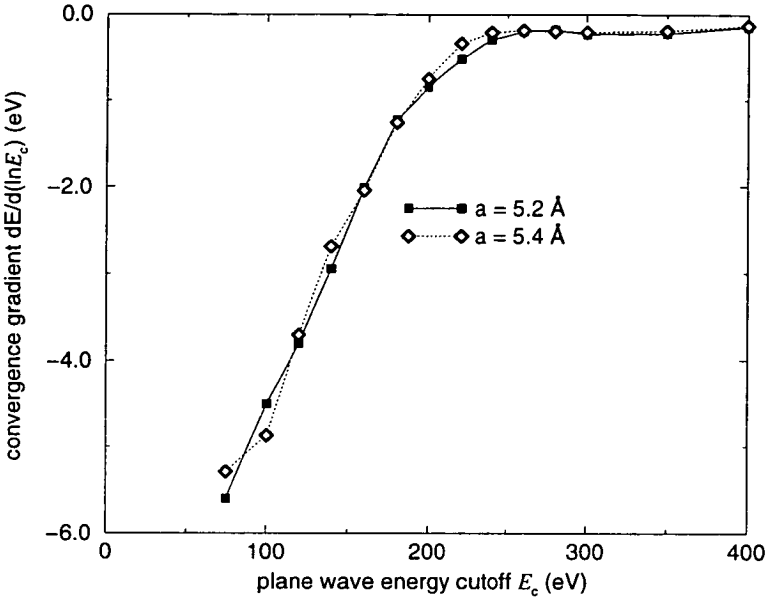


Figure 3.7. Convergence gradient ($\partial E/\partial \ln \mathcal{E}_c$) versus cutoff energy $\mathcal{E}_c$ for two atoms of Si (calculated by central difference method from data in Figure 2.2) at two different volumes.

structure determined by cell relaxation is confirmed by a fresh simulation at the original cutoff, to ensure that the stresses are sufficiently small; sometimes the optimisation must be performed in two or more steps to ensure the cutoff does not stray too far from that originally prescribed.

The convergence gradient measures the change in energy due to better description of the charge in the plane waves with highest kinetic energy, i.e. the rapidly fluctuating wavefunctions in the deep pseudopotential near to the ionic cores. It is thus completely dependent on the pseudopotential. However, it ought to be independent of cell volume or lattice parameter, and this has been found for all systems thus far investigated (as illustrated in Figure 3.7). As expected, it has also been found to be independent of k-point sampling within the limits of uncertainty, as shown in Figure 3.8, but finer **k**-point sampling reduces the error from the fit, since the error in the homogeneous plane wave distribution assumption is smaller in this case (Figure 3.3). The convergence gradient is not, however, completely independent of the environment of each ion, since the charge

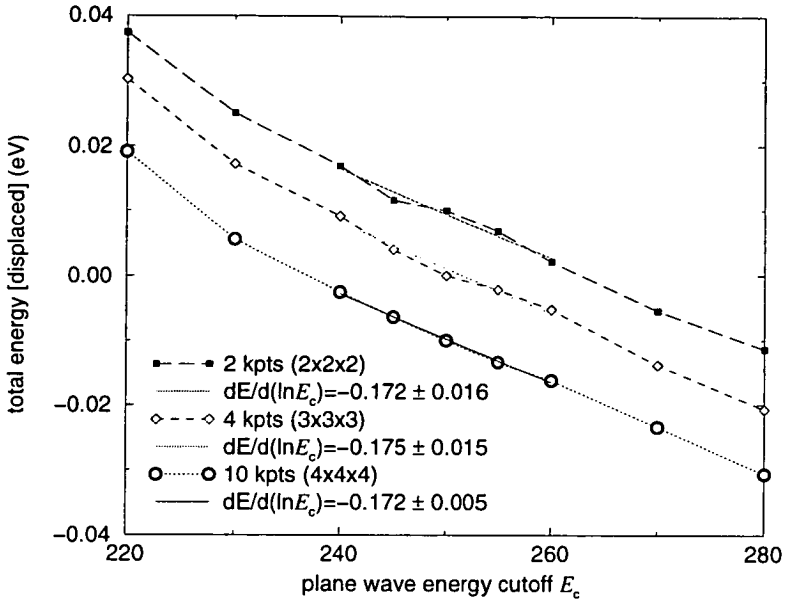


Figure 3.8. Determination of the convergence gradient $(\partial E/\partial \ln(\mathcal{E}_c)\ \mathcal{E}_c$ by linear fitting, for two atoms of Si around 250 eV, at different densities of **k**-point sampling; total energy values have been rigidly shifted to allow easy comparison.

distribution surrounding an ion is affected by the type of bonding involved; even with norm conservation, no pseudopotential is fully transferable. The convergence gradient is, of course, specific to the pseudopotentials in use, and is often almost entirely due to the pseudopotential with the highest minimum cutoff energy.

Efficient structural relaxation requires that the unit cell and the ionic coordinates relax at roughly the same rate [68]. However, if the ionic relaxation proceeds by the conjugate gradients method, the speed of convergence can be hard to predict, and great variation in the speed of ionic relaxation has been found, especially when performed simultaneously with cell relaxation. The conjugate gradients method is designed for linear systems; in non-linear problems it may only find the minimum very slowly. Ideally, the normal modes of the system should be used as the variables for ionic minimisation, since they will be closest to giving harmonic behaviour, but usually only the fractional ionic positions $\{\mathbf{s}_I\}$ and cell shape h are available as variables. When the cell and the ions are being relaxed simultaneously, instead of dealing with $U(\{\mathbf{R}_I\})$,

the energy surface is in fact $U(\{\mathbf{s}_I\}, h)$, further separating variables which are probably strongly coupled. It was thus sometimes found more efficient to relax the two sets of coordinates separately until the forces and stresses were less than 0.1 eV/Å and 0.1 eV/Å^3 respectively, after which it becomes less problematic to relax all parameters simultaneously.

Many of the structural relaxations performed were followed by phonon calculations, described in the following Section, which require accurate calculations of forces for distortions from equilibrium. It will be shown to be important that any 'residual forces' on the relaxed structure are as small as possible. For this reason, once the stresses were less than, typically, 10^{-4}eV/Å^3, the relaxation of the cell was terminated, and optimisation continued with only relaxation of the ions, until the forces were smaller than 10^{-4}eV/Å. This usually did not significantly affect the stresses, but if they were increased by the subsequent ionic relaxation, to beyond the threshold given above, then further simultaneous relaxation was performed.

3.3 Conclusions

We have described both established methods and new developments for finding the equilibrium structural parameters of a crystal phase using the forces and stresses calculated from first principles. Although the use of plane waves in pseudopotential codes such as CASTEP facilitates force calculation, there is a correction to the stress which is required at all but the very highest levels of basis set completeness. The necessity for this correction to be determined by a fit to calculated energies produces an uncertainty in its value which may affect the choice of cutoff and k-point sampling scheme.

The structural parameters may be optimised simultaneously, although due to the varying efficiency of computational schemes to minimise non-linear functions, separating the ionic and cell degrees of freedom may sometimes still be necessary. The degree of convergence towards the structural minimum which is required will vary according

to any further calculations which are desired, such as the determination of the normal modes which will be described in the next Chapter.

Chapter 4

Vibrational properties from first principles

> *When certain approximations are made, the dynamics of a three-dimensional crystal is not a particularly difficult subject, but it is a very algebraic one with a profusion of indices and suffices and it is easy to lose one's way among one-dimension crystal or linear chain.*
>
> W. Cochran, *The Dynamics of Atoms in Crystals* [7]

The vibrational properties of a crystal determine a wide range of macroscopic behaviour: thermal properties such as the specific heat, transport properties like resistivity and velocity of sound, and the interaction with radiation, for example in infra-red absorption and Raman scattering. The optical effects are important tools for following the progress of phase transitions. Very low frequency modes are implicated in structural phase transitions, and the presence of unstable modes in a proposed structure can suggest alternative, more stable structures. *Ab initio* calculations of phonon frequencies and eigenvectors are thus very useful not only in predicting macroscopic properties, but also in tracing mechanisms of phase transitions. Phonon eigenvectors are not experimentally accessible, beyond determination of the symmetry and polarisation, so theoretical work can complement experimental investigation [69].

Once the equilibrium structure has been found, as described in the previous Chapter, the vibrational properties of the crystal may be determined by finding the variation of energy with ionic displacement. The large number of ionic coordinates in a typical unit cell would require many individual displacements to be made. However, the phonon modes may be found much more efficiently if the forces are known [42], as described below. Such calculations have been performed with plane-wave pseudopotential methods for over fifteen years [58] but are now becoming routine even for complex structures; in this Chapter the general theory and its implementation with the CASTEP code are described.

4.1 Basic theory of lattice dynamics

We first review the basic theory of lattice dynamics, following the notation of Maradudin *et al.* [24]. An infinite crystal is divided up into primitive cells (labelled by l), each containing r atoms (labelled by κ) having mass m_κ. As the lattice vibrates, each atom is displaced from equilibrium by an amount $\mathbf{u}\binom{l}{\kappa}$. The position of each atom is thus given by

$$\mathbf{R}\binom{l}{\kappa} = \mathbf{x}_l + \mathbf{x}_\kappa + \mathbf{u}\binom{l}{\kappa} \tag{4.1}$$

For small displacements, the harmonic approximation may be applied, defining *harmonic force constants* Φ:

$$F_\alpha\binom{l}{\kappa} = \sum_{l'}\sum_{\kappa'}\sum_{\beta} \Phi_{\alpha\beta}\binom{l\ \ l'}{\kappa\ \ \kappa'} u_\beta\binom{l'}{\kappa'} \tag{4.2}$$

or, alternatively,

$$\Phi_{\alpha\beta}\binom{l\ \ l'}{\kappa\ \ \kappa'} = \frac{-\partial^2 E}{\partial u_\alpha\binom{l}{\kappa}\, \partial u_\beta\binom{l'}{\kappa'}} \tag{4.3}$$

where E is the total energy, and $\mathbf{F}$ is the force exerted on the ions when atoms are displaced from equilibrium by $\mathbf{u}$; α and β run over x, y, z.

A normal mode with angular frequency ω is defined as a motion such that

$$\mathbf{F}\binom{l}{\kappa} = -m_\kappa \omega^2 \mathbf{u}\binom{l}{\kappa} \tag{4.4}$$

By defining *mass reduced coordinates* $\epsilon\binom{l}{\kappa} = \sqrt{m_\kappa}\mathbf{u}\binom{l}{\kappa}$, and using (4.2), the equations of motion implied in (4.4) may be written as an eigenproblem:

$$-\omega^2 \epsilon\binom{l}{\kappa} = \sum_{l}{}' \sum_{\kappa'} \frac{\Phi\binom{l\ l'}{\kappa\ \kappa'}}{\sqrt{m_\kappa m_{\kappa'}}} \epsilon\binom{l'}{\kappa'} . \tag{4.5}$$

We thus seek polarisation vectors (eigenvectors) $\epsilon\binom{l'}{\kappa'}$ which describe the relative amplitude of displacement for all atoms under a vibration with frequency ω.

As in the bandstructure formalism, the periodicity of a perfect crystal may be exploited. We expect phonon solutions, in which two equivalent atoms, having the same κ but in different primitive cells, will differ in the phase of their motion by an amount $\mathbf{q}\cdot\Delta\mathbf{x}$ where $\Delta\mathbf{x}$ is the difference in the positions of their primitive cells. The solutions to (4.5) are thus classified by their periodicity over primitive cells. The eigenvectors are therefore expressed:

$$\epsilon\binom{l}{\kappa} = \epsilon\binom{0}{\kappa} \exp(i\mathbf{q}\cdot[\mathbf{x}_l - \mathbf{x}_0]) \tag{4.6}$$

where $\mathbf{x}_0$ is the position of the origin of some arbitrary reference cell and $\mathbf{q}$ is the wavevector. If we define a factor $\mu_{\kappa\kappa'} = 1/\sqrt{m_\kappa m_{\kappa'}}$, and the Fourier-transformed dynamical matrix by

$$D_{\alpha\beta}(\kappa\kappa'|\mathbf{q}) = \sum_{l} \mu_{\kappa\kappa'} \Phi_{\alpha\beta}\binom{0\ l}{\kappa\ \kappa'} \exp(i\mathbf{q}.[\mathbf{x}_l - \mathbf{x}_0]) \tag{4.7}$$

then the original eigenproblem of (4.5) may be rewritten in terms of eigenvectors $\boldsymbol{\xi}_\kappa = \epsilon\binom{0}{\kappa}$:

$$-\omega^2 \boldsymbol{\xi} = \mathbf{D}(\mathbf{q})\boldsymbol{\xi} \tag{4.8}$$

For each wavevector $\mathbf{q}$, evaluation and diagonalisation of $\mathbf{D}(\mathbf{q})$ will give the frequencies of normal modes at $\mathbf{q}$, which may be assigned to the corresponding point of the Brillouin Zone. The eigenvectors give the polarisation vectors of each mode, and now only need to be specified for each κ, that is, for each ion in one primitive unit cell.

4.2 From *ab initio* forces to phonons

The theory of vibrations in the crystal lattice is written as if the interactions between all the atoms, separated by distances up to infinity, were known. In practice this is not possible. It is shown here that by using the supercell method with periodic boundary conditions, phonons at some specific points of the Brillouin Zone may be calculated from first principles. Early examples of similar methods concentrated on simple structures with high-symmetry phonons [70, 71, 72].

Pseudopotential plane-wave methods usually simulate only one or a few primitive cells, with periodic boundary conditions, but by means of bandstructure techniques, and the Ewald summation, Coulomb interactions between all ions and all electrons of an infinite crystal are included, as described in Chapter 2. If the supercell extends over l_1 primitive cells in some direction, cyclic (periodic) boundary conditions imply that $\mathbf{u}\binom{l}{\kappa} = \mathbf{u}\binom{l+ml_1}{\kappa}$, where m is any integer. The sum over all primitive cells in (4.2) may then be broken up so that the forces are given by

$$\mathbf{F}\binom{l}{\kappa} = \sum_{l'=1}^{l_1} \sum_{m=-\infty}^{\infty} \sum_{\kappa'} \Phi\binom{l\ \ l'+ml_1}{\kappa\ \ \kappa'} \mathbf{u}\binom{l'+ml_1}{\kappa} \tag{4.9}$$

If a wavevector $\mathbf{q}$ implies the same periodicity as these cyclic boundary condition impose, corresponding phonons may be 'frozen into' the supercell without violating boundary conditions. This is the basis of 'frozen-phonon' calculations, which formed some of the earliest applications of *ab initio* methods to lattice dynamics [42, 58, 73]. These aim to find the energy for excitation of one mode, but the eigenvector must thus be known beforehand (although simple mixed eigenvectors have also been found in this

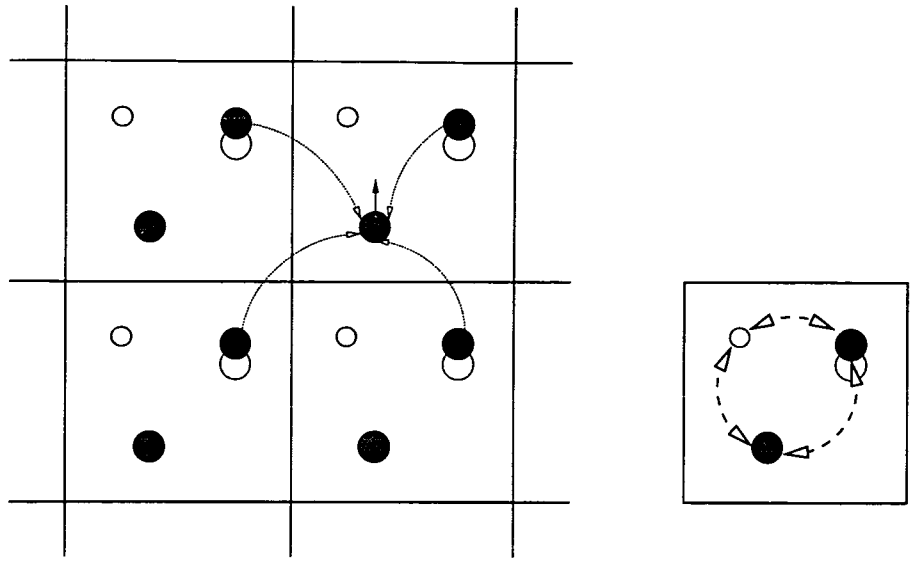


Figure 4.1. The periodic array of supercells may be considered as a single isolated supercell, with interactions only between the enclosed ions. This is valid only for wavevectors which are reciprocal lattice vectors of the supercell.

way [43]). However, it suggests that exact calculation of phonons commensurate with a given supercell is possible. This condition on the wavevector can be written:

$$\begin{aligned} \epsilon\binom{l+ml_1}{\kappa} = \epsilon\binom{l}{\kappa}\exp(i\mathbf{q}\cdot[\mathbf{x}_{l+ml_1}-\mathbf{x}_l]) &= \epsilon\binom{l}{\kappa} \\ \Rightarrow \exp(i\mathbf{q}\cdot\mathbf{x}_{ml_1}) &= 1 \end{aligned} \tag{4.10}$$

The Fourier transformed dynamical matrix defined in (4.7) may now be written

$$\begin{aligned} D(\kappa\kappa'|\mathbf{q}) &= \sum_{l=1}^{l_1}\sum_{m=-\infty}^{\infty}\mu_{\kappa\kappa'}\Phi\binom{0\ \ l+ml_1}{\kappa\ \ \kappa'}\exp(i\mathbf{q}\cdot[(\mathbf{x}_l-\mathbf{x}_0)+\mathbf{x}_{ml_1}]) \\ &= \sum_{l=1}^{l_1}\left\{\sum_{m=-\infty}^{\infty}\mu_{\kappa\kappa'}\Phi\binom{0\ \ l+ml_1}{\kappa\ \ \kappa'}\right\}\exp(i\mathbf{q}\cdot[\mathbf{x}_l-\mathbf{x}_0]) \end{aligned} \tag{4.11}$$

Force constants Υ which relate atoms only within the supercell only may be defined

by writing

$$\Upsilon\begin{pmatrix} l & l' \\ \kappa & \kappa' \end{pmatrix} = \sum_{m=-\infty}^{\infty} \Phi\begin{pmatrix} l & l'+ml_1 \\ \kappa & \kappa' \end{pmatrix}. \tag{4.12}$$

It may be seen that (4.11) represents a system consisting only of the atoms within one supercell, with effective force constants given by (4.12). These force constants, determined by the periodicity of the supercell, are exactly the combination of original force constants Φ involved in a distortion where one atom in every supercell is displaced, i.e. a simulation with periodic boundary conditions. The exact phonon frequencies may thus be obtained by treating the supercell as an isolated system of r atoms, connected by springs with effective force constants which may be found from the simulation, instead of as an infinite crystal, as illustrated in Figure 4.1.

The condition (4.10) determines which phonons may be calculated exactly in this way. It may be generalised by the cyclic boundary condition $\mathbf{q} \cdot \mathbf{A} = 2\pi m$ (m integer) for all lattice vectors $\mathbf{A}$ of the supercell. In other words, the wavevector must be one of the reciprocal vectors of the supercell. At each wavevector, there are $3r$ modes, where r is the number of ions in the primitive cell; $\mathbf{D}(\mathbf{q})$ thus has dimensions $3r \times 3r$.

An important case is that of a supercell consisting of only one primitive cell. The reciprocal vectors of the supercell are hence those of the lattice itself, and are therefore all Γ points of the Brillouin Zone. Thus only Γ point modes are obtainable from a simulation of only a single primitive cell, using this method. In an isolated system, such as a molecule, there is no interaction between different supercells, and the 'phonons at the Γ point' are just the molecular vibrational modes [28].

If the supercell is sufficiently large compared to the range of inter-ionic forces then the interaction between equivalent ions in neighbouring supercells is very small. For example, a supercell containing 64 atoms of silicon in the diamond (FCC) structure, with dimensions $2a = 10.8$ Å, was used to calculate vibrational modes by displacing one atom by 0.1% of the unit cell (0.0108Å), as will be described below. The decay of force constants with distance from the displaced atom is shown in Figure 4.2.

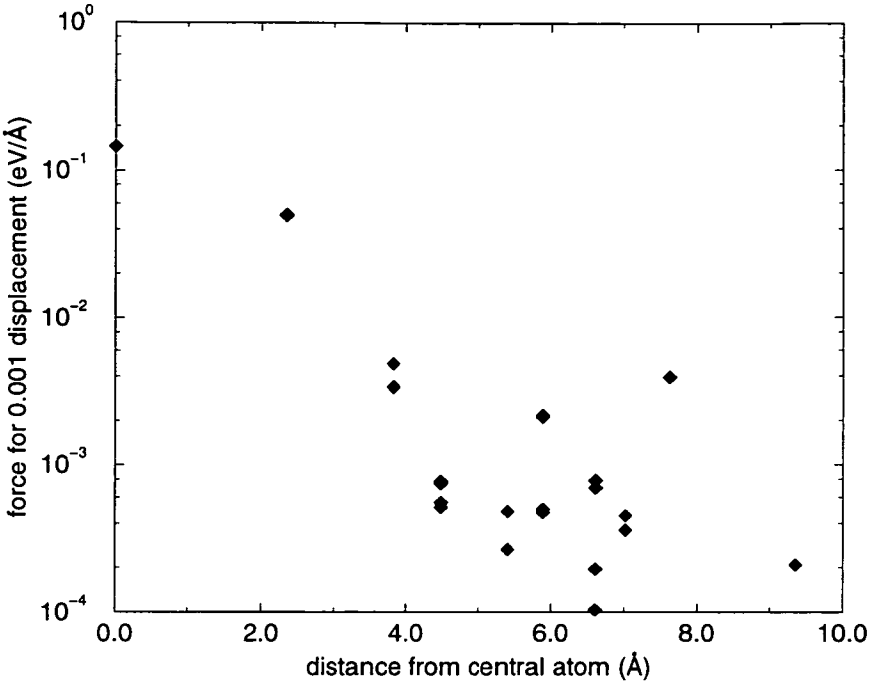


Figure 4.2. Magnitude of the force exerted on each atom in a 64-atom cell of diamond Si when one atom is displaced by 0.1% of the cell, versus the distance from the displaced atom. The force at zero distance corresponds to the self-interaction restoring force. The forces decrease sharply beyond the second nearest neighbours (at 3.8Å), but are still non-negligible as far away as 7.6Å. Some forces are zero by symmetry, and are not shown. The angular dependency of interaction is not shown.

Although the symmetry of the distorted structure was enforced during the simulation, the small amount of anharmonicity arising from a finite displacement can be seen to split degeneracy, e.g. in each of the third neighbour sets of atoms.

This leaves the problem of finding the force constants Υ from first principles. From (4.2), but considering that the sum over supercells is implicit in the calculation of the force,

$$\Upsilon_{\alpha\beta}\begin{pmatrix} l & l' \\ \kappa & \kappa' \end{pmatrix} = \frac{\partial F_\alpha \begin{pmatrix} l \\ \kappa \end{pmatrix}}{\partial u_\beta \begin{pmatrix} l' \\ \kappa' \end{pmatrix}} \tag{4.13}$$

If only the $\begin{pmatrix} l' \\ \kappa' \end{pmatrix}$th atom is displaced along one coordinate, β and by an amount η then

$$F_\alpha \begin{pmatrix} l \\ \kappa \end{pmatrix} = \eta \Upsilon_{\alpha\beta}\begin{pmatrix} l & l' \\ \kappa & \kappa' \end{pmatrix} \tag{4.14}$$

thus determining a series of values of Υ. All that is required to calculate a set of phonons is thus to move each ion in the supercell along each coordinate in turn, and calculate the Hellmann-Feynman forces on all the ions.

As it stands, $3r$ such displacements will need to be made, if Υ is to be completely determined and thus diagonalised. However, once again the symmetry elements of the crystal may be employed. If an element $\{\mathbf{S}|\mathbf{v}(S)\}$ maps atom j to J and atom κ to K then it can be shown [24] that pairs of force constants are related by

$$\mathbf{\Phi}_{JK} = \mathbf{S}\mathbf{\Phi}_{j\kappa}\mathbf{S}^T \tag{4.15}$$

Hence if all the symmetry elements of the point group are known, all the force constants can be deduced from a smaller number of simulations. In a structure with high symmetry, only a few such simulations are needed to construct the entire dynamical matrix, thus greatly reducing the computing time required.

Specific phonon frequencies were first extracted from first-principles pseudopotential calculations, by Yin and Cohen [42], who calculated four phonon frequencies for silicon and germanium. They used the 'frozen phonon' method, possible for high symmetry phonons in a simple lattice; the forces method of (4.14) and the energy approach of (4.3) were compared and found to be in good agreement, as would be expected. Anharmonicity was also investigated, and may in general be calculated by freezing in various amplitudes of a known phonon eigenvector. This is done for particular unstable modes in Chapters 6.

4.3 Code and practicalities

As part of the work described in this thesis, code was written to calculate phonon frequencies and eigenvectors at a given wavevector. using the Hellmann-Feynman forces from a set of ionic displacements. Each displacement must be of one atom only, and the whole set along either the Cartesian axes or parallel to the edges of the unit cell.

Displacements are assumed to be of a sufficiently small magnitude to be treated in the harmonic approximation.

The force constants are found by dividing the forces by the size of the displacement which caused them, according to (4.14), and are assigned to the relevant $\Upsilon_{\alpha\beta}\left(\begin{smallmatrix} l & l' \\ \kappa & \kappa' \end{smallmatrix}\right)$. The symmetry relationships given by (4.15) are then used to copy these force constants to the related elements of $\mathbf{\Upsilon}$. If this process assigns all elements of the dynamical matrix $\mathbf{\Upsilon}$, enough displacements have been made. Running the program with arbitrary forces can thus quickly confirm or determine by trial and error a minimum set of displacements which need to be made. Once it is filled with *ab initio* force constants, the dynamical matrix is then scaled by the appropriate masses according to (4.5).

The symmetry elements and associated mapping list F_0 needed to complete the dynamical matrix were found by analysis of the supercell, using an existing program which determines special **k**-points from the symmetry. Occasionally a system has higher symmetry than that of the supercell used for simulation (e.g. a hexagonal system or an isolated molecule in an orthorhombic box) so some symmetry elements are not picked up from analysis of the supercell. Similarly, a supercell containing more than one primitive cell will have translations associated with the identity matrix which are not needed to specify the structure and are thus not in the generated list of symmetry elements. Such additional elements must be added to the list of symmetry operations which may be used in (4.15).

Diagonalisation of the dynamical matrix then yields the phonon frequencies and mass-reduced eigenvectors corresponding to all the normal modes of the supercell. If required, a wavevector $\mathbf{q}$ may be chosen, the Fourier transformed dynamical matrix $\mathbf{D}(\mathbf{q})$ formed, and only those phonons at $\mathbf{q}$ calculated. This splits the vibrational modes in a large supercell over different points of the Brillouin Zone, and gives eigenvectors with fewer components, which describe the displacements in only one primitive cell.

To obtain phonons which obey the space group symmetry, including the correct degeneracies, the original forces must display the correct symmetries. If they do not,

the dynamical matrix will not have perfect symmetry, and the phonons will break symmetry in both frequencies and eigenvectors. In plane-wave codes such as CASTEP, accurate forces are only obtained with a symmetric k-point set, and after converging the wavefunctions extremely well: the degree of convergence (number of self-consistent iterations) required for accurate forces is much greater than that for accurate energies (see Section 3.1.1). It was found that symmetrisation of the forces within the CASTEP simulation was also advisable: even fully converged forces do not always display exactly the correct symmetry. When calculating forces arising from displacements, the lower symmetry of the distorted structure should be used.

It is also important that displacements are made within the harmonic limit. This can be determined by making smaller and smaller displacements until the force is proportional to the displacement. If the ion is not on a plane of reflection symmetry, displacements in both positive and negative directions should be made, and the force constants averaged: this will overcome any first-order anharmonicity. Displacements of order 0.1% of the lattice parameter were generally found to be sufficiently small.

Although in principle displacements should be made from equilibrium, in practice it can take a very large number of iterations to relax the ions so that the forces are negligible. If there are small 'residual forces' on a structure at equilibrium, these should be subtracted from the forces after displacement: this procedure is only valid within the harmonic limit. The process of averaging over positive and negative displacements removes this requirement. The size of these residual forces, which can be considered as a source of noise since the system is never truly harmonic, determines the smallest displacement which may be usefully made; if the forces from displacement are not at least an order of magnitude larger than the residual forces, large errors are likely to occur. It was found that residual forces less than 10^{-3} eV/Å were usually adequate for accurate force calculations. Symmetrisation of the forces in the equilibrium structure was found to be helpful here, since those forces which should be zero by symmetry will be enforced to be so. When calculating equilibrium residual forces to be subtracted from

displacement forces, an equivalent simulation should be used, not just the end results of the structural optimisation. In CASTEP this implies identical plane-wave cutoff and k-point sampling; the effective cutoff may change during a cell relaxation.

4.3.1 Newton's Third Law

Newton's Third Law of action and reaction demands that when one atom is internally displaced there is no net force on the system. This is equivalent to the requirement for phonon vibrations to take place without oscillation of the centre of mass since this would require an external driving force, or that the energy is invariant with respect to a equal displacement of all atoms, that is, a rigid shift of the entire system. Within the formalism of harmonic lattice dynamics defined by (4.2), it is thus expected that

$$\begin{aligned} \sum_{l\kappa} F_\alpha \begin{pmatrix} l \\ \kappa \end{pmatrix} &= \sum_{l\kappa} \sum_{l'\kappa'\beta} \Phi_{\alpha\beta} \begin{pmatrix} l & l' \\ \kappa & \kappa' \end{pmatrix} u_\beta \begin{pmatrix} l' \\ \kappa' \end{pmatrix} = 0 \\ \Rightarrow \sum_{l\kappa} \Phi_{\alpha\beta} \begin{pmatrix} l & l' \\ \kappa & \kappa' \end{pmatrix} &= 0 \text{ for all } \alpha, \beta, l', \kappa'. \end{aligned} \tag{4.16}$$

and this restriction is known as the 'sum rule' for force constants. If (4.16) holds, then three translational modes with exactly zero frequency must be normal modes of the system, with eigenvectors

$$\mathbf{u} \begin{pmatrix} l \\ \kappa \end{pmatrix} = \mathbf{u}. \tag{4.17}$$

In practice, the forces obtained from CASTEP calculations only approximately obey (4.16), for a number of reasons. Firstly, if the wavefunctions are not completely converged from an initial random starting point, the Hellmann-Feynman theorem is not strictly applicable and will give errors in the forces. Secondly, even for perfectly converged calculations, the charge density is interpolated between a discrete set of grid points used in the Fourier transform between real and reciprocal space. If the crystal is displaced slightly, ions will lie at different positions with respect to this grid, and the interpolation of charge density will produce slightly different results; the energy is thus

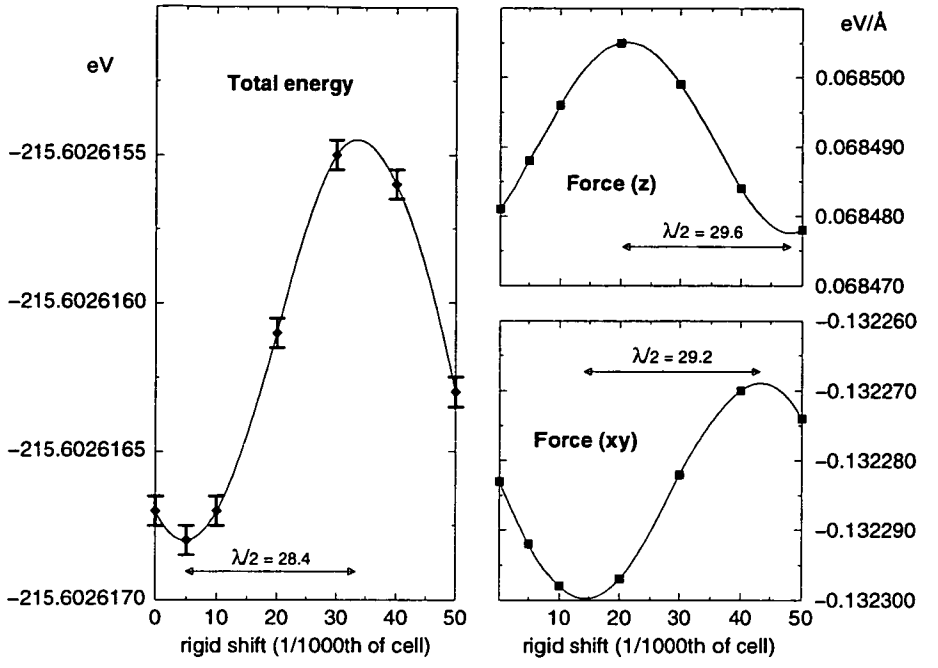


Figure 4.3. Total energy and force for a primitive unit cell (two atoms) of diamond-structured Si with a Fourier transform grid of 18 × 18 × 18 (energy cutoff of 250 eV); the force shown is that on an atom displaced from equilibrium by 0.005 of the unit cell dimension a. The calculated values vary periodically with an average $\lambda = 0.0582a$, compared to $a/18 = 0.055a$ or $a/17 = 0.059a$. This variation is thus assumed to be due to the discretisation of the charge density. The amplitude of the oscillation is very small: approximately 7×10^{-7} eV in the total energy and 15×10^{-6} eV/Å in the forces.

only very nearly conserved. The force will also vary very slightly with the position of the origin with respect to the FFT grid. Figure 4.3 illustrates the variation in energy and force as the entire crystal is displaced, in a system with one atom displaced from equilibrium (as it would be to measure force constants); it can be seen that this error is very small.

The amplitude of fluctuations in the force is comparable to the amount of violation of the sum rule typically observed in sets of forces from CASTEP simulations. The violation of the sum rule in the system shown in Figure 4.3 was also found to fluctuate on the same scale as the forces, although in this case the sum rule could be automatically satisfied by enforcing symmetry in the calculation. When a FFT grid coarser than

required by the energy cutoff was used, the amplitude of this oscillation increased considerably; conversely, when it was increased beyond that required, oscillation became undetectable. In the general case, any violation of the sum rule is therefore closely related to the charge density discretisation and interpolation. The largest contributions to the energy fluctuations came from the kinetic and pseudopotential energies.

There are, of course, other errors in the forces due to the method. Anharmonicity affects the forces both directly, since a finite displacement must be made, and indirectly if any small forces remaining in the relaxed structure are subtracted, or in the process of averaging over a series of displacements. The pseudopotential approximation, the local density approximation and k-point sampling of course all influence the forces in some way, but none of these should directly violate Newton's Laws.

It is of course possible to enforce the sum rule on a set of forces obtained from a simulation simply by setting the on-diagonal 'self-interaction' term such that the overall sum is zero. However, for $\alpha \neq \beta$ this may result in a violation of the Hermitian symmetry of $\mathbf{\Phi}$. This can be shown by considering setting one specific term in this way:

$$\Phi_{\alpha\beta}\begin{pmatrix}1 & 1\\ 1 & 1\end{pmatrix} = -\sum_{l\neq 1,\,\kappa\neq 1} \Phi_{\alpha\beta}\begin{pmatrix}l & 1\\ \kappa & 1\end{pmatrix} \tag{4.18}$$

The Hermitian symmetry of the matrix demands that $\Phi_{\alpha\beta}\left(\begin{smallmatrix}1 & 1\\ 1 & 1\end{smallmatrix}\right) = \Phi_{\beta\alpha}\left(\begin{smallmatrix}1 & 1\\ 1 & 1\end{smallmatrix}\right)$ but the latter term should itself obey a sum rule analogous to (4.18), so it is also required that

$$\sum_{l\neq 1,\,\kappa\neq 1} \Phi_{\alpha\beta}\begin{pmatrix}l & 1\\ \kappa & 1\end{pmatrix} = \sum_{l\neq 1,\,\kappa\neq 1} \Phi_{\beta\alpha}\begin{pmatrix}l & 1\\ \kappa & 1\end{pmatrix} \tag{4.19}$$

Corresponding pairs of terms in these sums are *not* equal in the general case, although for structures of high symmetry, the two sums may in fact be over the same values in a different order. In general, however, there is no unique way of satisfying (4.19) and any method will necessarily interfere with other related to other force constants.

In practice the discrepancy from the sum rule should be small, as should any degree of violation of Hermitian symmetry: the violation can be minimised by ensuring that

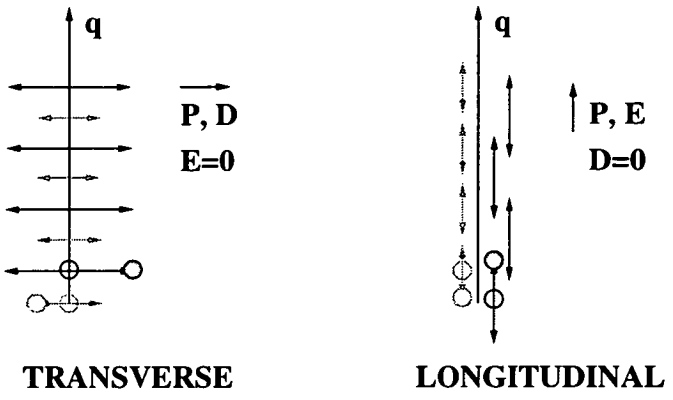


Figure 4.4. Electromagnetic fields in transverse and longitudinal optic phonons for long wavelength **q** in a polar crystal, in which the phonon displacements generate a dipole. The conditions give in (4.22) determine that the electric field **E** is non-zero for an LO phonon but zero for a TO phonon; this difference gives rise to a frequency splitting.

simulations proceed with symmetry enforced whenever possible, and that calculations are fully converged with respect to energy cutoff, number of **k**-points etc. Under these circumstances, the exact methods used to enforce the sum rule and Hermitian symmetry should have negligible effect on the forces.

In the particularly high-symmetry system of diamond-structured Si, a force constant method was used by Wei and Chou [70] with a least squares fit to solve for the optimum elements of a fully symmetrical dynamical matrix from the forces obtained via simulation.

4.3.2 LO–TO splitting at the zone centre

It was stated in Section 2.4 that longitudinal optic phonons may generate a macroscopic electric field in polar crystals, and that since this electric field may not be represented using periodic boundary conditions, phonon calculations would not find the correct frequency. This effect is explained here, (following the treatment of [48]) together with some of the approaches which may be taken to overcome this limitation.

In a long-wavelength optic phonon, ions of different charge may move in opposite

directions, generating a dipole and giving rise to a macroscopic polarisation $\mathbf{P}$. This is related to the electric field $\mathbf{E}$ and electric displacement $\mathbf{D}$ by the dielectric constant ϵ:

$$\mathbf{D} = \epsilon\mathbf{E} = \mathbf{E} + 4\pi\mathbf{P}. \tag{4.20}$$

In addition, Maxwell's equations in the absence of free charge (i.e. in insulators) dictate that both

$$\nabla \cdot \mathbf{D} = 0 \tag{4.21a}$$

$$\text{and } \nabla \times \mathbf{E} = 0. \tag{4.21b}$$

In the general case, the dielectric constant ϵ may be a tensor, but in the simple case of a cubic crystal, both $\mathbf{E}$ and $\mathbf{D}$ must be parallel to $\mathbf{P}$, and hence to the ionic displacements. For phonon solutions, the requirements of (4.21) then become

$$\mathbf{D} = \mathbf{0} \quad \text{or} \quad \mathbf{D}, \mathbf{E} \text{ and } \mathbf{P} \perp \mathbf{q} \tag{4.22a}$$

$$\text{and } \mathbf{E} = \mathbf{0} \quad \text{or} \quad \mathbf{E}, \mathbf{D} \text{ and } \mathbf{P} \parallel \mathbf{q}. \tag{4.22b}$$

In a transverse optic phonon, $\mathbf{P} \perp \mathbf{q}$, so from (4.22) it must be concluded that $\mathbf{E} = \mathbf{0}$ and hence from (4.20) that $\epsilon = \infty$. Since such phonons generate zero macroscopic electric field, the restriction imposed by periodic boundary conditions has no effect. However, in a longitudinal optic phonon, $\mathbf{P} \parallel \mathbf{q}$, so (4.22) implies that $\mathbf{D} = \mathbf{0}$, there is thus an electric field of strength $\mathbf{E} = -4\pi\mathbf{P}$, and so $\epsilon = 0$. There is thus an additional restoring force on the ionic displacements in the phonon, so the frequency will be higher than that of the transverse modes. This electric field is suppressed by periodic boundary conditions, so that the frequency of this mode reverts to that of the transverse modes.

In a cubic crystal, this effect will be manifested by optic phonons at Γ being calculated to be triply degenerate if periodic boundary conditions are imposed, as might

naïvely be expected, with the frequency of the TO phonon. This has been acknowledged in previous simple zone-centre phonon calculations [74, 75]. The polarisation and resulting electric field would be largest in very ionic crystals, such as the alkali halides, and of course will be zero in purely covalently bonded crystals where the effective charges and dipoles are zero.

At wavevectors other than $\mathbf{q} = 0$, an electric field with the periodicity of the phonon is still generated, but the supercell used to calculate such phonons must, by (4.10), be large enough to contain an integer number of wavelengths, so that there is no net dipole over the supercell. Any LO-TO splitting away from $\mathbf{q} = 0$ is thus correctly calculated by the technique described in this thesis.

In a cubic crystal with only two ions in the primitive unit cell, the Lyddane-Sachs-Teller relation [48] may be used to relate the LO and TO frequencies in a simple cubic crystal, via the static dielectric constant ϵ_0 and the index of refraction ϵ_∞:

$$\frac{{\omega_{LO}}^2}{{\omega_{TO}}^2} = \frac{\epsilon_0}{\epsilon_\infty} \tag{2.28}$$

The dielectric constants ϵ_0 and ϵ_∞ are, however, non-trivial to calculate from density-functional simulations [76], although experimental values may be used with ω_{TO} to estimate ω_{LO}. In crystals with more ions in the unit cell, the Born effective charges are needed [77], which also require techniques not yet implemented in the codes used in this work.

It is possible to calculate the LO-TO splitting by including dipole-dipole interactions, within the formalism of density-functional perturbation theory (DFPT) [78, 79] but this was outwith the scope of this thesis. It is hoped to include this functionality in CASTEP in the future.

4.4 Imaginary modes

The harmonic approximation for vibrational motion assumes that a single variable x moves under a harmonic potential $E_{\text{vib}} = kx^2/2$ where k is the 'spring constant'. If $E_{\text{vib}} \geq 0$ (i.e. $k \geq 0$) then the vibration is stable, with a frequency ω such that $E_{\text{vib}} = m\omega^2 x/2$ and $F = -m\omega^2 x$. In the phonon picture, x represents the instantaneous ionic displacement along the phonon eigenvector at each ion and ω is the frequency of oscillation.

However, if instead $E_{\text{vib}} \leq 0$, then any displacement away from $x = 0$ is unstable and results in increasing displacement, not in stable oscillation. Following the formalism for stable oscillation shows that this is represented by $\omega^2 < 0$ and thus an imaginary frequency.

Any structure which has such an unstable phonon thus has a displacement pattern which is inherently unstable (some authors use the term 'dynamically unstable') and thus a mechanism for lowering its energy. If there is any thermal motion of the ions, this phonon must increase in amplitude beyond that of stable phonons. This may result in a phase transition if the displacement pattern is stabilised at some amplitude: this may happen as displacement exceeds the harmonic limit. If an *ab initio* calculation yields imaginary frequencies the proposed structure cannot be stable and an alternative structure must be sought.

Such information supplements that obtained from $E(V)$ curves, which only indicate whether one structure is energetically stable with respect to another. With the use of phonon information, it may be identified that a structure is dynamically unstable before the lower-energy structures are found. Even low-frequency stable modes can be a indicator that at some volume (and hence some pressure) an energy-lowering distortion may become possible, since coupling to elastic constants may further lower the frequency of the mode.

When an unstable phonon is found, more stable structures may be found by displacing the ions according to the corresponding phonon eigenvector, with some estimated

amplitude, and allowing the ions to relax to equilibrium. Any initial symmetry which previously prevented this displacement is now broken, and new structural degrees of freedom are available. The cell parameters should also be re-relaxed to equilibrium as it is likely that freezing in the unstable mode will induce an internal stress. This is especially important if the space group of the resulting distorted structure has lower symmetry than the initial unit cell. This use of phonon information is illustrated in the following Chapters.

4.5 Applications

In addition to results presented in subsequent Chapters, the methods of relaxing the unit cell and finding vibrational properties which have been implemented as part of this work have been used for a variety of other systems. Some of these have been investigated in collaboration with others, and are summarised here since they were not executed solely by the author.

4.5.1 Silicon

Diamond-structured silicon has extremely high symmetry, with a primitive unit cell of only two atoms, so it is straightforward to construct a supercell containing many primitive unit cells and thus obtain phonons at several points in the Brillouin Zone. Several important aspects of such calculations are illustrated.

The force constants shown in Figure 4.2 were used, obtained from a calculation in a supercell of 64 atoms using the CETEP code. In such a large cell, it was essential to enforce symmetry during the simulation to suppress noise in the calculation; proper convergence of the forces required almost double the number of iterations needed to converge the energy. However, as noted above, remaining anharmonicity may break some elements of full cubic symmetry: the structure with one displacement has only four of the full 48 cubic symmetry elements, which are enforced during simulation, so

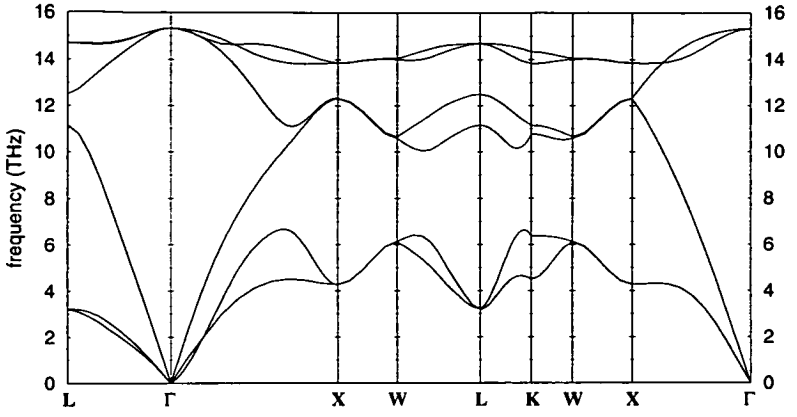


Figure 4.5. Phonon dispersion curve for diamond-structured Si assuming that atoms further away than 10 Å do not interact; results agree well with experiment apart from degeneracy breaking in the TA modes along ΓL.

any of the remaining symmetries may be violated.

Strictly, (4.10) dictates that such a supercell allows accurate calculations only at Γ, X, L, W and halfway along Δ (ΓX) and Σ ($\Gamma K X$). However, Figure 4.2 shows that the forces have practically decayed to zero by the edge of the cell, so it may be assumed that there is negligible interaction between neighbouring supercells. The sum of (4.11) may then be used as an approximation to (4.7) for any wavevector $\mathbf{q}$, giving $\omega(\mathbf{q})$ for the entire Brillouin Zone. The sum rule of (4.16) was enforced as shown in (4.18), inducing slight symmetry breaking. The dispersion curve shown in Figure 4.5 was obtained.

It was found that although the inevitable noise in the forces produced noticeable degeneracy breaking and non-zero acoustic modes at Γ when the 192×192 matrix was diagonalised, when the 6×6 matrix at each $\mathbf{q}$ was formed and diagonalised, the errors reduced in significance. This is probably because the extra symmetry elements in the perfect structure, which are not in the space group of the distorted structure, have the effect of leaving some symmetry unenforced in the forces from CETEP, but on performing such extensive summation, many such errors simply cancel out. However, it may also be affected by the smaller computational rounding errors during such a relatively simple

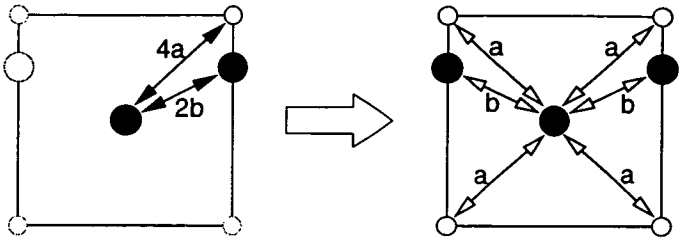


Figure 4.6. Periodic boundary conditions dictate that an atom on the edge of the unit cell is considered only once, creating an asymmetric set of interactions. Symmetry can be recovered by assuming each interaction is in fact the sum over a set. For example, the interaction of an atom with another at separation $(\frac{1}{2}, y, z)$ (where $y, z < \frac{1}{2}$) is replaced by two half-strength interactions with atoms at $(\frac{1}{2}, y, z)$ and $(-\frac{1}{2}, y, z)$.

diagonalisation.

There is significant degeneracy breaking along branches between the strictly allowed wavevectors, which is most obvious in the ΓL branch. This is due to the fact that the sum in (4.11) is not being performed over a perfectly symmetrical set of unit cells: due to the use of periodic boundary conditions, an atom in the face of the supercell is included at only one side in (4.11). However, the degeneracy can be recovered, and the approximation of using (4.11) improved, with a symmetrical set of interactions, generated by splitting the calculated forces into sums over periodic images when these are equidistant. This is illustrated in Figure 4.6. This generates the dispersion curve shown in Figure 4.7. It is noticeable that the ΓL and ΓX TA branches are flattened by this procedure, and are then in better agreement with experiment [80]; the apparent crossover of the two acoustic branches along ΓKX in Figure 4.5 is also eliminated.

There remains some noticeable degeneracy breaking at W; this was found to be rectified if dynamical matrix elements were averaged over all force constants related by symmetry. This process did, however, slightly reduce the accuracy of the zero frequency translation modes at Γ, since the large number of symmetry elements resulted in a very large number of summations.

Many empirical models used to calculate dispersion curves assume that interactions

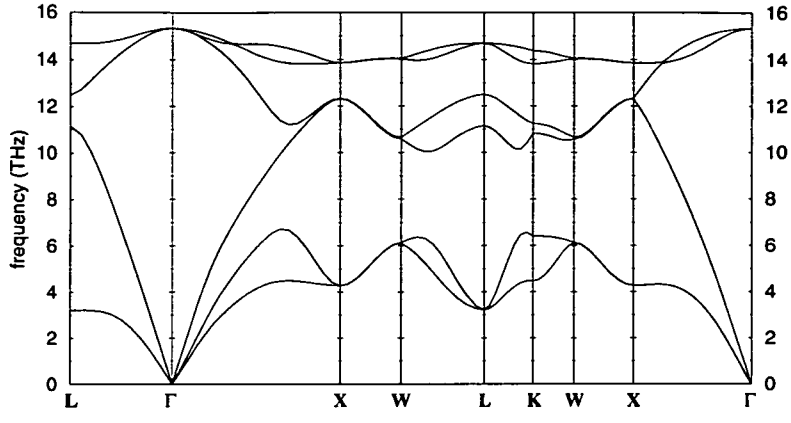


Figure 4.7. Phonon dispersion curve for diamond-structured Si, with degeneracy recovered by splitting interactions over equidistant periodic images of each shell.

are very short-ranged: typically, interactions up to only second neighbours are considered [81, 82, 83]. The interatomic force constants calculated above may be used to investigate the validity of such models by including in the dynamical matrix only interactions between atoms separated by less than a certain distance; the remaining force constants are set to zero. The sum rule given in (4.16) requires that the self-interaction force constants are then reset.

The simplest possible such approximation is that in which only nearest neighbours interact. This imposes the requirement that all interactions are central [81], in which case (4.19) is automatically satisfied; the forces obtained from CETEP with symmetry enforced demonstrate this constraint exactly. The dispersion curve obtained is shown in Figure 4.8.

The cutoff applied to the interatomic force constants to obtain only nearest neighbour interactions is rather smaller than the size of the unit cell, so the complete shell of four first neighbours is included. It is interesting that if the frequencies are squared, to recover eigenvalues of $\mathbf{D}(\mathbf{q})$, then they lie symmetrically either side of $\frac{1}{2}\nu^2_{\Gamma,O}$. This is a consequence of the equivalence of the nearest neighbour model to only three non-zero values (disregarding sign) for dynamical matrix elements ($\{\kappa = \kappa', \alpha = \beta\}$, $\{\kappa \neq \kappa', \alpha = \beta\}$, $\{\kappa \neq \kappa', \alpha \neq \beta\}$) in some appropriate basis set, and is explained in

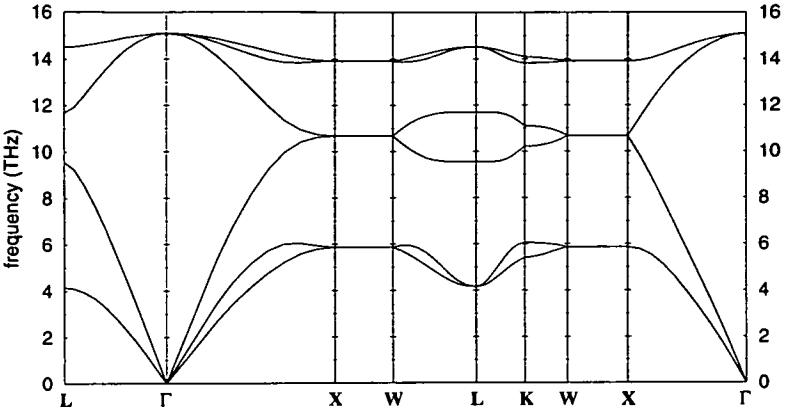


Figure 4.8. Dispersion curve of silicon assuming a nearest-neighbour (central) interaction only. Of particular note are the completely flat bands along XW.

more detail in Appendix B.

Between the limit of all available interactions and those of nearest neighbours only, successive complete shells (at increasing distance) of interactions were added. Figure 4.9 gives the frequencies of the TA modes at points along ΓL and ΓX, showing that the convergence to the final result is slow and oscillating. However, only successive shells up to fourth neighbours correspond to perfect shells in an infinite crystal, due to the use of periodic boundary conditions in a finite supercell. Interactions with more distant sets of atoms are instead mixtures between shells rather than purely atoms at one separation only.

The phonon density of states may also be used to compare these models: phonon frequencies are calculated at a fine grid of wavevectors throughout the Brillouin Zone and a histogram formed. Figure 4.10 shows the results from a 100×100×100 grid of wavevectors using different numbers of neighbours. The singularity at approximately 4.3 THz is due to the flatness of the TA branch near X, and requires interactions with several shells of neighbours to be included: most features are only present with at least five shells of neighbours.

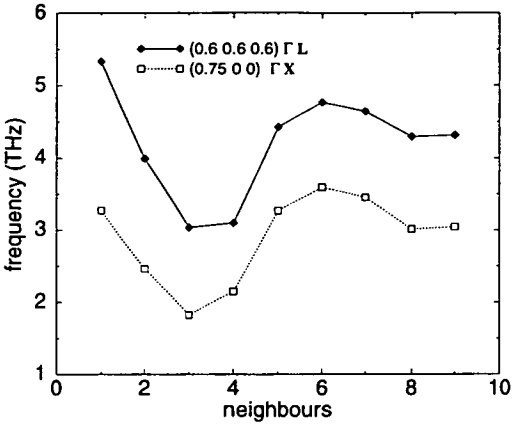


Figure 4.9. Frequencies of transverse acoustic modes along ΓL and ΓX (nearer to the zone boundary than the zone centre in each case), as successive shells of interactions are added. Many shells are needed for proper convergence.

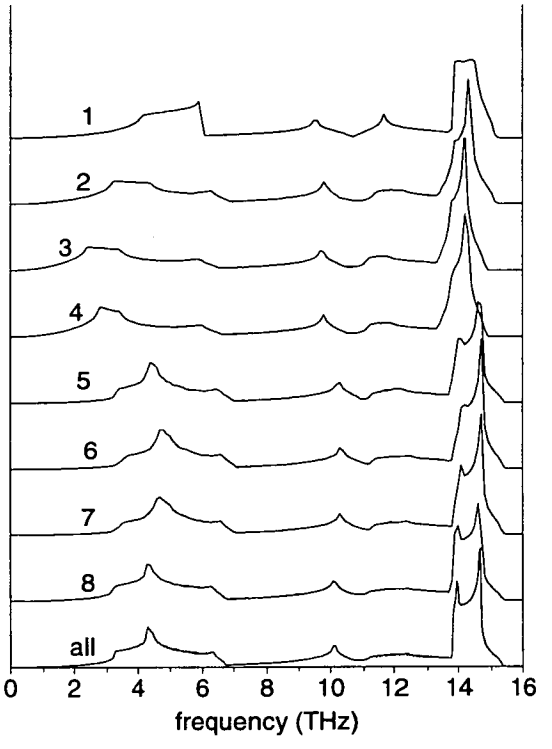


Figure 4.10. Phonon density of states for diamond-structured Si using increasing numbers of interaction shells. Frequencies were calculated over a 100^3 mesh of wavevectors in a parallelopiped formed by the reciprocal lattice vectors, and a histogram of resolution $\Delta\nu$=0.1 THz performed, except for the case of all neighbours, where a 150^3 mesh was used, and $\Delta\nu$=0.0375 THz.

4.5.2 Layered semiconductors GeSe and GeS

H.C. Hsueh used the cell relaxation implementation to determine the equilibrium structure of GeSe [69] and GeS [11] under increasing pressure. These isostructural layered semiconductors have orthorhombic structures with two ionic parameters: to find the equilibrium structure for even one pressure by fitting an energy surface to the lattice parameters would have required extremely large amounts of CPU time. He determined the structural trends under compression, showing that the most compressible axis is not necessarily that normal to the layers. These were compared with results from X-ray powder diffraction experiments and found to be in good agreement.

The frequencies and eigenvectors at each pressure were found using the code developed as part of this thesis. Layered structures are often assumed to vibrate as rigid layers, and corresponding mode assignments are made to experimentally-determined frequencies. Only the frequencies and symmetry of the phonons can be deduced from experiment, so it is not possible to test this assumption experimentally. However, by calculation of phonon eigenvectors from first principles using CASTEP and the analysis described above, it was shown that at even moderate pressures this approximation breaks down and the layers interact producing more complicated displacement patterns. Furthermore, the calculated frequencies were shown to be within a few percent of those from Raman scattering experiments, when the latter were extrapolated to zero temperature.

4.5.3 Alkali halides

High symmetry phonons of sodium chloride and potassium chloride were calculated by H. Akbarzadeh of Isfahan University of Technology, according to a scheme devised by the author which is a slight variation of the technique described above. The equilibrium lattice parameters for NaCl were approximately 5% smaller than experimental values (5.32 Å compared to 5.64 Å), but for KCl a_0=6.48 Å compared to 6.29 Å experimentally. The underestimation in NaCl may be due to the use of the LDA in a highly ionic system,

ω	NaCl		KCl	
(10^{13} rad s^{-1})	calc	expt	calc	expt
Γ LO	3.07	4.9	–	4.0
Γ LA	3.07	3.2	2.92†	2.8
ϵ_0	5.90		4.84	
ϵ_∞	2.34		2.19	
Γ LO (LST)	4.87	5.1	4.34	4.2
X LO	3.77	3.6	3.12	3.0
X LA	2.62	2.7	1.96	2.0
X TO	3.46	3.3	3.00	2.9
X TA	1.65	1.6	1.01	1.1

†: Extrapolation from four points along ΓX.

Table 4.1. Angular phonon frequencies ω (10^{13} rad s^{-1}) at Γ and X points of the Brillouin Zone for NaCl and KCl. Calculated values are obtained from CASTEP simulations (with H. Akbarzadeh); experimental data taken from [80]; values of dielectric constants ϵ_0 and ϵ_∞ from [48]. Γ LO (LST) indicates a LO frequency calculated from the Lyddane-Sachs-Teller relation (2.28).

but the reason for overestimation in the case of KCl is not clear.

Table 4.1 shows the angular frequencies of zone centre and zone boundary (X) modes. The agreement between theory and experiment is reasonable, given that the lattice parameters were significantly smaller than experiment, but the degeneracy between LA and LO modes at Γ can be seen in NaCl. The Γ LO frequency may be corrected by using the Lyddane-Sachs-Teller relation (2.28) with the experimental values of ϵ_0 and ϵ_∞, but this is not a very satisfactory approach for *ab initio* calculations.

4.5.4 Other applications

The method described in this Chapter for determining phonon frequencies from CASTEP or CETEP simulations has been used by several members of the Condensed Matter Group at the University of Edinburgh, in the study of various systems. Applications in which the author was heavily involved are described elsewhere in this thesis. In addition, the equilibrium structure and zone centre vibrational frequencies of $AgGaSe_2$ chalcopyrite were determined using these methods by Karki *et al.* [13]; good agreement

with experiment was obtained.

The vibrational code has also been used by Stewart Clark and co-workers to determine the vibrational frequencies of isolated molecules, including ethane, ethene and ethyne, benzene and a liquid crystal molecule [28]. Where experimental data is available, the theoretical frequencies were found to be in agreement within a few percent.

Chapter 5

Magnesium silicate perovskite

"Wiggle yourselves.. among the three, and make headway the best you can."
'Dow Jr.' (E.G. Paige), *Jerdan Yankee Humour* (1853)

5.1 Introduction

Magnesium silicate makes up most of the material in the Earth's lower mantle, and its properties and phase transitions determine much of the density distribution, thermal properties and discontinuities of the mantle [84, 85]. Such discontinuities, which may be chemical (e.g. involving substitution of iron for magnesium) or physical (a change in crystal structure), have been observed in the lower mantle at several depths, via seismic data [86, 87]. There is still considerable debate as to whether convection in the upper and lower mantles, caused by temperature gradients, occurs within two layers or one single layer [88, 89]. If any of the observed discontinuities were due to temperature-induced phase transitions in $MgSiO_3$ this would establish the temperature at the corresponding depths, and thus help to determine the mantle dynamics.

The phase of $MgSiO_3$ believed to be stable in the mantle has one of the perovskite structures. Such structures are displayed by many compounds of the form ABX_3, and exhibit a wide variety of behaviour, with a multitude of possible phase transitions [90].

Despite much study, with a variety of theoretical techniques (reviewed in [85, 91]), the properties of $MgSiO_3$ are still not fully understood. Attempts to produce samples for X-ray diffraction under mantle pressures have met with difficulty, not least because of the low X-ray scattering cross-section of these ions [92, 93, 94]. However, by comparison with other perovskites, especially $SrTiO_3$, an orthorhombic–tetragonal–cubic series of transitions may be envisaged. This has had some support from observations of twinning in the orthorhombic phase, after quenching from high temperatures [95].

We consider three phases related by successive symmetry-breaking transitions as the temperature is lowered: the hypothetical cubic phase ($Pm3m$) at high T, the observed orthorhombic ($Pbnm$) phase (low T) and an intermediate tetragonal phase ($I4/mcm$). The vibrational dynamics of these modes is very important since transitions between these phases are likely to proceed via soft-mode mechanisms [91]. Density-functional theory calculations of the frequencies and eigenvectors of the high-symmetry phonons of all three phases are presented in this Chapter, followed by investigations to determine which phonons would be involved in transitions between the three phases.

The triply degenerate zone-centre phonons of several other perovskites have already been found from first principles [74, 75, 96]. The eigenvectors of unstable phonons calculated by King-Smith and Vanderbilt [96] correctly predict the symmetry of distorted phases with one formula unit in the unit cell (i.e. those in which only zone-centre phonons can freeze in). The orthorhombic phase of $MgSiO_3$ contains twenty atoms, as opposed to the five atoms in the cubic unit cell. However, the geometry of the orthorhombic unit cell is related to that of the cubic phase by two 'cell-doubling' procedures [90]. Any phonons involved in a transition between cubic and orthorhombic phases must therefore lie either at the zone centre or the zone boundary of the cubic phase, at points determined by the sense of the cell-doubling. A simulation of more than one unit cell of the cubic crystal is thus required to calculate these zone-boundary phonons.

The CASTEP code described earlier in this thesis was used, employing norm-conserving, nonlocal, Kleinman-Bylander pseudopotentials [34, 97] and the local density approximation. A plane-wave basis set cutoff energy of $\mathcal{E}_{\text{cutoff}}$ =500 eV was used, corresponding to approximately 4000 plane waves at each k-point for a twenty atom cell. A Monkhorst-Pack k-point scheme [26] was used to give 4 k-points in the Brillouin Zone.

The structure of this Chapter is as follows: 5.2 describes the general features of the perovskite phases; Sections 5.3, 5.4 and 5.5 present the results of structural optimisation and phonon calculations for cubic, tetragonal and orthorhombic phases respectively. Section 5.6 determines which phonons of the cubic phase are important for any phase transitions, and considers the role of the tetragonal phase. Section 5.8 describes first-principles molecular dynamics simulations which attempt to bound the transition temperature for the tetragonal–orthorhombic transition. Section 5.9 discusses the conclusions which have been reached.

5.2 Perovskite structures

All phases of $MgSiO_3$ perovskite essentially consist of octahedral 'cages' of oxygen ions, linked by sharing oxygens at the corners, and each containing a silicon (B) ion in the centre. In the spaces between the octahedra there are 12-fold coordinated magnesium (A) ions. In the cubic phase ($Pm3m$), these octahedra are aligned along the cubic axes, as shown in Figure 5.1. However, in the orthorhombic structure, the octahedra are rotated around the silicon ions, and the magnesium ions are displaced (also shown in Figure 5.1), giving space group $Pbnm$. This is the form of $MgSiO_3$ observed under ambient conditions [98] and assumed to be the dominant phase in the lower mantle. The tetrahedral structure is an intermediate, having octahedral rotations about only one axis, with $I4/mcm$ symmetry; the reasons for its consideration are given in Section 5.4.2.

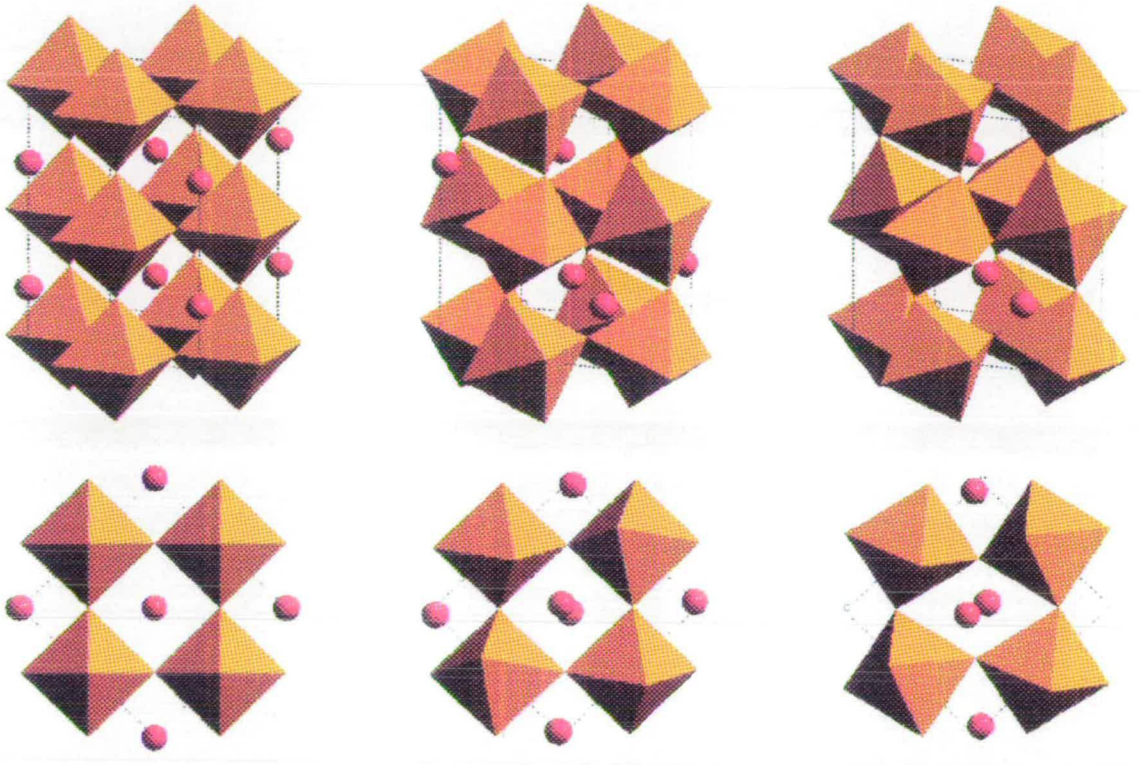


Figure 5.1. Cubic, tetragonal and orthorhombic structures of $MgSiO_3$ perovskite. Twenty atoms of each structure are drawn (four unit cells of the cubic phase; one of the orthorhombic). Silicon ions are enclosed in octahedral cages of oxygen atoms, with magnesium ions in interstices. It can be seen that the three phase are related primarily by rotations of these SiO_6 octahedra; this is discussed fully in the text.

The oxygen octahedra are nearly perfectly rigid, and thus are often treated as distinct entities in discussions of perovskite structures [90]. The 'rigid unit' model of such structures [99] assumes that they are incompressible and cannot be distorted. This has important consequences for coupling between rotation of the octahedra and the strain, and also for the exact nature of the normal modes of the system, as will be discussed more generally in Chapter 6. The applicability of this model to $MgSiO_3$ perovskite is also investigated further in Chapter 6.

5.3 Cubic phase

The cubic phase has a unit cell of only one formula unit (five atoms) and the highest symmetry, so requires the least computational effort of the three phases. There are no structural parameters to determine, so the effects of pressure may be easily investigated, and the high symmetry also reduces both the number of **k**-points in the reduced Brillouin Zone and the number of displacements required to find the vibrational modes. Since it shares many features with the tetragonal and orthorhombic phases, the study of the cubic structure will introduce the important concepts for all three phases.

5.3.1 Equilibrium structure

The quantum-mechanical stress with the Pulay correction was used to relax the unit cell to equilibrium, as described in Chapter 3. The ionic positions are fixed by the cubic symmetry, which was enforced throughout the simulation. The structure stabilised at a lattice parameter of $a_0 = 3.44$ Å. Approximately twenty molecular dynamics steps were required to obtain the equilibrium cell.

This method was checked by the more established technique of performing simulations at different lattice parameters, to find the bulk modulus. This confirmed a minimum energy at $a_0 = 3.44$ Å. The bulk modulus K is sensitive to the fitting procedure: a quartic polynomial fit (R=0.99992) gave K=255.0±0.3 GPa, $K' = \partial K/\partial P = 4 \pm 1$,

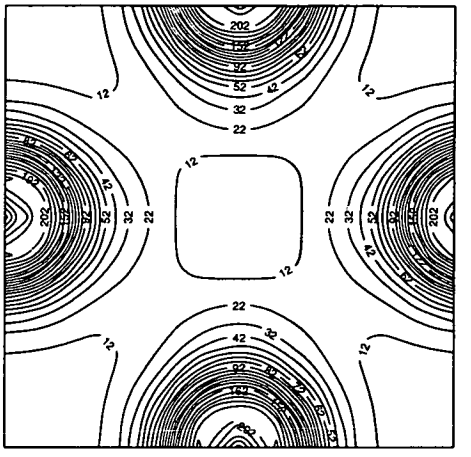


Figure 5.2. Charge density through a plane containing Si ion in the centre, oxygens at edges, and Mg ions at the corners, showing the large degree of ionicity.

in agreement with [100] and a Murnaghan fit gave $K = 241 \pm 3$ GPa and $K' = 3.6 \pm 0.1$. These high results are comparable with experimental values for the orthorhombic phase: $K = 266 \pm 6$ GPa and $K' = 3.9 \pm 0.4$; $MgSiO_3$ has been noted to have bulk moduli higher than most other oxide perovskites [85]. The bonding is mostly ionic, as shown by the plot of charge density in Figure 5.2; this is also in agreement with theoretical findings reviewed elsewhere [85].

5.3.2 Phonons

Zone centre

A simulation of only one primitive cell (five atoms) can only give reliable information about the fifteen zone centre phonons, due to the use of periodic boundary conditions over a distance smaller than the range of the forces (see Chapter 4). Four displacements were needed to construct the zone-centre dynamical matrix: one each for the silicon and magnesium ions, and two for the oxygens (parallel and perpendicular to cell face). A separate simulation was performed for each distortion, with the appropriate reduced symmetry. The entire dynamical matrix was constructed for one unit cell using the

	ω	eigenvector				
	(THz)	Mg	OI	OII	OIII	Si
Γ_{15}	26.33	0.024	-0.894	0.224	0.224	0.315
Γ_{15}	13.37	-0.253	0.070	-0.398	-0.398	0.783
Γ_{15}	0.000	0.492	0.400	0.400	0.400	0.529
Γ_{15}	$6.01i$	-0.833	0.189	0.363	0.363	0.084
Γ_{25}	5.189	0.000	0.000	-1.000	1.000	0.000

Table 5.1. Frequencies and eigenvectors of the zone-centre modes of cubic $MgSiO_3$, at $a = 3.44$ Å. Displacements are given in mass-reduced coordinates. The imaginary frequency represents an unstable mode.

method described in Chapter 4, employing the full cubic symmetry. On diagonalisation, five sets of triply degenerate Γ phonons were obtained; the phonons are strictly at the zone centre so these are TO and TA modes, since the LO/TO splitting is not accessible (Section 4.3.2).

The eigenvectors of the Γ-point modes are shown in Table 5.1, in mass-reduced coordinates $\boldsymbol{\xi}_\kappa = \sqrt{m_\kappa}\mathbf{u}_\kappa$. It is noteworthy that there is an unstable mode, having an imaginary frequency, which is dominated by motion of the magnesium (A) ions against the rest of the crystal. This is in contrast to many other perovskites, in which it is the B atoms which 'rattle' inside the oxygen cage [90, 96]. Since there is no evidence of strong covalent Si–O bonds in $MgSiO_3$, this motion is probably due to the relative sizes of the oxygen, silicon and magnesium ions [92, 96]: magnesium is one of the smallest cations. The effective size of an ion can, however, be dependent on the choice of Mg pseudopotential, as described in Section 5.7.

Three acoustic translation modes are expected with zero frequency; this requires all the forces from the displacement of one atom to sum to zero (Section 4.3.1). This was partially achieved by adjusting the force on each coordinate when displaced, and hence some of the on-diagonal elements of the dynamical matrix, by less than 1%, giving an indication of the accuracy of the calculations. Anharmonicity was investigated by performing a second set of simulations with twice the displacements. This gave force

constants which differed from the originals by less than 0.1%, so anharmonicity was considered negligible. The effect of k-point sampling on phonons of other perovskites has been found to be significant for other perovskites [96]. However, on increasing the k-point sampling from 4 points in the Brillouin Zone to 32, the changes in frequency were approximately 0.2–0.4 THz and the changes in eigenvectors 10% at most (in those not entirely determined by symmetry); there were no qualitative changes. The effect of k-point sampling is smaller in simulation of larger cells, such as those used in most of this work.

Zone boundary

A larger simulation is needed to investigate phonons away from the zone centre. A supercell of four primitive unit cells ($a' = b' = \sqrt{2}a_0$; $c' = 2a_0$; 20 atoms) was chosen, since this reflects the geometry of the orthorhombic phase. The lattice parameter a_0 obtained above was used, and the force constants obtained in a similar manner to the five atom cell. This cell allows direct calculation of phonons at the Γ, X, M and R points of the cubic Brillouin Zone, by (4.10). It should be noted that the omission of LO/TO splitting only occurs at Γ, so phonons at other points are all calculated correctly.

Because the supercell was not cubic, displacements in the z direction as well as the x direction were required to construct the Fourier-transformed dynamical matrix $\mathbf{D}(\mathbf{q})$ at the zone centre and boundaries, as described in Chapter 4. The frequencies obtained are shown as circles in Figure 5.3. The breaking of cubic symmetry due to the tetragonal shape of the supercell has the results that phonons are not perfectly degenerate, as can be seen, for example, at the R point of the Brillouin Zone. This is one indication of the level of anharmonicity and error in the force calculations.

The eigenvectors of all the phonons were also obtained from the dynamical matrix. The most unstable modes were found at R and M (around $12i$ THz), consisting of rotations of near-rigid octahedra around the silicon atoms. At M, the mode involves

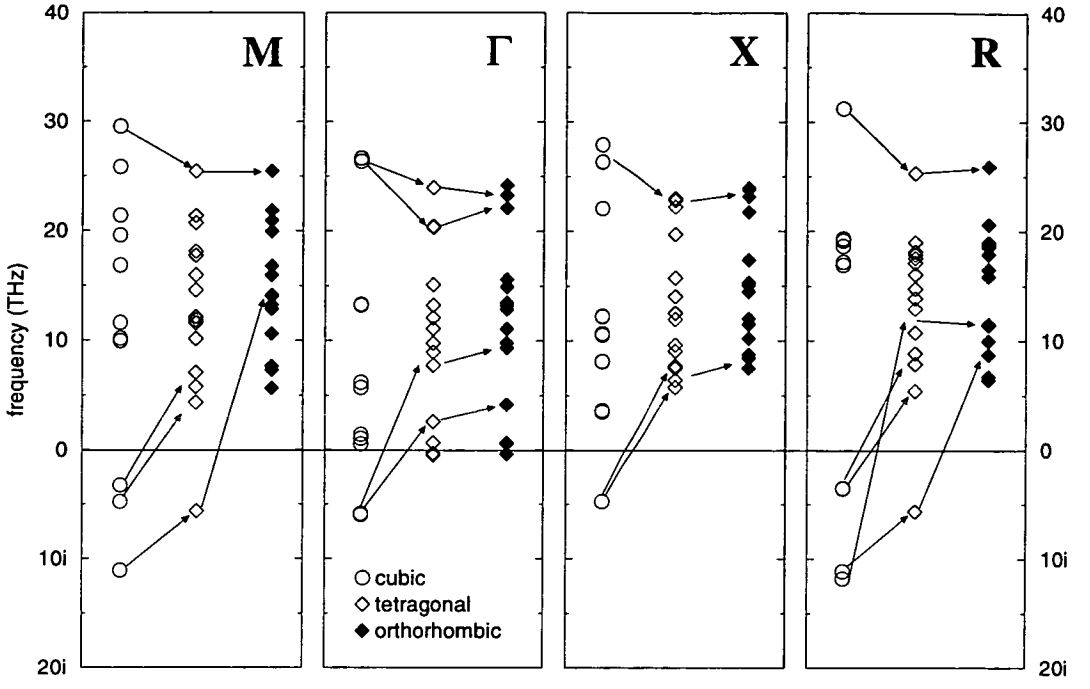


Figure 5.3. Cubic, tetragonal and orthorhombic phonons at each part of the Brillouin Zone of the corresponding cubic phonon. There are unstable phonons at all parts of the Brillouin Zone in the cubic phase, but all the phonons in the orthorhombic phase are stable. The tetragonal structure is described in Section 5.4.2 and has only two unstable modes. The horizontal axes trace the progression from cubic to orthorhombic; the arrows link phonons with similar eigenvectors as the structure becomes more distorted. Loss of degeneracy is due to the choice of supercell.

rotation about z, and is denoted M_2 by symmetry; rotations at R are around all three axes and denoted R_{25}. These modes have been previously predicted to be the only zone-boundary rigid unit modes in cubic perovskites [99] (see Chapter 6). These results are also consistent with the unstable modes found by other theoretical methods [86, 88, 92], but those workers found no other unstable phonons, whereas in this work several unstable modes involving magnesium displacement are found (around $5i$ THz). The stable modes mostly involve squashing and distorting the octahedra and displacement of magnesium ions.

Unstable modes can be described by a simple model consisting of quartic local potentials coupled by harmonic interactions [1]; this model will be described more

fully in Chapter 7. A predominantly unstable vibrational band is obtained if the local potential has a deep double well form, giving order-disorder type transitions, but if the local potentials have a single central minimum (displacive type) at least part of the band will be stable [8, 101]. The high-temperature phase in order-disorder systems only has an order parameter of zero when averaged over the whole system; on a local scale the coordinates hop between positive and negative values.

In the case of cubic $MgSiO_3$, it was found that there are unstable phonon modes at all points investigated, and although there is considerable mixing and crossing between the several branches on which these modes lie [92], some of these branches are likely to be completely unstable. Any phase transition involving these branches would thus have considerable order-disorder character, especially if they are nearly dispersionless. In particular, the set of unstable modes involving Mg displacement around $5i$ THz may be imagined from Figure 5.3 to lie on such a band. An idealised case may be considered, in which the Mg ions occupy rigid interstices which which they do not interact, and are only weakly coupled with each other. If the potential within each site has more than one minimum, the resulting phonons will lie on unstable, almost dispersionless bands, and participate in an order-disorder transition over the sites. However, there is in fact considerable mixing between different bands, which complicates this scenario: some bands contain both octahedral rotation modes and Mg displacement modes (at different parts of the Brillouin Zone) and the prediction of transition character from the phonon branches is therefore less straightforward.

Effect of pressure on frequency

The five-atom cubic cell is sufficiently small and of high symmetry that the Hellmann-Feynman forces from one distortion could be found in under half an hour on an Alpha AXP workstation. In addition, the ionic positions are independent of volume. It was thus straightforward to investigate the variation of Γ mode frequencies with cell size and obtain estimates of the Grüneisen parameters $\gamma_i = -\partial(\ln\omega_i)/\partial\ln\Omega$, which are

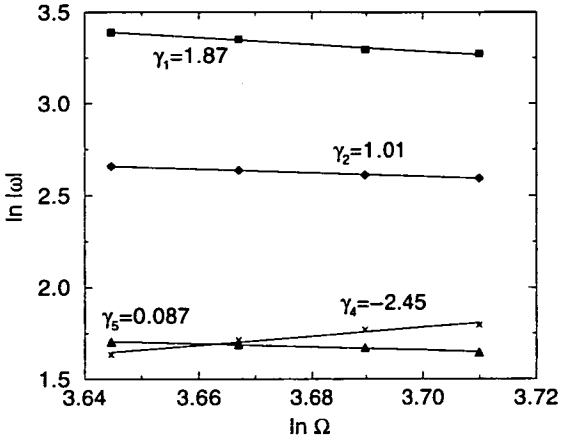


Figure 5.4. Variation of frequency of cubic zone-centre phonons with cell volume, to give Grüneisen parameters.

important for predicting the thermal properties of materials. The results are shown in Figure 5.4. Larger cells are more unstable with respect to the soft magnesium mode, which is consistent with the simple 'rattling' picture suggested above. The frequencies of all other Γ phonons decrease as the cell is enlarged, i.e. have positive γ, with similar magnitudes to those obtained experimentally in the orthorhombic phase [85]. These results act as further confirmation of the method.

The effect of pressure on the M_2 unstable mode was also investigated, by introducing some amplitude of the octahedral rotation at several different cell volumes, and calculating the restoring force from first principles. It was found that in this case the magnitude of the imaginary frequency increased with pressure. Under compression the structure therefore becomes more unstable with respect to this distortion, in contrast to the behaviour of the unstable mode at Γ. This coupling between the unit cell and a mode involving octahedral rotation may be explained in term of rigid unit modes, which are discussed in detail in Chapter 6. These results are in agreement with the trends found by [92] and imply that increased pressure does not favour the cubic phase.

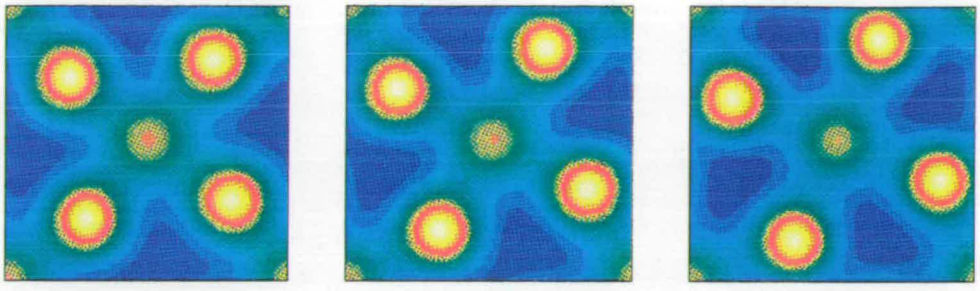
5.4 Phonons of intermediate structures

Introducing some amplitude of either the R_{25} or M_2 rotational mode into the cubic phase creates a tetragonal structure: the rotation axis becomes the principal axis, although the space group is dependent on the particular mode. At some amplitude of each mode, the structure should stabilise to a local equilibrium, creating a new phase whose properties may be investigated. Since phonon modes strictly imply ionic motion at constant volume, the lattice parameters of the cubic phase were initially retained. At the local equilibrium of one of these modes, there may still be some unstable phonons, which correspond to unstable modes of the cubic phase which have not been frozen in. In both phases, the normal modes were calculated using a similar procedure to that for the cubic phase.

5.4.1 Freezing in M_2 mode

We investigated the configuration formed by freezing the M_2-rotation of the oxygen octahedra into the cubic phase, giving a unit cell of 10 atoms ($a' = b' = \sqrt{2}a_0$, $c = a_0$) with space group $I4/mmm$. The charge density distributions for different amplitudes of this mode are shown in Figure 5.5. An equilibrium amplitude was found, with zero forces on ions, but this is likely to be only a saddle point of the overall energy surface. Beyond this amplitude, charge overlap between oxygen ions in unlinked octahedra develops, shown by charge density in regions previously empty.

Only phonons at Γ and M of the cubic phase are commensurate with this larger cell. Of these, only modes containing distortions in the xy plane were considered, since phonons involving distortions parallel to z will be uncoupled to those along xy by symmetry, and are likely to be very similar to those in the undistorted cubic phase. Phonons with the same symmetry as the frozen mode, involving only OII distortion, remained normal modes, although their frequencies were changed. Other eigenvectors are linear combinations of those of the original cubic Γ and M modes.



(a) (b) (c)

Figure 5.5. Electronic charge density across an (001) plane near to the centre of a SiO_6 octahedron, with various amplitudes of the cubic M_2 rotational phonon frozen in. Blue regions indicate low density. A small amplitude gives (a); the equilibrium amplitude (fractional displacement of oxygen is 7%) has charge density (b); and further rotation causes overlap between oxygen ions in different octahedra as in (c).

The M_2 octahedral rotation mode, which had a frequency of $11.1i$ THz in the cubic structure, is now stable with a frequency of 13.5 THz. In the simplest model of freezing in an unstable mode (see Section 5.6.2), an unstable frequency of $i\omega_0$ would lead to a stable frequency of $\sqrt{2}\omega_0$. The stable frequency found here is a little softer than this, indicating a small departure from a pure quartic potential. In the work described below it was found that this phonon stiffened further in the full orthorhombic structure, to 14.1 THz.

The mode with the highest frequency in the cubic structure was still the highest, but the frequency was reduced from 29.6 THz to 24.3 THz. A degenerate pair of unstable modes were found, predominantly consisting of displacements of the magnesium ions, with the degree of instability ($3.66i$ THz) slightly decreased from the corresponding mode in the cubic structure ($4.76i$ THz).

5.4.2 Freezing in R_{25} mode

When some amplitude of the cubic R_{25} mode is frozen into the cubic supercell, a twenty atom body-centred cell (space group $I4/mcm$) may be constructed, which reflects the geometry of the orthorhombic phase. The values of the lattice parameters are kept as

those of the cubic twenty atom cell. The cubic R_{25} mode is triply degenerate; rotation was chosen about the y axis of the twenty-atom cell (xy axis of the cubic unit cell), since this is the rotation observed in the orthorhombic phase.

At large amplitudes of R_{25}, non-negligible forces were found to be exerted on the magnesiums, so the harmonic eigenvectors are no longer valid; the general R_{25} mode should therefore be thought of as including some magnesium motion. To reach a local equilibrium, the magnesium ions were therefore also relaxed and were found to move along x with approximately 25% of the O$_{\text{I}}$ displacement. This additional magnesium displacement is very close to the eigenvector of another of the unstable phonons of the cubic phase at R (in fact, like the R_{25} mode, an alternative choice of eigenvectors in a degenerate triplet). The phase formed in this way may thus described by two ionic structural parameters, or, alternatively, by introducing non-zero amplitudes of two cubic phonons.

The sixty phonons commensurate with this cell were found. These are the Γ and X modes of this phase, but will be linear combinations of the Γ, X, M and R phonons of the cubic phase, since this cell is quadrupled with respect to the cubic five-atom cell. The eigenvectors of this phase were compared to those of the cubic phonons, by taking scalar products between eigenvectors of the two phases. Each phonon of the tetragonal phase was thus identified with its closest match in the cubic structure; in what follows phonons of all phases are denoted by the location in the cubic Brillouin Zone of their cubic match. The tetragonal and orthorhombic phonons are plotted against this part of the cubic Brillouin Zone in Figure 5.3 to enable comparisons to be made.

Only two unstable modes were found in this structure: the R and M modes corresponding to rotation about the z axis, with the latter slightly more unstable ($5.64i$ THz). The eigenvector of the M mode consists of M_2-type octahedral rotation together with some motion of Mg and O$_{\text{I}}$. We thus conclude that this is strongly related to the original cubic M_2 mode, although it is much less unstable. All other modes which were unstable in the cubic structure, including the magnesium displacement modes, are

stable in this structure. Since there are no longer unstable phonons at all points of the Brillouin Zone, a soft-mode transition from this tetragonal phase to the orthorhombic would have more displacive character than any cubic–tetragonal transition.

5.4.3 Proposal of intermediate phase

Both these phases contain some of the structural features found in the orthorhombic phase, so they may be considered as possible intermediates between the cubic and orthorhombic phases. Of the two modes used here, the R_{25} mode is the most unstable in the cubic phase, and the resulting tetragonal structure has lower energy than the $I4/mmm$ phase discussed above. On cooling from the cubic phase, the R_{25} mode would thus be expected to freeze in first; conversely, on heating the orthorhombic phase there would be sufficient thermal energy to reach the $I4/mcm$ phase before the other. This structure was thus chosen as a possible intermediate tetragonal phase between cubic and orthorhombic. The structure is shown in Figure 5.1, and in the remainder of this Chapter we consider the possibility that it forms a distinct thermodynamic phase.

5.5 Orthorhombic phase

5.5.1 Structure

The orthorhombic phase has much less symmetry than the cubic phase, and hence requires more structural relaxation to find the equilibrium structure. As in the cubic phase, the three cell edges were allowed to move under the internal and Pulay stresses, but the ions were also relaxed under the Hellmann-Feynman forces, as described in Chapter 3. This simulation was started from previous *ab initio* values [100] and the symmetry of the *Pbnm* space group was enforced throughout relaxation. There are ten structural parameters, and so this relaxation required a considerable amount of computational effort: several hundred hours of CPU time on a Alpha AXP were needed. As discussed in Chapter 3, the non-linear nature of the energy as a function of the cell

	this work	previous [100]	experiment [94]	
P	20 GPa	10 GPa	0 GPa	10.6 GPa
a(Å)	4.635	4.711	4.777	4.710
b(Å)	4.833	4.880	4.927	4.873
c(Å)	6.771	6.851	6.898	6.790
Mg_x	0.5157	0.5174	0.5131	0.511
Mg_y	0.5603	0.5614	0.5563	0.557
O^1_x	0.1155	0.1128	0.1031	0.099
O^1_y	0.4572	0.4608	0.4654	0.464
O^2_x	0.1914	0.1928	0.1953	0.196
O^2_y	0.1968	0.1995	0.2010	0.201
O^2_z	0.5594	0.5582	0.5510	0.561

Table 5.2. Structural parameters of equilibrium structure of orthorhombic perovskite, with assumed effective pressures deduced from the structure and experimental applied pressures.

and ionic parameters reduces the efficiency of the relaxation scheme. Near equilibrium, the movement of the cell was terminated so that the ionic positions could be found to sufficient accuracy for phonon calculations.

The structural parameters obtained are given in Table 5.2. The unit cell is smaller than most previous work, but the structural parameters follow the reported trend of an increase in distortion under compression [92, 100, 102]. Following Bukowinski and Wolf [88] the rotation angles $\theta = \cos^{-1}(a/b)$ of R_{25} and $\phi = \cos^{-1}(\sqrt{2}a/c)$ of the M_2 rotations are compared (Table 5.3), effectively assuming that the SiO_6 octahedra remain rigid as they rotate, causing a change in cell vectors (see Chapter 6). This confirms that the structure presented here is the most distorted as well as the smallest.

However, the Si–O bondlength is only 1% smaller than that observed experimentally, which is typical for calculations using the LDA [19]. These results suggested that the structure found by the present calculations is under an effective external pressure, since both volume and structure are affected. The effect of the LDA has been previously compared to a pressure in this and other systems [101, 103], and generalised-gradient corrections have also been similarly treated [104].

	$\theta(M_2)$ (°)	$\phi(R_{25})$ (°)	Si–O (Å)
this work	16.5	14.5	1.77
other *ab initio* [100]	15.1	13.5	1.79
MEG [88]	10.4	7.0	1.904
SSMEG [92]	10.8	9.8	1.776
expt, 0 GPa [94]	14.3	11.6	1.792
expt, 10.6 GPa [94]	14.9	11.9	1.777

Table 5.3. Octahedral rotation angles inferred from cell shape, and mean Si–O bondlength. This work gives the most distorted structure, but also has the smallest cell, supporting the approximation of the effect of the LDA as an external pressure.

We can use this approach in comparing the present work to previous density-functional calculations of $MgSiO_3$: the cell parameters of a previous *ab initio* calculation by Wentzcovitch *et al.* [100], when compared to experimental results [94], suggest that they are equivalent to a pressure of 8–10 GPa. The effect of pressure on cell parameters calculated in [100] suggests that the cell parameters from the present work correspond to an additional 10–12 GPa. The experimental bulk modulus and equilibrium volume give a pressure of 20 GPa for the volume found in this work, which is consistent with these estimates. Hence it is assumed in what follows that the present orthorhombic structure is under an effective pressure of 20 GPa.

The structural parameters are harder to assign to a consistent pressure, since the experimental data [94] does not show the simple monotonic trends predicted by Wentzcovitch *et al.* and does not extend to 20 GPa. However, the experimental data does show increasing octahedral rotation under pressure, and most of our structural parameters continue the calculated trends, so the structural parameters are also consistent with an effective pressure. The remaining small differences between our calculations and those of Wentzcovitch *et al.* are probably due to differences in pseudopotentials.

In addition to the displacements of ions within the cell, the orthorhombic unit cell has a density 8% higher than the cubic phase, although the oxygen octahedra are larger by about 9% in volume. There is therefore some coupling between the displacements of

the ions and the strain of the unit cell. The angle between Si–OI and Si–OII bonds is 2° from perpendicular, so there is also a small degree of distortion of the SiO_6 octahedra.

To examine this coupling between distortion and strain, the ionic positions were relaxed within the equilibrium cell of twenty atoms of the cubic structure, i.e. using $a = b = \sqrt{2}a_0$, $c = 2a_0$. Under these constraints, the octahedral rotation was about 5% less than that of the full orthorhombic cell found above. However, the octahedra were 18% larger than in the cubic phase, with approximately 1° angular distortion. This will be discussed further in Chapter 6.

5.5.2 Phonons of orthorhombic phase

The orthorhombic symmetry required that displacements were made along x, y and z to construct the entire dynamical matrix. The forces remaining on the orthorhombic structure after relaxation were less than 0.001 eV/Å, and so they were subtracted from the forces from simulations with ionic displacements, assuming that the harmonic approximation applies to sufficient accuracy (Chapter 4). To minimise anharmonic effects, the resulting force constants were averaged over pairs between the same atoms, which must be equal by the Hermitian symmetry of the dynamical matrix; this process changed the phonon frequencies by a maximum of 2% for the optic phonons.

As expected, there were no unstable modes. Only the Γ phonons of the orthorhombic Brillouin Zone can be found, since only one unit cell is simulated. The zone-centre corresponds to the Γ, X, M and R points of the cubic phase, due to the quadrupling of the unit cell. Like those of the tetragonal phase, these phonons may thus also be matched with their closest cubic modes and plotted in Figure 5.3 against the appropriate part of the cubic Brillouin Zone.

The phonon with the highest frequency at each point of the Brillouin Zone is the same in cubic, tetragonal and orthorhombic phases (an octahedral breathing mode at R), but has a frequency of 25.9 THz in the orthorhombic phase, compared to 29.6 THz in the cubic phase. At lower frequencies the orthorhombic eigenvectors are often linear

this work	Raman expts [85]	
A_g modes	21.9 GPa	0 GPa
25.5	28.5[a]	27.0
20.0	22.4[b]	20.7
18.8	18.3	16.1
15.2	17.3	15.0
11.4	12.7	11.4
10.0	11.6	11.1[b]
8.6	9.45	8.49
	9.00	8.34

a: extrapolated from higher pressures
b: using average pressure shift of other modes, and experimentally uncertain

Table 5.4. Calculated frequencies (in THz) of A_g modes of orthorhombic phase, which are expected to give strong Raman signals, and experimentally observed Raman modes. Five of the seven A_g modes show frequencies close to observed modes.

combinations of several cubic phonons. The phonons which were unstable in the cubic phase are now all stable, and mostly have low frequency. The octahedral rotation modes described in the cubic and tetragonal phases were still clearly identifiable. The eigenvectors were used to determine the symmetry of the phonons (g or u): the 24 symmetric modes were found, which are the Raman modes, and within this group seven A_g modes can be identified. These are likely to include the strongest Raman signals, and are compared with experimental values in Table 5.4. There is generally a good agreement, although only five of the A_g phonons have frequencies close to the observed Raman modes: previous detailed analysis of calculated eigenvectors [85] also assigned only five A_g modes to Raman signals.

Any remaining effects of anharmonicity were further reduced by recalculating some force constants using half the original displacement. This only had a small effect on the calculated phonons and eigenvectors: the maximum change in frequency was 0.5 THz when force constants from Mg and O1 z displacements were recalculated in this way.

5.6 Roles of individual phonons

In Section 5.4.2, the tetragonal $I4/mcm$ phase was described as formed by freezing two phonons into the cubic phase. It was also shown that the M_2 and R_{25} phonons which are unstable in the cubic phase become stable after gaining some non-zero permanent amplitude. The roles of individual phonons in transitions between the three phases, and the energy surfaces defined by treating their amplitudes as variables are investigated here.

5.6.1 Cubic phonon modes contributing to orthorhombic structure

The tetragonal phases described in Section 5.4 were formed by freezing a permanent amplitude of one or two phonons into the cubic phase. Likewise, the positions of the ions in the stable orthorhombic phase may also be described in terms of a combination of phonons from the cubic phase; if the correct set of phonon wavevectors is chosen, this corresponds merely to a change of variables. Pure phonon modes involve only displacements at constant volume, so only the fractional atomic positions in the orthorhombic cell are used (i.e. as if it had the same cell parameters as the cubic cell) and a small change in cell parameters is in principle allowed once these positions are fixed. The mass-reduced displacement from the cubic configuration, $\mathbf{d}_\kappa$, is found for each atom κ in the stable orthorhombic structure:

$$\mathbf{d}_\kappa = \sqrt{m_\kappa}[\mathbf{R}_\kappa(\text{orth}) - \mathbf{R}_\kappa(\text{cubic})] \tag{5.1}$$

Phonon coefficients c_j may then be defined such that

$$\mathbf{d}_\kappa = \sum_j c_j \mathbf{p}_\kappa^j \tag{5.2}$$

where $\mathbf{p}$ are the mass-reduced eigenvectors of the cubic crystal. It is straightforward to find the coefficients c_j since the eigenvectors $\mathbf{p}$ are orthonormal, and for the equilibrium

rank according to frequency		frequency (THz)	c_0 ($\sqrt{\text{a.m.u.}}$ Å)	E_0 (eV)	description
R	(59-60)	$11.8i$	6.48	5.98	octahedral rotation about xy
M	(57)	$11.1i$	4.35	2.38	octahedral rotation about z
X	(50+51)	$4.73i$	3.24	0.240	mostly Mg displacement
R	(47-48)	$3.48i$	0.841	0.0042	mostly Mg displacement
X	(30+31)	10.5	0.780	stable	Mg and O1 displacement
M	(13)	19.6	0.210	stable	octahedral squash

Table 5.5. Calculation of the coefficients of cubic phonons frozen into the orthorhombic distorted phase. Phonons are labelled by the point of the cubic Brillouin Zone at which they occur, and their rank out of all sixty phonons. The depth E_0 of the minima of the well has been calculated for unstable phonons, using c_0 and the frequency, assuming no coupling. Where two ranks are indicated, phonons form degenerate pairs but the amplitude and energy are for the active combination. Stable phonons correspond to a well with only one minimum, at the origin, so E_0 is zero.

orthorhombic cell they are denoted c_0.

This is not a perfect description of the structure, since the harmonic eigenvectors $\mathbf{p}^j$ do not always describe the mode at large displacement. However, the Γ, X, M and R cubic phonon eigenvectors do span the complete set of all possible strain-conserving distortions in the twenty atom cell. All distortions not involving strain, that is, those that involve ionic displacements only, can be expressed by using the cubic phonons at these points of the Brillouin Zone as a basis set.

The results of this analysis are shown in Table 5.5. Four of the fifteen unstable modes and two of the 42 stable phonons of the cubic phase have significant non-zero coefficients. Phonons consisting predominantly of magnesium displacement are present much more weakly than those involving octahedral rotations, and were shown in Section 5.4.2 to be strongly coupled to the rotational modes. Hence we concentrate on the two octahedral rotation phonons in discussing the orthorhombic phase. The coefficients were also calculated for the distorted structure in the cubic cell, where they were all found to be smaller.

The tetragonal phase proposed in Section 5.4.2 may only contain non-zero amplitudes of phonons at Γ and R of the cubic phase. The two phonons in Table 5.5 at R are exactly those contributing to this tetragonal phase. The other four modes (at X and M) can therefore be assumed to freeze in from the tetragonal intermediate to form the orthorhombic phase, providing a natural pathway for a cubic–tetragonal–orthorhombic series of transitions; the rotational M_2 mode, which is one of these four, was found to still be unstable in the tetragonal phase.

5.6.2 Energy surfaces of rotational modes

The energy surface defined by taking these dominant modes as variables is now examined, to gain insights into the ways in which they interact to form the orthorhombic phase. The simplest model for a single unstable mode which becomes stable after freezing in to a certain amplitude is a potential of a double well form. The energy in this model takes the form

$$E(c) = -\alpha c^2 + \beta c^4 \tag{5.3}$$

where α and β are positive constants and c is the amplitude of one phonon mode. $E(c)$ has a central maximum and a minimum on either side, as shown in Figure 5.6. When applied to, for example, the M_2 unstable mode in cubic perovskite, the cubic phase corresponds to the maximum at $c = 0$, but there are minima at $c' = \sqrt{\alpha/2\beta}$ which correspond to the metastable distorted phase described in Section 5.4.1. In the absence of phonon-phonon or phonon-strain coupling, the phonon coefficients c_0 in the orthorhombic phase (Table 5.5) would correspond to these minima, and the depths of the associated wells E_0 are estimated using this assumption. In the case of degenerate pairs, the combination found to actually contribute to the orthorhombic phase is used as a single mode for simplicity.

However, if two or more such phonons are present in a system, with amplitudes c_1 and c_2, there will in general be a coupling between them. In the simplest case this may

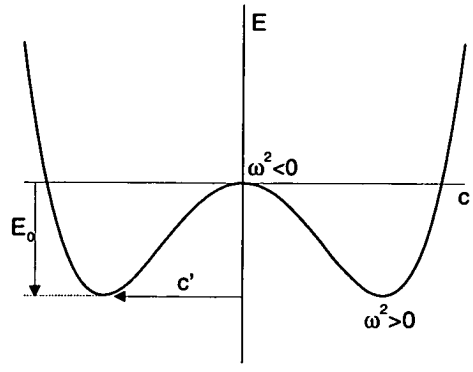


Figure 5.6. Double well model of energy as a function of phonon displacement as described by $E(c) = -\alpha c^2 + \beta c^4$. The cubic phase corresponds to the central maximum and the stable orthorhombic phase corresponds to one of the minima.

be represented by a quadratic coupling h:

$$\begin{aligned} E(c_1, c_2) &= E_1(c_1) + E_2(c_2) + E_h(c_1, c_2) \\ &= -\alpha_1 c_1^2 + \beta_1 c_1^4 - \alpha_2 c_2^2 + \beta_2 c_2^4 + h c_1^2 c_2^2 \end{aligned} \tag{5.4}$$

The minima of this system are not at (c_1', c_2') as defined above for the individual wells of (5.3), but depend on the coupling h. If the anharmonicities of the two modes are similar, i.e. $\beta_1 \simeq \beta_2$, and we take $\gamma_h = h/\sqrt{4\beta_1\beta_2}$, the minimum along c_1 occurs at

$$c_1^2 \simeq \frac{c_1'^2 - \gamma_h c_2'^2}{1 - \gamma_h^2} \tag{5.5}$$

so γ_h is identified as a dimensionless number indicating the strength of the coupling. For example, a negative h results in minima which are deeper and further from the origin than those of the one-phonon wells. At the origin (i.e. around the point of zero distortion) the frequencies of the normal modes are independent of any coupling, but at the saddle points corresponding to the equilibrium amplitude c' of one mode, the instability of the other mode is increased by negative coupling or decreased by positive coupling.

The coefficients c_0 obtained from the orthorhombic structure, using Equation (5.2), are those of the global minimum for all phonon modes, rather than the individual minima of the isolated wells. In general all the modes may be coupled at finite amplitudes. However, only those contributing to the orthorhombic phase with non-zero amplitude, as found in Section 5.6.1, need to be considered. In particular, the two octahedral rotation modes M_2 and R_{25} were studied.

To investigate phonon coupling in $MgSiO_3$, a single phonon may be frozen into the cubic phase; the *ab initio* energies and forces can then parametrise the model described in (5.3) to give equilibrium amplitude c'. The amplitude c_0 of each cubic phonon found in the orthorhombic structure can then be compared with the local c' for that mode in isolation.

The eigenvectors of the M_2 octahedral rotation remain unchanged as the amplitude increases. However, in the general case there may be coupling to other modes when the phonon amplitude is beyond the harmonic limit: for example, it was found in Section 5.4.2 that at finite amplitudes the R_{25} mode is coupled to a mode involving Mg motion along the x axis. Since we are principally concerned with the coupling between the two rotational modes, this Mg motion was included when finding the c' for the R_{25} mode.

Table 5.6 shows the results of this process for the M_2 and R_{25} phonons. The M_2 rotation was also combined with a change of the size and shape of the unit cell, such that the the octahedra remained undistorted. This simulates a rigid unit mode rather than a phonon, as will be discussed in Chapter 6.

We find that in the case of the M_2 rotation, the orthorhombic structure has a smaller rotation angle than that of the isolated phonon in equilibrium. The equilibrium amplitude of the isolated R_{25} rotation was harder to identify, since coupling with the Mg displacement mode was included, but was also smaller in the orthorhombic structure than in the isolated phonon. The coupling h between these modes, defined in (5.4), is therefore positive.

phonon	c'	c_0(orth)	c_0(cubic)
M_2	5.45 ± 0.02	4.35	4.14
M_2 + strain	5.1 ± 0.3	4.35	4.14
R_{25}	6.6 ± 0.4	6.48	6.18

Table 5.6. Normal mode coordinates of cubic rotational phonons in the fully relaxed orthorhombic phase, c_0(orth), ionic relaxation in a cubic cell, c_0(cubic), and of the equilibrium structure when the phonon is frozen in alone, c'. In both orthorhombic cases, the structure is not as distorted as the minima of the isolated phonons. Only the magnitudes of the coordinates are given.

Each of the octahedral rotation modes, when frozen in, thus stabilises the other. In both cases, this effect is even stronger when the unit cell is not relaxed, so this is a true phonon-phonon interaction rather than an interaction via the strain. This is in agreement with similar investigations using modified electron gas theory [88] but differs from the negative h value found in [86], in which it is not clear whether coupling with Mg distortion modes was considered. It is also consistent with the reduction in the instability of the M_2 mode when the R_{25} mode is frozen in, as described in Section 5.4.2.

We used the values of ω_R/ω_M, $c'(M)$, $c_0(R)$, and $c_0(M)$ to parametrise the model in (5.4), to give $\beta_1 = 0.01683$, $\alpha_2 = 1.13$, $\beta_2 = 0.01150$ and $h = 0.0085$ when α_1 is normalised to unity. The resulting well is shown in Figure 5.7, with annotations showing the points representing each of the three phases. The coupling strength $\gamma_h^2 \simeq 0.09$; the differences between the equilibrium amplitudes of isolated and coupled phonons are relatively small. When the c_0 values from relaxation in a cubic cell were used, the coupling increased to $\gamma_h^2 \simeq 0.15$, reflecting the finding that the normal mode coefficients were found to be even more reduced from their isolated values.

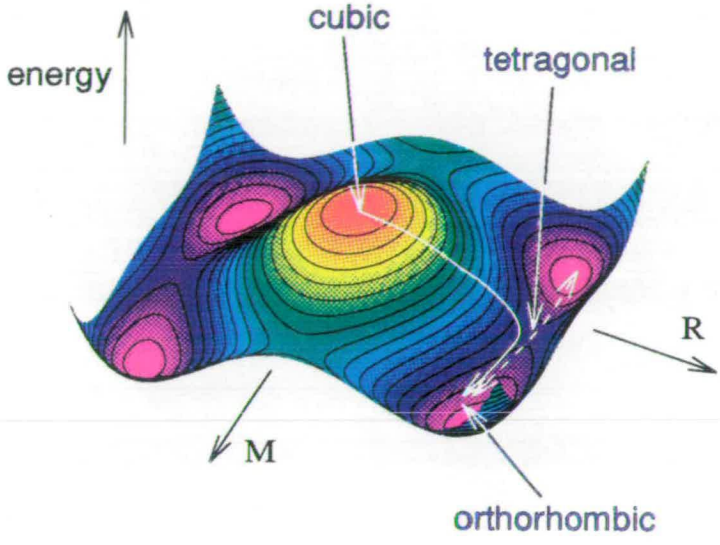


Figure 5.7. Coupled two mode system of (5.4) parametrised for the rotational modes from *ab initio* calculations. The points of the energy surface corresponding to the three perovskite phases are marked. The solid curve denotes a cubic→tetragonal→orthorhombic transition; the dotted curve represents motion at some temperature where the tetragonal phase could be observed. However, a cubic→orthorhombic transition could also occur directly via a route which would have the steepest gradient.

5.7 Pseudopotentials for magnesium and oxygen

Many of the very early applications of density functional theory and pseudopotentials concentrated on the properties of silicon, partly because of the tremendous theoretical effort which was aimed at understanding semiconductors. Both the $3s$ and $3p$ valence states of silicon are screened from the nucleus by the $n = 2$ shell, allowing a weak pseudopotential to be constructed. Modern silicon pseudopotentials typically require a basis set cutoff of only 150 eV. However, the problem of producing reliable pseudopotentials for magnesium and oxygen is somewhat harder: in this section some of the problems involved are demonstrated and attempts to overcome them are discussed. Fuller specifications of pseudopotentials are given in Appendix A.

5.7.1 Oxygen pseudopotentials

The $2p$ valence electrons of the oxygen atom are not screened from the nucleus by any core electrons with $l = 1$ angular momentum. The potential felt by the $2p$ electrons is thus very deep, and hence a fully transferable pseudopotential is likely to require a large basis set. An early potential [36] which was used in the work described in the first part of this Chapter, here labelled #1, attempted to overcome this problem by optimising the pseudopotential for the O^{2-} ion, effectively sacrificing some transferability in order to decrease the cutoff energy required to around 500 eV. A cutoff of 650 eV was however found to be required to ensure full convergence in $MgSiO_3$.

Although the bonding in $MgSiO_3$ is largely ionic, there is a finite amount of electronic charge distribution between the oxygens in $MgSiO_3$, forming the octahedral 'cage', as shown in Figure 5.5. This suggests that a pseudopotential optimised for the oxygen ion may not be entirely appropriate, and may explain why the equilibrium orthorhombic phase found above is considerably denser than observed experimentally or predicted with other pseudopotentials. More recently, an alternative pseudopotential (#2) has been developed using the Q_c tuning method (Section 2.3) by Lee *et al.*, which, although it requires a cutoff energy of 700 eV, uses the atomic reference states for some components, is less specialised and therefore probably more transferable.

5.7.2 Magnesium pseudopotentials

The Group II elements have valence electronic configurations containing only s states, so to generate non-local pseudopotentials for angular momentum components other than $l = 0$ requires the use of excited states. The standard reference configurations of BHS [38] prescribe a configuration of $3s^{0.5}3p^{0.25}3d^{0.25}$ (ionisation, here to Mg^+, is always necessary to ensure a bound excited state). However, this state is of high energy, and so is likely to be considerably different from typical solid-state configurations. The transferability of Kleinman-Bylander pseudopotentials may then be compromised (Section 2.3).

Recently it has been suggested that for some applications it may be better to use a fully local pseudopotential for magnesium [30], i.e. to assume that the potentials acting on any p or d components of the solid-state charge density are closer to that on the atomic valence s components, than to those acting on p or d components in the excited ionised atom. However, this approach was not found to produce satisfactory pseudopotentials for larger cations, such as calcium or barium: in this case it was found necessary to bring an extra shell of electrons from the core into the valence states: this is discussed for the case of Ba in Section 6.5.

The Mg pseudopotential used in the work described so far (#1) takes the Kleinman-Bylander non-local form, and was generated using the conventional BHS excited states [36]. However, a fully local potential (#3) has been found to give more satisfactory results [30] (when compared to, for example, all-electron calculations) and is evaluated below.

5.7.3 Effects on calculations of $MgSiO_3$ and MgO

Pseudopotentials are commonly evaluated via comparison of the logarithmic derivative (see Section 2.3) and solid-state tests of only one compound. However, the transferability is not completely established in this way. The equilibrium lattice parameters of cubic $MgSiO_3$ perovskite and MgO, and the stability of a displacement of the magnesium ions in $MgSiO_3$, were calculated using a number of different potentials. It should be noted that the stability of the normal mode dominated by Mg motion will not be quite the same as that of an isolated distortion. In addition to those potentials described above, a Mg pseudopotential with d local instead of s (#2), and a pseudopotential for oxygen with s, p and d non-local (#3) using a different Q_c tuning strategy [30] were also assessed. The results are shown in Figure 5.8.

The variation in equilibrium volume is significant. It may be seen that the choice of magnesium pseudopotential has the greatest effect on the volume of MgO, whilst the choice of that for oxygen affects mainly the volume of $MgSiO_3$. This effect alone shows

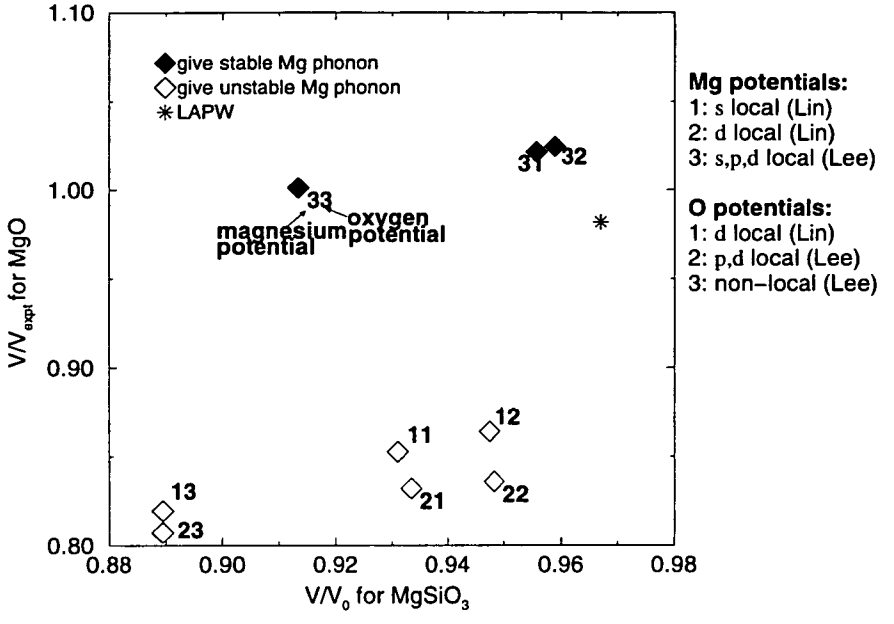


Figure 5.8. Variation in equilibrium volume of $MgSiO_3$ and MgO with choice of pseudopotential, and the stability of a Mg displacement in $MgSiO_3$. All simulations used a cutoff energy of 800 eV and a single **k**-point in the reduced Brillouin Zone. The reference volume for $MgSiO_3$ was taken as 43.8 Å, by assuming that the original calculations underestimate the volume in the cubic and orthorhombic phases by the same amount.

that testing these potentials on only one of these compounds is not sufficient to demonstrate reliability, even though they are both quite ionic. It may also be noted that some combinations of potentials overestimate the lattice parameter of MgO, in contrast to the usual expectation that LDA calculations underestimate the volume. Some explanation of the different effects on these structures may be gained from approximating the ions as hard spheres and identifying which directly determine the volume.

In $MgSiO_3$, the Mg ions occupy large 12-fold coordinated interstices in the structure, which are larger than required by the ionic radius of Mg. The SiO_6 cages thus determine the size of the structure, with the result that the oxygen pseudopotential has a strong effect on the equilibrium volume. The Mg potential has less effect on determining the volume, if the magnesium ion has a small radius. However, by using the fully local potential the Mg displacement is stabilised, and the volume is also increased,

suggesting that this potential gives a larger ionic radius. In this pseudopotential, the same potential is assigned to the p and d electrons as is used for the s electrons. Any p and d components thus effectively feel repulsion from $1s$ electrons, so would have higher energy than expected. The wavefunctions in the solid state are thus likely to be even more s-like, and may be enlarged since they are not constricted by p or d orbitals.

In contrast, both species are six-fold coordinated in MgO, and the Shannon ionic radii [105] suggest that the magnesium ions cannot fit into interstices between face-centred oxygen ions. The lattice parameter is thus determined by the sum of both radii, rather than being dominated by the oxygens as in $MgSiO_3$. However, as can be seen in Figure 5.8, the choice of magnesium pseudopotential produces the widest spread of equilibrium volumes. This again suggests that the effective ionic radii differ significantly over the three potentials used here.

The combination of the fully local magnesium potential (#3) and the p, d local oxygen potential (#2) can be seen to give the closest results to those obtained from all-electron (LAPW) calculations [86, 106]. LAPW results are expected to be affected only by the LDA, consistent with the slight underestimate of the lattice parameters found here. However, the original oxygen potential with d local (#1) gives similar results with this Mg potential. The equilibrium structure of orthorhombic $MgSiO_3$ was therefore recalculated using these potentials (by Karki *et al.*) [107]), and found to be only 0.2% smaller in volume than experiment. In the remainder of this work, these pseudopotentials were used. Since using the original oxygen potential has little effect on the volume, we expect the properties of $MgSiO_3$ which are dominated by the SiO_6 octahedra to be mainly unaffected.

The effect on the cubic phase is similar: the equilibrium lattice parameter using these pseudopotentials was found to be 3.49 Å, compared to 3.45 Å with the original set. The two highest phonon frequencies at Γ were found to change by less than 5%, and that of the Γ_{25} mode by 15% (1 THz). However, the greatest effect was found on the mode dominated by Mg displacement ($2.6i$ THz compared to $6.0i$), as expected.

The M_2 and R_{25} rotational phonons were also recalculated in the cubic phase, and found to be slightly less unstable (frequency change was less than 10%), following the trend of stability increasing with volume (Section 5.3.2). The energy difference between the tetragonal and orthorhombic phases was also found to decrease, consistent with the general decreased instability of the structure.

It is clear that testing pseudopotentials on one structure alone is not necessarily sufficient to demonstrate reliability. However, exhaustive testing on a number of different compounds whose structures are known experimentally is contrary to the aims of truly *ab initio* techniques. Traditionally, matching the logarithmic derivative with that of the isolated atom should ensure that scattering properties are reproduced, but this still only demonstrates that a pseudopotential is reasonably accurate over a limited range of energies. Alternative, more reliable methods of evaluating the performance and reliability of pseudopotentials without extensive solid-state testing are yet to be identified.

5.8 Molecular dynamics simulations

The investigations described in this Chapter have deduced the probable mechanism of any transitions between cubic, tetragonal and orthorhombic phases. However, since they simulate the zero temperature state, they do not predict at what temperature any transitions take place, and if the tetragonal phase does in fact form a distinct intermediate phase over any temperature range. In order to find the transition temperature, molecular dynamics simulations at finite temperature are required.

Traditionally, molecular dynamics simulations have used an empirical potential to calculate the forces between atoms in any given configuration. However, in 1985 the Hellmann-Feynman forces obtained from plane-wave pseudopotential calculations were first used instead [49], and *ab initio* MD has become an established technique [20]. Such simulations are particularly useful when the type of bonding involved is not precisely

known or when bonds are being broken and re-formed, because empirical potentials cannot adequately describe such situations. The large computational requirements of such simulations still mean that only modest numbers of atoms may be considered, rather than the $\mathcal{O}(10^6)$ which may be simulated using empirical potentials, but with increasing computational power such first-principles simulations are becoming more routine.

In order to follow the dynamics of the $MgSiO_3$ perovskite system at finite temperatures, the same supercell configuration as before was used, containing four formula units. Four **k**-points in the reduced Brillouin Zone were used, with a cutoff of 700 eV for the plane-wave basis set, and pseudopotentials as determined in Section 5.7 and used by Karki *et al.* [18, 107]. As previously explained, this cell has the correct geometry to contain the orthorhombic, tetragonal and cubic phases which have been described; it should thus be possible to detect any transitions between these phases.

The ions were initially assigned their positions in the tetragonal phase, with the cell vectors (a=b=$\sqrt{2}a_0$,$c = 2a_0$) of the cubic phase. In the phonon calculations described above, no normal modes were found with frequencies above 30 THz, i.e. with period less than 33 fs. A timestep of 1.0 fs was thus used, so that even the highest frequency modes should be accurately represented. All ions were initially given a random velocity according to a prescribed temperature, and the equations of motion integrated forwards in time with a Verlet algorithm. The lattice vectors were not changed in response to the stress during the simulations as this would have required much larger amounts of computer time.

After an initial period of equilibration, the temperature may be measured from the time-averaged kinetic energy. If the atoms start from their equilibrium positions, equipartition would be expected[1] (i.e. actual temperature is half of that prescribed), but any phase transitions during the simulation may release additional energy. It is

[1] It may be noted that if the centre of mass is constrained to be at rest, the number of degrees of freedom is only $3N - 3$ rather than $3N$, affecting the application of the theory of equipartition.

expected that the tetragonal–cubic phase transition would require a temperature above the melting point (T_M is about 1800 K at 0 GPa), so instead we are interested in the orthorhombic–tetragonal phase transition, with some transition temperature T_c.

The CETEP code was used on the T3D parallel supercomputer at the Edinburgh Parallel Computing Centre. These simulations require extremely large amounts of CPU time, so it was not possible to perform as many iterations as might be desired at each temperature to produce highly accurate statistics. However, the results presented here should be sufficient to identify the progress of the transition and bound the transition temperature.

5.8.1 Following transitions using normal modes

As described previously in this Chapter, each metastable phase can be described in terms of freezing in ionic displacements, following the eigenvectors of selected phonons of the cubic phase. All modes are referred to by their location in the cubic Brillouin Zone and rank according to frequency (from §1 to §60). According to this scheme, one combination of each of the degenerate pairs §47/48 (Mg displacement) and §59/60 (R_{25} SiO_6 rotation) form the tetragonal phase. To then form the orthorhombic phase requires contributions from §13, §57 (M_2 rotation) and one combination each of §30/31 and §50/51, all at X or M in the cubic phase. The contributions and nature of these six phonons were summarised in Table 5.5.

The state of the system during molecular dynamics simulations is therefore described in terms of the average coefficients $\{\langle c_j(t)\rangle_t\}$ of the cubic normal modes rather than the positions of the ions $\{\mathbf{R}_{l,\kappa}(t)\}$. This is done in a similar way to that described in Section 5.6, by projecting the displacement from the cubic phase onto each cubic phonon eigenvector in turn. A total distortion $|\mathbf{d}|^2$ was also calculated, using (5.2),

$$|\mathbf{d}(t)|^2 = \sum_j c_j^2(t) = \sum_{l,\kappa} m_\kappa [\mathbf{R}_{l,\kappa}(t) - \mathbf{R}_{l,\kappa}(\text{cubic})]^2 \qquad (5.6)$$

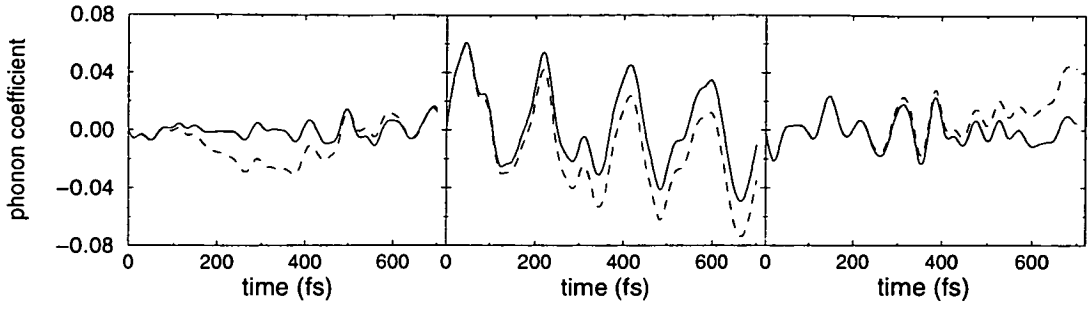


Figure 5.9. Coefficients (in a.m.u.Å) of the three translational phonons during a MD run before (dashed line) and after (solid line) correction for the drift of the centre of mass is applied.

which when time-averaged indicates the average deviation from the cubic phase. The term 'coefficient' is used in preference to 'amplitude' to emphasize that it is the time-averaged displacement which is important, rather than the amplitude of oscillations about it.

If all cubic normal modes were also normal modes of the tetragonal and orthorhombic phases, perfectly harmonic oscillation would be observed. However, as described already, some mixing between modes occurs, as well as degeneracy breaking; the coefficients $c_j(t)$ of cubic phonons will vary non-sinusoidally for any modes which deviate significantly from normal modes of the orthorhombic phase. However, most modes are close enough to lead to useful analysis, and so the cubic phonons are retained as reference modes for consistency with other sections of this work. The mode coefficients are thus taken as order parameters for modes expected to participate in the transition.

The random initial velocities assigned to the ions are chosen such that the centre of mass is stationary, but a small amount of motion does develop due to the use of a finite timestep, which causes imperfect integration of the equations of motion. Such drifting may be expressed as a non-zero coefficient of the three translational phonon modes §43, §44 and §45. A correction was thus made by subtracting the displacement of the centre of mass before projecting the atomic displacements onto the phonon eigenvectors; this

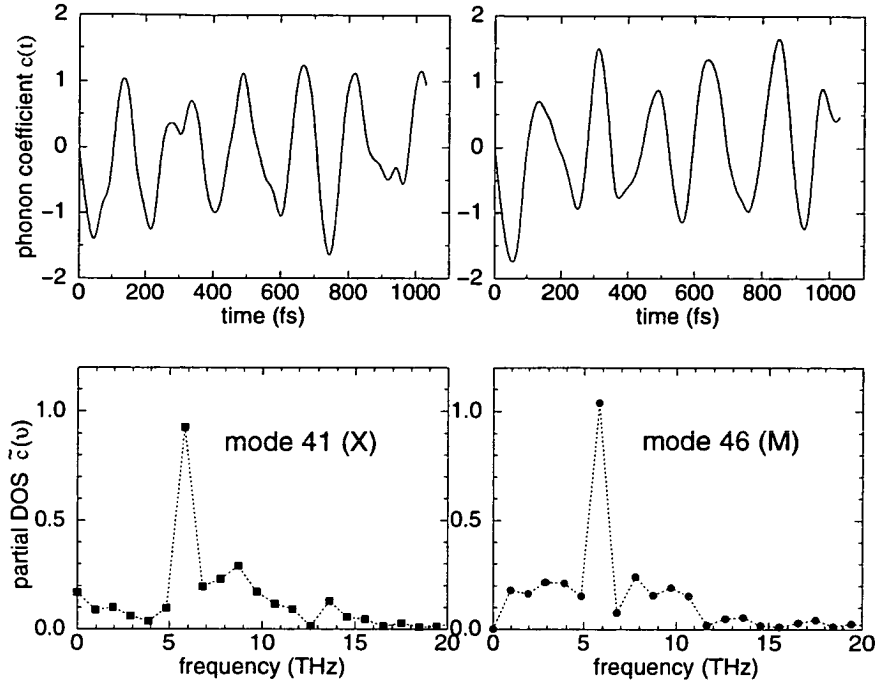


Figure 5.10. Phonon coefficients $c_j(t)$ and partial density of states (Fourier transform) $\tilde{c}_j(\nu)$ for two modes which give the clearest single frequencies in a MD simulation around 850 K. Both modes involve Mg displacement along the z axis. The plotted points demonstrate the maximum resolution available from a 1 ps simulation.

appears to only slightly decrease the amplitude of oscillation but recovers zero time-averaged coefficients (Figure 5.9). The oscillations observed in the translational modes must therefore be due to noise and breaking of orthogonality in the eigenvectors, so using a smaller timestep would not produce significantly improvement. However, these oscillations are much smaller than the $\langle c_j(t)\rangle_t$ of the phonons we wish to follow, and the effect of this procedure on the non-translational modes is negligible.

In principle, the Fourier transform of each coefficient $c_j(t)$ will generate a partial density of states (DOS) $\tilde{c}_j(\nu)$, which for a perfectly harmonic mode would show a single frequency. However, a MD simulation of time τ will give a frequency resolution of $1/\tau$, typically 1 THz in these simulations. Furthermore, few of the phonons are perfectly harmonic, and the changes in the structure during the simulation introduce other Fourier components and will change the frequencies of some modes. The longest

simulation was done at a temperature of around 850 K, and the two phonons found to give the clearest partial DOS are shown in Figure 5.10 by way of example. These modes are thus the least affected by the transition to the orthorhombic phase which takes place during the simulation (described in Section 5.8.3). Both modes involve Mg displacements along z, so would be expected to be only weakly coupled to those involved in the transition which are in the xy plane. The change in frequency for these modes between the tetragonal and orthorhombic phases is less than 0.3 THz. Longer simulations would, however, be required for useful results from such a process.

5.8.2 Low temperature

At temperatures below T_c, the orthorhombic phase is expected to be stable; this requires the freezing in of the four modes listed above, which account for the additional structural parameters of the orthorhombic phase.

An initial temperature of 800 K was assigned, but although the system first appeared to equilibrate around 430±5 K (as might be expected by equipartition), after 300 fs the temperature increased to 550±5 K. This was accompanied by a distinct and significant increase in the total distortion $\langle|\mathbf{d}|^2\rangle_t$ from 40.82±0.09 to 52.0±0.5 a.m.u.Å^2, as shown in Figure 5.11.

The coefficients of selected phonons during this simulation are shown in Figure 5.12. The R modes which form the tetragonal phase, §59/60 and §47/48, have significant coefficients throughout as would be expected. It can be seen that during the first half of the simulation, the third R_{25} rotational phonon §58 (which is still unstable in the tetragonal phase) acquires a large coefficient, whilst the M_2 mode §57 is present only weakly. However, after 300 fs the M_2 mode acquires a large average coefficient, and oscillates around that value without crossing $c(t) = 0$, whilst mode §58 returns to small oscillations about zero. Accompanying this change are increases in the coefficients of §13, §30/31 and §50/51, i.e. those known to be present in the orthorhombic phase. The system is thus identified as orthorhombic in this part of the simulation, and this

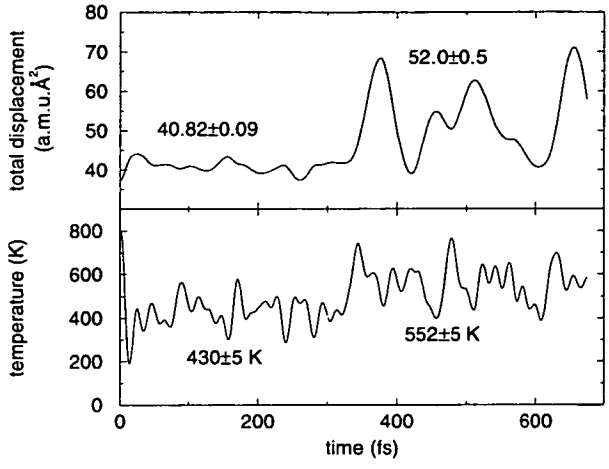


Figure 5.11. Total distortion from cubic $|\mathbf{d}(t)|^2$ (upper graph) and temperature as found from kinetic energy (lower) during a simulation at low temperature.

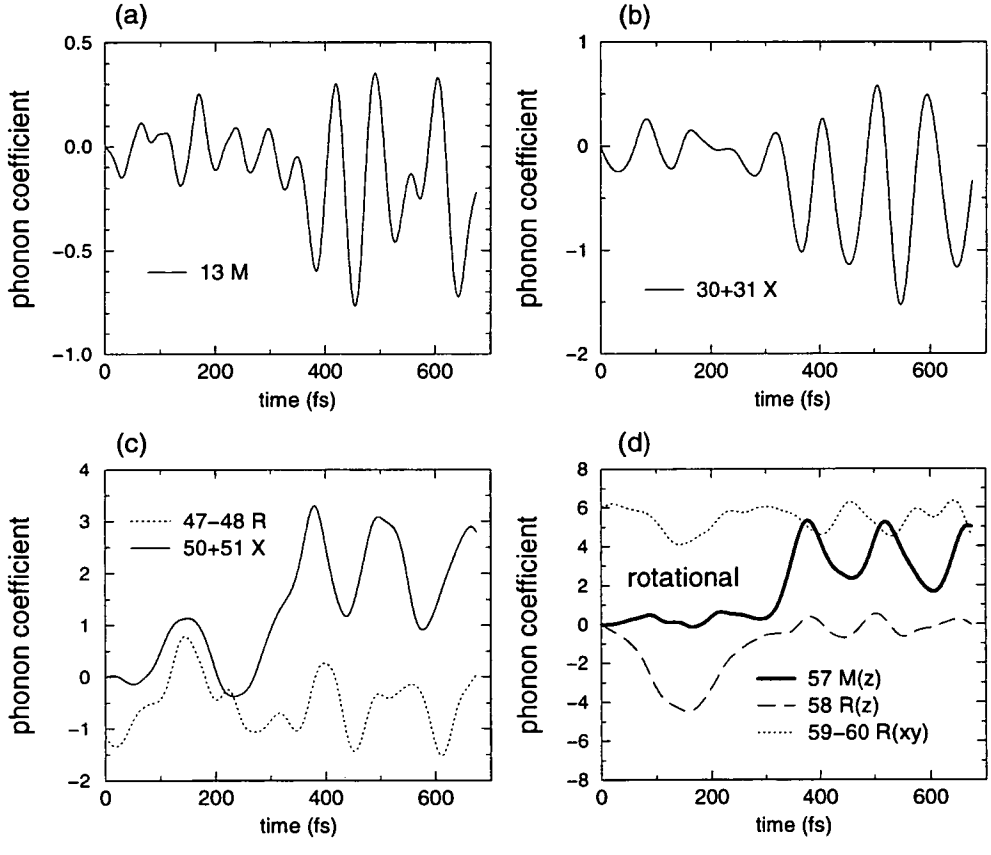


Figure 5.12. Phonon coefficients (in a.m.u. Å^2) for all modes with significant $\langle c(t)\rangle_t$ during the low-temperature MD simulation. The modes in (d) primarily involve rotations of the SiO_6 octahedra.

is expected to be the equilibrium phase. The transition into the orthorhombic phase releases energy, explaining the increase in temperature, and is also responsible for the increase in total distortion.

The M_2 phonon §57 as shown here has a frequency of approximately 6.9 THz. This is about half of the frequency calculated in the orthorhombic phase; the difference may be ascribed to anharmonicity and coupling, since the oscillations are by no means small. Although different pseudopotentials were also used in the initial calculations, the frequencies of the rotational phonons in the cubic phases were not found to change by more than a few percent.

5.8.3 Intermediate temperatures

Figure 5.13 shows the total distortion and temperature for a simulation at an intermediate temperature; the system appeared to equilibrate around 850 K after about 300 fs. The dominant phonons are shown in Figure 5.14. It can be seen that the M_2 mode §57 freezes in but almost immediately changes sign, as does §50/51 and to a less dramatic extent, §30/31 and §13. The coupling between these four modes is evident from the simultaneous switching of sign. This jump from one of the orthorhombic phases to an equivalent demonstrates the increasing thermal energy invested in these modes. In contrast, the modes at R which constitute the tetragonal phase, §47/48, §58 and §59/60, are mostly unaffected by these changes.

At a higher temperature, the third R_{25} phonon was observed to play more than a momentary role in the dynamics. This simulation equilibrated at 1440 K (Figure 5.15). The total distortion was larger than in any of the previous simulations (around 57 compared to typically 52 a.m.u.Å^3), but the phonon coefficients show that the structure should still be considered to be orthorhombic. As the temperature increases towards T_c, the orthorhombic distortions, and thus also $|\mathbf{d}(t)|^2$, would be assumed to slowly decrease, but the increased amplitude of oscillations at a high temperature may be responsible for the larger time-averaged total distortion.

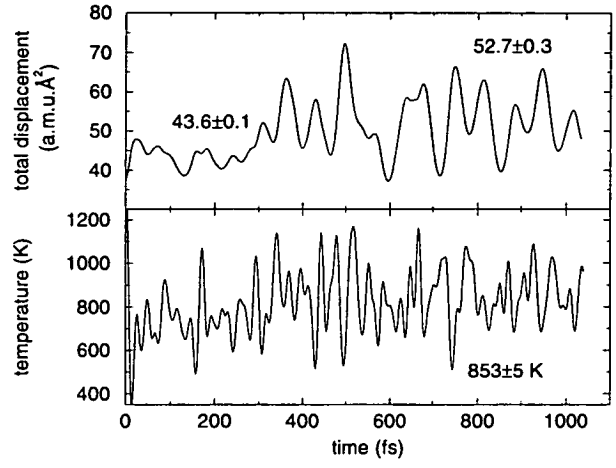


Figure 5.13. Total distortion from cubic (upper) and temperature as calculated from kinetic energy (lower) during the MD simulation at intermediate temperature.

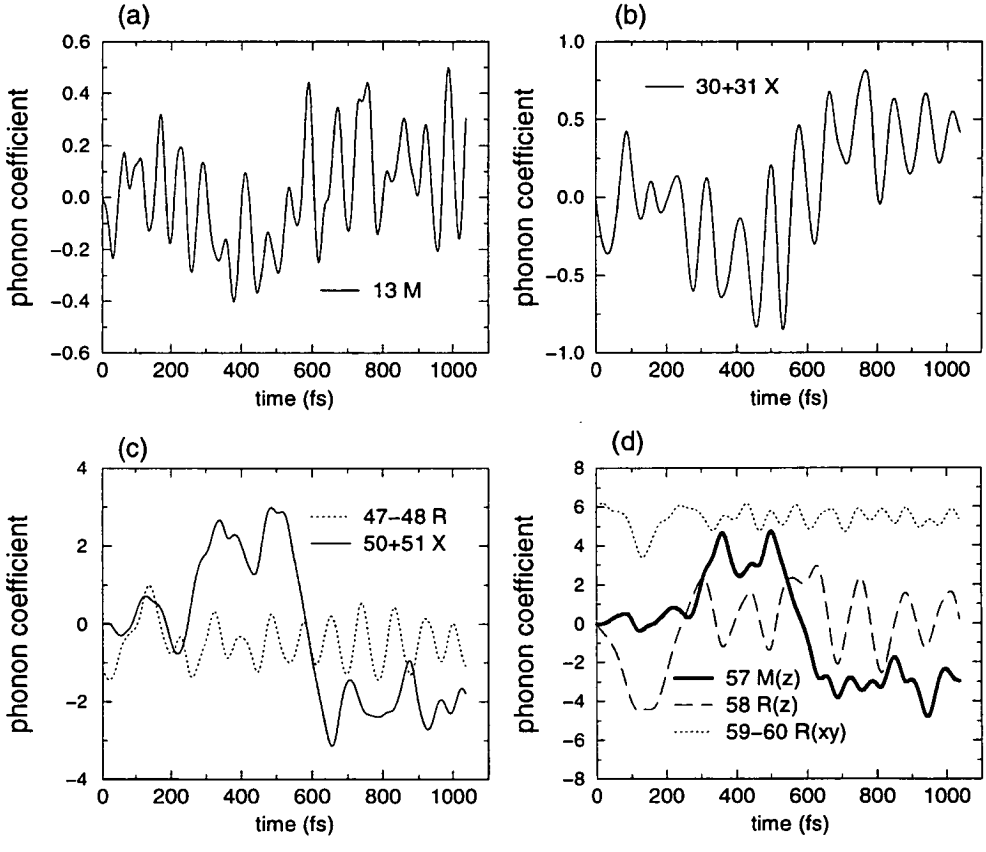


Figure 5.14. Phonon coefficients (in a.m.u.Å^2) for all modes with significant average amplitude during the intermediate MD simulation. The modes in (d) primarily involve rotations of the SiO_6 octahedra.

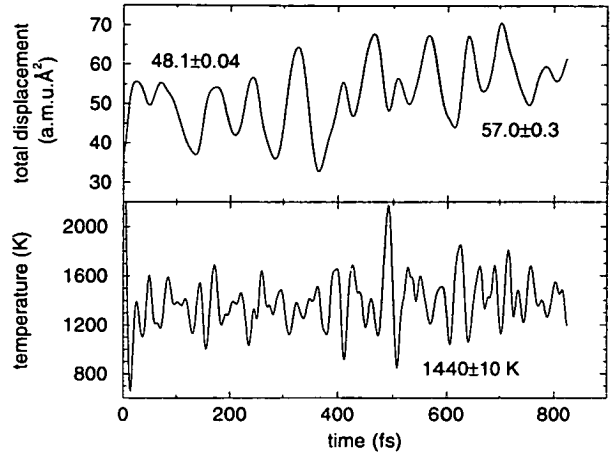


Figure 5.15. Total distortion from cubic (upper) and temperature as calculated from kinetic energy (lower) during the higher intermediate MD simulation.

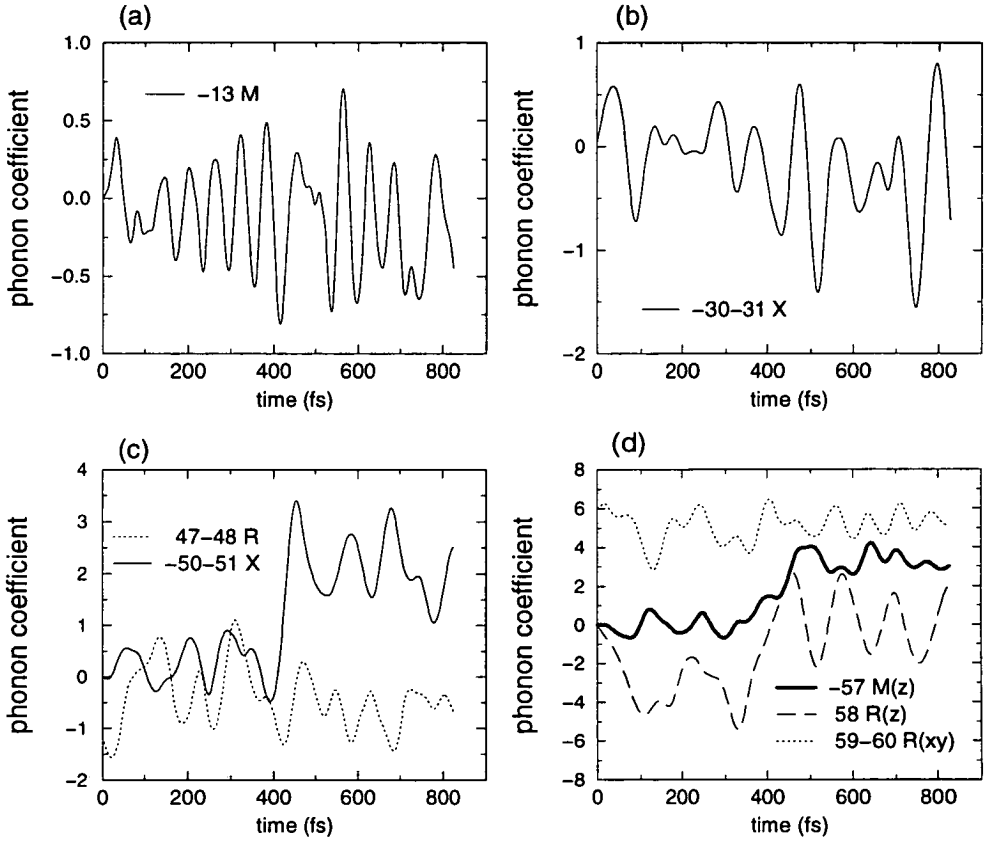


Figure 5.16. Phonon coefficients (in a.m.u. Å^2) for all significant modes during the higher intermediate MD simulation. The modes in (d) primarily involve rotations of the SiO_6 octahedra.

The phonons which form the orthorhombic phase from the cubic (all except §59/60 and §47/48) were actually observed to have coefficients of opposite sign to previous simulations, but since the eigenvectors have arbitrary sign, the sign of the coefficients may be changed freely. This has been done before plotting Figure 5.16. As in the other simulations, the equilibrium system contains significant amplitude of §57 and other expected phonons. It can be seen that the third R_{25} phonon (§58) undergoes two complete periods of oscillation before reverting to an average coefficient of zero, but thereafter still has a large amplitude of oscillation. However, the structure must be close to the orthorhombic phase for this phonon to be stable as observed.

5.8.4 High temperature

At 1750 K, a significant contribution from §57 and related modes is seen (Figures 5.17 and 5.18), indicating that the system is still orthorhombic. The transition temperature must thus be still higher than 1750 K, and thus above the melting temperature at zero pressure.

However, the metastable tetragonal phase is observed for the first 650 ps of the simulation, and includes significant contributions from mode §58. Since modes §58,§59 and §60 are triply degenerate in the cubic phase, and §58 is almost as unstable as §57 in the tetragonal, this is understandable. It can be seen that the coefficient of §58 increases at the expense of §59/60, so only two of these modes constitute the tetragonal phase. The structure formed by freezing in all three of these modes also has lower symmetry than the tetragonal, so would be considered as an alternative to the orthorhombic phase, rather than as a possible intermediate towards the cubic.

5.8.5 Implications for T_c

It appears that at zero pressure an orthorhombic–tetragonal transition would require a higher temperature than the melting temperature. If coupling to the strain is allowed, the energy difference between these phases increases, thus further increasing T_c. At

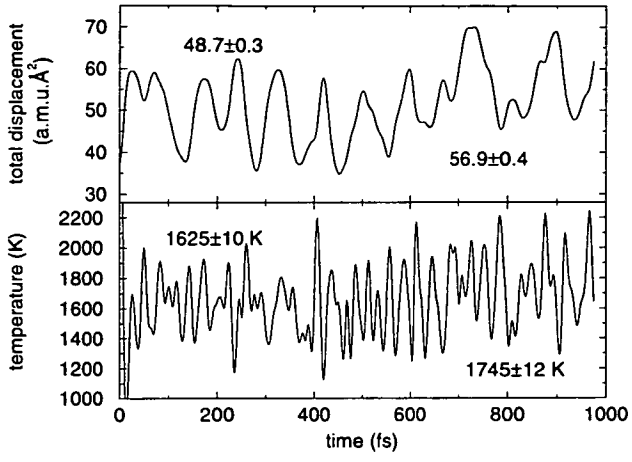


Figure 5.17. Total distortion from cubic (upper) and temperature as calculated from kinetic energy (lower) during the MD simulation at high temperature.

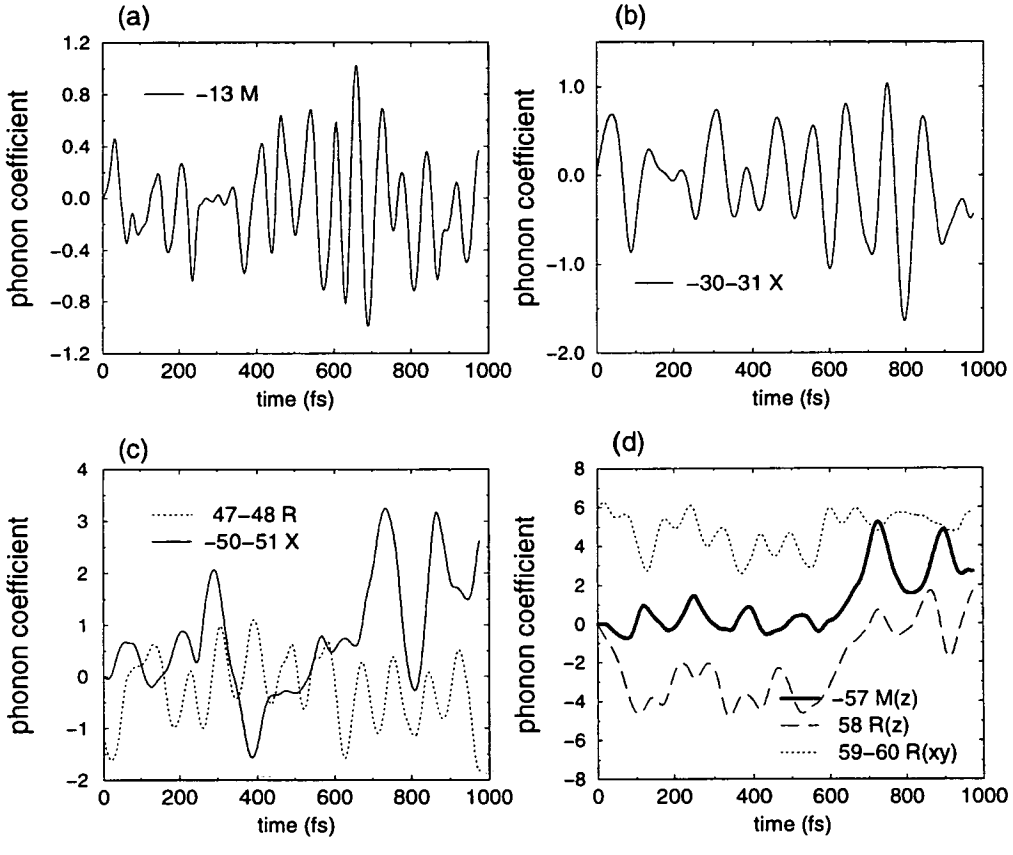


Figure 5.18. Phonon coefficients (in a.m.u. Å^2) for all significant modes during the high temperature simulation. The modes in (d) primarily involve rotations of the SiO_6 octahedra.

higher pressures, such as those found in the mantle, the denser orthorhombic phase would be increasingly favoured, although the melting temperature also increases. These MD studies thus suggest that the transition does not occur in the mantle.

However, these simulations are not only limited by neglecting coupling to the strain, but also by the simulation of only twenty atoms. Phase transitions usually involve long-range fluctuations, which are explicitly excluded by periodic boundary conditions. Furthermore, the finite duration of these simulations, necessary because of the extreme levels of computational resources required, means that the long-term behaviour may not be adequately represented. More extensive simulations would therefore improve the reliability of these predictions.

5.9 Conclusions

The structures and important sets of phonons of cubic, tetragonal and orthorhombic phases of $MgSiO_3$ have been found from first principles. The unstable phonons of the cubic structure are distributed throughout the Brillouin Zone, but all become stable after the transition to the orthorhombic structure.

By freezing in the most unstable phonon of the cubic phase, and another strongly coupled to it, a tetragonal intermediate may be formed, which displays some of the structural features of the orthorhombic phase. The eigenvectors of the most unstable phonon of this phase, and those coupled to it, contain the remaining distortions necessary to reach the orthorhombic phase. Rotations of octahedra play the largest part in the transition, and are the most unstable.

A two stage transition pathway between the cubic and orthorhombic phases may be expressed in terms of only six phonons, and small changes in the unit cell. Molecular dynamics simulations of one of these stages are therefore analysed in terms of these phonons, with the result that a transformation from a metastable tetragonal phase to the orthorhombic is clearly identifiable. However, this was always observed for

temperatures up to 1750 K, which is close to the melting temperature. These results suggest that the transition temperature for the tetragonal–orthorhombic transition is thus likely to be too high for the tetragonal phase to occur in the mantle. Further study would be necessary on larger systems for longer times to confirm this prediction.

Chapter 6

Rigid Unit Modes

There are, as I have said, some minds which can go on contemplating with satisfaction pure quantities presented to the eye by symbols, and to the mind in a form which none but mathematicians can conceive.

There are others who feel more enjoyment in following geometrical forms, which they draw on paper, or build up in the empty space before them.

James Clerk Maxwell, *Scientific Papers* 1890/1891

6.1 Introduction

Many minerals contain four- or six-fold coordinated sites, which are often drawn, for the sake of simplicity, as tetrahedral or octahedral 'coordination polyhedra' respectively. The description of certain structures, for example quartz and perovskites, in terms of coordination polyhedra has been widespread for over thirty years [90]. The advantages of picturing structures in this way is obvious: in many cases it makes the essential features much clearer and greatly reduces the number of degrees of freedom which need to be considered. When a number of minerals have similar structures (as is the case with the perovskite structure), reducing the number of parameters needed to describe each structure may draw attention to trends which affect a number of compounds.

However, it may be that these constructions play a more than a merely pictorial role in describing the dynamics and phase transitions of such systems. The concept of rigid

units having implications for the dynamics of such systems, rather than acting merely as convenient ways to describe complex structures, is somewhat more recent [108]. The *rigid unit mode* approximation assumes that such polyhedra behave as if it were not possible to distort them. In this methodology, coordination polyhedra are treated as perfectly rigid, and may not be distorted; the consequences of this approximation are then deduced. In particular, by finding any vibrational modes which are allowed under this constraint it is hoped to determine related structures, between which there may of course be phase transitions.

Rigid unit mode (RUM) analysis was first performed for the β phase of quartz [108, 109] and explained the phase transition between α and β phases by a zone-centre RUM, and an observed incommensurate phase transition by RUMs along a whole line in reciprocal space. Recently a computational algorithm has been developed [99] to detect possible rigid units modes in a given structure: this will be described and used in this Chapter. This approach can be applied to a number of minerals [110] and, for example, has successfully explained diffuse scattering seen in X-ray diffraction experiments [109].

In this Chapter the main features of rigid unit modes are summarised, and the ways in which the *ab initio* techniques described in previous Chapters may be used to investigate the validity of the rigid unit mode approximation are discussed. In some systems, the RUM model may be only useful to qualitatively describe the general features of a structure or a phase transition, but in other cases more quantitative predictions may be made.

The perovskite structure of ABX_3 type minerals, which was described in Chapter 5, is usually described in terms of corner-sharing BX_6 octahedra. These are perfect octahedra in the cubic perovskite ($Pm3m$) structure, but have often been found to be slightly distorted in lower-symmetry structures. The various space groups which could be formed by tilting the octahedra were categorised by Glazer [111] and are discussed thoroughly in a useful review [91]. The unstable rotational modes discussed in Chapter 5 consist of just such rotation (or tilting) and are good candidates for behaving as

rigid unit modes; we discuss below the evidence for this approximation.

Recently it was reported that ZrW_2O_8 has a negative thermal coefficient of expansion, which is unusual since cubic symmetry was preserved, indicating that this behaviour is isotropic. This is interpreted [14, 16] as being due to the presence of rigid unit modes which, as will be shown in this Chapter, decrease the volume on excitation. However, this explanation relies on the stiffness of the rigid units in the structure, which is assumed in order to invoke the presence of RUMs. ZrW_2O_8 is made up of WO_4 tetrahedra and ZrO_6 octahedra, which are corner-linked in an open framework structure. Barium zirconate perovskite also contains ZrO_6 octahedra, so investigations of this much simpler structure were carried out to estimate the stiffness of these units, using insights gained from the study of $MgSiO_3$.

Silica (SiO_2) exhibits several phase transitions under pressure. The structure may be considered as being composed of rather distorted SiO_6 octahedra, linked by both corner- and edge-sharing. However, previously it has only briefly been discussed in terms of rigid units [112, 113]. The structural details of some of these transitions are examined with a view to determining whether any of them bear the hallmarks of rigid unit modes.

6.2 Signature features of RUMs

Rigid unit modes have several identifying features, which are described here so that they may be used to determine whether the behaviour of a given structure is dominated by RUMs. RUMs may affect the behaviour of one phase under pressure or temperature, and also explain the mechanisms of soft-mode phase transitions which involve RUMs. The role of low-frequency and unstable phonons in such transitions is described in Chapters 4, 5 and 7; here we examine the eigenvectors of the important modes in more detail. The methods by which features related to RUMs may be detected are also outlined.

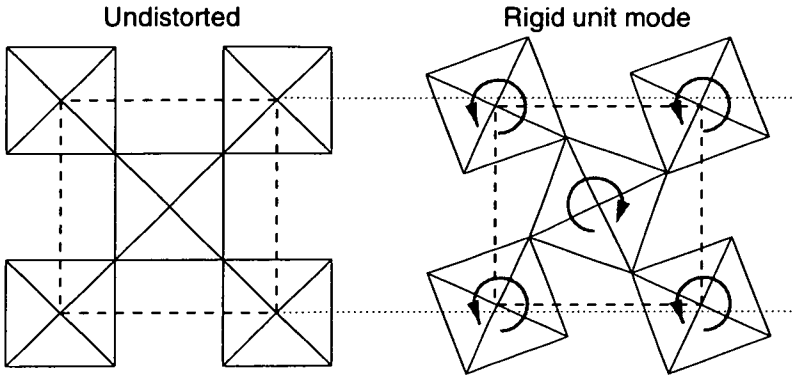


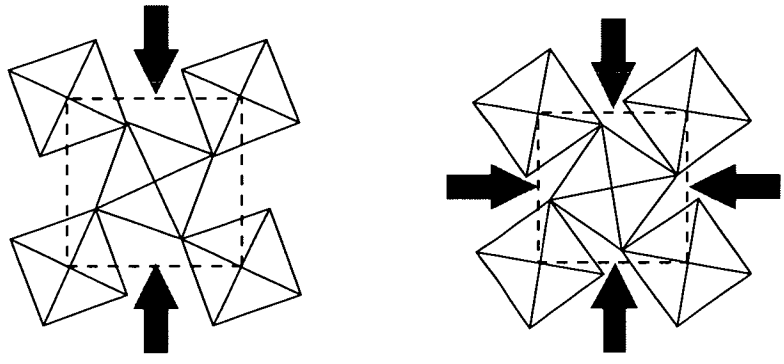
Figure 6.1. Illustration of the M_2 rigid unit mode in cubic perovskite, in which the size of the octahedra is maintained, but consequently the unit cell must contract along the axes perpendicular to the rotation axis (drawn perpendicular to the page).

6.2.1 Preservation of bond lengths and bond angles

By definition, all bondlengths and bond angles of the rigid units must be preserved in RUMs; the volumes of individual octahedra (or tetrahedra) must thus also be preserved. Bondlengths and bond angles may easily be determined from *ab initio* structural determinations, in different phases between which a RUM phase transition is suspected, or in one phase at a variety of pressures. Some authors have relaxed the strict definition of a rigid unit, to instead include modes in which the bondlengths (and volume) of polyhedra change, but bond angles are relatively unaffected, so that the shape of the polyhedra is conserved [114, 115]. However, here the definition is restricted to only incompressible and undistortable polyhedra.

6.2.2 Decrease in cell volume

If rigid units rotate as RUMs, continuing to share atoms at the corners, the unit cell vectors perpendicular to the rotation axis must change in length, as shown in Figure 6.1. Compared to the phase with highest symmetry, in which there is zero amplitude of such a mode, the volume must decrease on excitation of the mode [116]. Since in the ideal case the rigid units remain undistorted and uncompressed, the fraction of the unit cell



(a) (b)

Figure 6.2. Two methods of compression in the cubic perovskite: (a) permitting solely uniaxial strain; (b) strain along both xx and yy, which allows coupling to the the M_2 rigid unit mode (rotation axis drawn perpendicular to the page as in Figure 6.1).

filled by the rigid units must thus increase. In the simple case illustrated in Figure 6.1, for example, $a = b = a_0 \cos\theta$ for rotation angle θ about the c axis, giving a second-order decrease in lattice vector with θ; the c/a ratio of lattice parameters would therefore also increase. It will be shown in Section 6.6.1 that in some systems there is a first order coupling.

6.2.3 Elastic moduli

The RUM illustrated in Figure 6.1 permits a very low-energy compression mechanism for lattice parameters perpendicular to the rotation axis, assuming that bond-bending between octahedra requires little energy. In the limit that the 'hinges' between rigid units have zero bond-bending spring constants, (but that units may not overlap) the entire structure could collapse to half its initial volume without energy cost. Figure 6.2 shows two strains applied to this system: when only $\epsilon_{yy} = \epsilon$ is non-zero, the lattice vectors do not permit the RUM to freeze in, giving

$$E = \frac{1}{2} c_{11} \epsilon^2 > 0 \tag{6.1}$$

and hence $c_{11} > 0$. However, if $\epsilon_{xx} = \epsilon_{yy} = \epsilon$, then the mode can allow compression

without any input of energy, i.e.

$$E = \frac{1}{2} \begin{pmatrix} \epsilon & \epsilon \end{pmatrix} \begin{pmatrix} c_{11} & c_{12} \\ c_{12} & c_{11} \end{pmatrix} \begin{pmatrix} \epsilon \\ \epsilon \end{pmatrix} = (c_{11} + c_{12})\epsilon^2 = 0 \quad (6.2)$$

$$\Rightarrow \quad c_{12} < 0 \quad (6.3)$$

which explicitly demonstrates that there is an elastic instability. Although this may only happen when the mode is excited to some finite amplitude (isotropic pressure on the high-symmetry structure could only result in compression of individual bonds), at any non-zero temperature this will be the case.

In real systems, the bond-bending force constants, though finite, will usually be much lower than for the bond-stretching required to distort polyhedra. For a RUM model to be applicable, bond-bending between units must require rather lower energy than distortion of bond angles within a polyhedron. A compression mechanism which involves only inter-unit bond-bending will then still have very low energy, corresponding to at least one rather small elastic modulus.

6.2.4 Differences between RUMs and harmonic phonons

Rigid unit modes are often described as 'phonons', but strictly they do not describe the same motion. Harmonic phonons are conventionally written in terms of atomic displacements at constant volume. The harmonic theory of lattice vibrations expresses all interactions in terms of a harmonic energy, with force constants Φ as given in Chapter 3; the atoms vibrate around a set of mean positions which minimise the energy, thus fixing interatomic distances and the lattice parameters. Thermal expansion requires an anharmonic energy function, and is thus incompatible with purely harmonic phonon theory; the specific heat may be treated harmonically, but also requires anharmonicity for a full explanation. In practice, the phonon frequencies may be determined at a range of volumes (corresponding to a range of pressures), for vibrations around mean positions which may themselves change with volume. The variation of frequencies with

volume is expressed by the Grüneisen parameters, which may be used to used to determine thermal expansion. However, the important point for the present discussion is that phonon eigenvectors are restricted to motion at constant volume.

Pure rigid unit modes, however, preserve the size and shape of the polyhedra, so the unit cell must change as illustrated in Figure 6.1. Freezing in a pure rigid unit mode should have no effect on the polyhedral volume but would involve shrinking of the unit cell; a pure phonon mode would preserve the cell volume but would enlarge the polyhedra. Phonons and RUMs may thus be seen to be at opposite extremes of strain-mode coupling.

6.3 RUM analysis with CRUSH

The CRUSH package [99, 117] has been developed by Dove *et al.* to detect RUMs in a given structure. Coordination polyhedra are assumed to behave as rigid units—that is, all atoms in a given polyhedron remain in fixed relation to each other, so the only variables considered are the rigid translation of each unit and its three rotation angles. Shared atoms which link two polyhedra are replaced by pairs of 'half atoms', one on each of the two vertices which should be linked. Each pair of split atoms is then joined by a spring, which is the only force constant in the model (Figure 6.3). The dynamical matrix is formed and hence the normal modes found [118]. A distortion which allows all split atom pairs to remain joined thus has zero energy, and so a normal mode which similarly preserves the linkage between rigid units has zero frequency: such a mode is exactly a rigid unit mode as defined by Dove *et al.* Those modes in a real system which involve distortions of the octahedra (e.g. 'breathing' or 'squashing', as shown in Figure 6.3) are mapped by CRUSH to a smaller number of modes in which the linking atoms are split, and which thus have a non-zero frequency: the value of this frequency is only a very rough indication of the related frequencies in the real system.

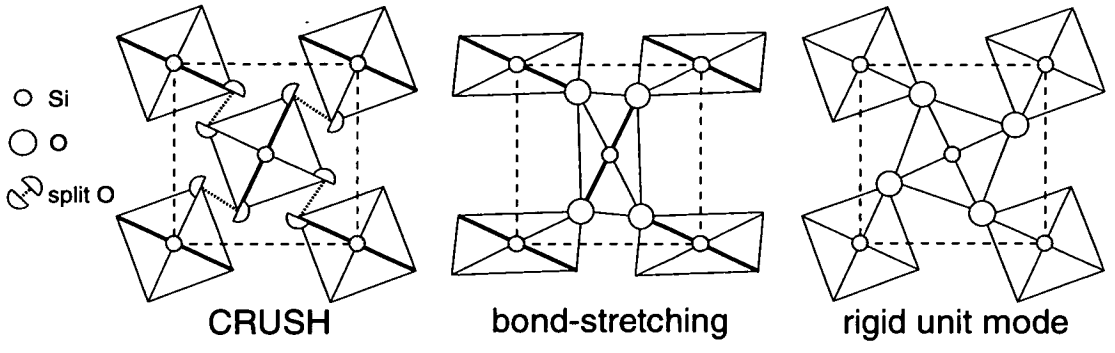


Figure 6.3. Representation of structural distortions in CRUSH and a more realistic model. The first two diagrams show a mode where all octahedra rotate through the same angle, as defined by the heavy lines. In CRUSH this necessitates separation of split atoms since the octahedra are undistorted, giving a non-zero frequency; but in a real system the octahedra would distort in order to remain corner-sharing. The third diagram shows the rigid unit mode in which (to a harmonic approximation) the oxygens do not need to split, thus giving a zero frequency in CRUSH.

However, the strain is not included as a variable, despite coupling to the strain being an important feature of many RUMs, as discussed in Section 6.2.2. The consequences of this omission, and whether the CRUSH approach is still useful, are discussed here. In a simple rotational mode such as that shown in Figure 6.1, the energy can be written in terms of the lattice parameter and rotation angle (taken as amplitude of the mode) as $E(\theta, a)$. The effect of exciting such a mode must then take account of any change of a with θ:

$$\begin{aligned} \frac{dE}{d\theta} &= \frac{\partial E}{\partial \theta} + \frac{\partial E}{\partial a}\frac{da}{d\theta} \\ \Rightarrow \frac{d^2E}{d\theta^2} &= \frac{\partial^2 E}{\partial \theta^2} + 2\frac{\partial^2 E}{\partial \theta\, \partial a}\frac{da}{d\theta} + \frac{\partial^2 E}{\partial a^2}\left(\frac{da}{d\theta}\right)^2 + \frac{\partial E}{\partial a}\frac{d^2a}{d\theta^2}. \end{aligned} \tag{6.4}$$

The frequency of a mode should be calculated from the full second derivative $d^2E/d\theta^2$ on the left hand side of (6.4), and will depend on the coupling between the cell and the rotation, via $a(\theta)$. In a rigid unit mode for which a zero frequency is expected, this full derivative should be zero. A pure phonon, however, would have $a(\theta) = a_0$, so that only the first term on the right hand side, the partial second derivative, contributes to

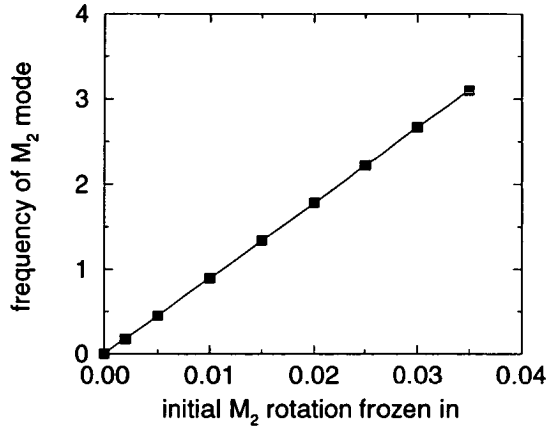


Figure 6.4. Frequency of the M_2 rigid unit mode calculated using the CRUSH program, in a perovskite structure which already contains some amplitude of that mode. Only the rotation around the undistorted structure is found to have zero frequency, since strain is excluded as a variable. The frequency of the lowest non-RUM mode was independent of this rotation amplitude.

the frequency. Since there is no coupling to the strain included in CRUSH, only this first contribution, $\partial^2 E/\partial\theta^2$, is calculated, effectively treating all modes as phonons.

However, it can be envisaged that a mode may be stable as a phonon but unstable, or merely of zero frequency, as a RUM. This implies that the partial second derivative $\partial^2 E/\partial\theta^2$ is positive, but the full derivative is zero or negative. In the particular case of the mode drawn in Figure 6.1, all terms in (6.4) are zero around $\{\theta = 0, a = a_0\}$, and hence the mode was drawn without split atoms in Figure 6.3. Away from this high-symmetry structure, however, $\partial a/\partial\theta \neq 0$, so all terms contribute apart from the final term of (6.4) which should always be zero for distortions around the equilibrium volume ($\partial E/\partial a = 0$). In Appendix C it is shown that in a simple bond-stretching model of this mode, the term linear in $da/d\theta$ is negative and cancels out the others, to give a RUM with zero frequency, but a stable phonon. Figure 6.4 shows the frequency of this mode calculated from CRUSH (i.e. using only the partial derivative) when some amplitude of that mode has already been introduced.

The absence of mode-strain coupling in CRUSH thus has the results that not all

RUMs give a strictly zero frequency in the CRUSH analysis. Only RUMs which are independent from the strain to first order are assigned zero frequency; this is also applicable to systems other than the simple one described here. The frequencies of such modes, however, will be both strongly dependent on any initial amplitude of the RUM displacement frozen in, and are typically much smaller then the other modes. They may thus still be clearly identified as being RUMs, and would have zero frequency if coupling to the strain were included.

6.3.1 Corner- and edge- sharing

The majority of structures are linked by sharing one atom between two units, so that in CRUSH the linkage may be represented by one pair of split atoms [99]; the linkage may occur between corners or along edges (giving split pairs at two adjacent corners). However, in structures such as rutile, there is more complicated linking, such that an atom is shared between three polyhedra. Since CRUSH makes no provision for atoms split into more than two 'pieces', such linking was represented in this work by more than one split pair at the same position, so that each pair covers links with one other unit. In practice, it was found that this did not affect the number or nature of the rigid unit modes found (i.e. the incomplete linking using only one pair at each vertex was sufficient) although the frequencies of other modes are naturally changed due to the additional spring constants introduced.

6.4 Perovskites

The perovskite structure ABX_3 is one of the archetypal systems for a polyhedral description, containing BX_6 octahedra which are almost always depicted as distinct entities. Three phases of magnesium silicate perovskite were discussed in Chapter 5: the tilting of the octahedra which relates the phases was illustrated in Figure 5.1. The orthorhombic phase was considered to be formed from the cubic by two successive

distortions which primarily consist of tilting of the octahedra: the applicability of the rigid unit mode model to these distortions will be considered here.

6.4.1 Previous evidence for RUMs

Describing perovskite structural phase transitions in terms of rotation and/or tilting of octahedra is not a new idea: the various space groups which could be reached by combinations of such distortions were set forth by Glazer [111] and described further by Hemley and Cohen [91].

There has been a reported trend in the structural parameters of $MgSiO_3$ towards increased distortion under pressure [92, 100, 102], with which the *ab initio* structural determinations presented here were consistent. X-ray diffraction experiments [94] also showed increasing octahedral rotation under pressure, consistent with rotation favouring a reduction in volume.

The analysis of Bukowinski and Wolf [88] defined rotation angles from ratios of the lattice parameters (e.g. $\cos\theta_M = \sqrt{2}a/c$ for the M_2 rotation), which was used in Table 5.3, effectively assuming that the rotations are purely rigid unit modes. Furthermore, the modes implicated in the transition have been widely described in terms of octahedral rotation [88, 92], and have been used in LAPW[1] calculations as structural parameters [86] where a full structural relaxation would have been impractical.

However, more detailed analysis is necessary to determine whether the octahedral description is a concept useful only in describing the structure, or whether it accurately predicts the behaviour of $MgSiO_3$. The freedom of computational simulations to investigate structures which are metastable or experimentally inaccessible allows this kind of investigation.

[1] Linear Augmented Plane Wave calculations, an all-electron method which is extremely computationally intensive.

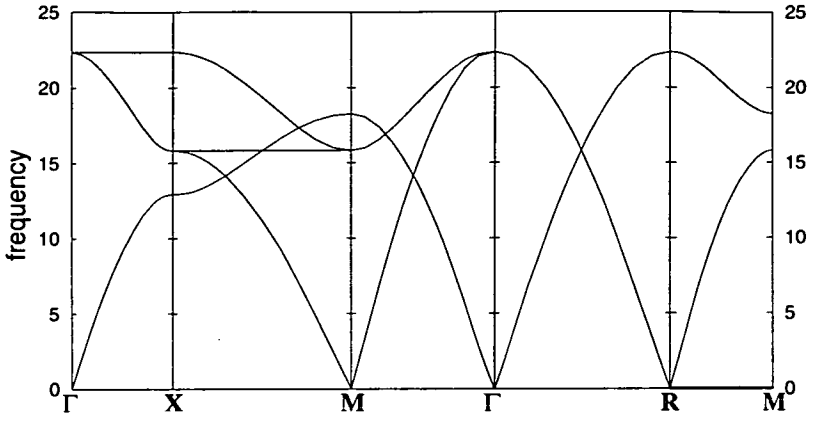


Figure 6.5. Dispersion curve obtained from rigid unit mode analysis of the cubic perovskite structure using the CRUSH package: The six branches correspond to the three translational and three rotational degrees of freedom of the single octahedron in each unit cell. Zero frequencies indicate rigid unit modes, including the three trivial translations at Γ.

6.4.2 CRUSH analysis

Rigid unit mode analysis of the cubic perovskite structure was performed as one of the early demonstrations of the CRUSH package [99], but is repeated here and discussed in more detail for completeness. The cubic perovskite structure contains perfect octahedra, and is thus ideal for such analysis. The dispersion curve obtained is shown in Figure 6.5.

Rigid unit modes are found at R, M and along the line joining these points. At R, there are three degenerate such modes, corresponding to rotations around three axes: in each, the rotations of successive octahedra along a rotation axis (perpendicular to the plane of Figure 6.1) have alternating directions. At each M, there is only a single mode, in which rotations along this axis are all in phase; there are, of course, three M points in the Brillouin Zone. The modes between R to M thus differ only in the phase relation between successive planes of linked octahedra which rotate around an axis perpendicular to the planes. In reality, there is usually a small but finite coupling between these planes [119] but in the CRUSH treatment there is no corresponding force

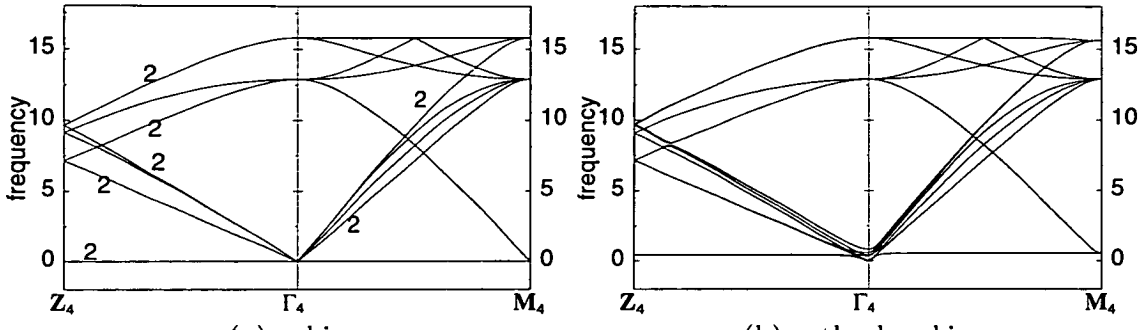


Figure 6.6. Dispersion curves obtained from RUM analysis of the perovskite structure using CRUSH, with a cell containing four formula units ($a = b = a_0, c = \sqrt{2}a_0$). The cubic Brillouin Zone is mapped to that of this larger cell by $\{\Gamma, X(z), M(z), R\} \rightarrow \Gamma_4$; $\{M(x), M(y)\} \rightarrow M_4(z)$; $M(z)$-$R(z)$ to Γ_4-Z_4-Γ_4. The lowest twelve (out of 24) branches are shown; degeneracies are shown in the cubic phase. The RUMs have exactly zero frequency in the cubic phase but are still clearly identifiable in when the atoms are assigned fractional coordinates as in the orthorhombic phase.

constant.

In the *ab initio* investigation in Chapter 5, both the M_2 and the R_{25} mode were in turn frozen into the cubic structure and the phonons of the resulting lower symmetry structure were determined. An equivalent process has been followed with the CRUSH analysis, to determine which modes are inter-related. As discussed in Section 6.3, introducing some amplitude of a strain-coupled RUM will make the frequency non-zero and strongly dependent on the amplitude frozen in; related RUMs may also be affected. This is not the same effect as the stabilisation of modes by introducing an unstable mode found in the *ab initio* work, merely an artefact of the geometry. Nevertheless, it may give useful insights into the mode coupling.

Introducing a finite amplitude of the M_2 mode about z gave that mode, the R_{25} mode with the same rotation axis, and all intermediate modes non-zero frequencies which varied with amplitude. All these modes have equal frequencies due to the zero inter-planar coupling described above. This frequency was found to be independent of lattice parameters a and c and was shown in Figure 6.4. However, the other two R_{25} modes and M_2 modes still had zero frequency.

Freezing in the R_{25} mode (which was found to be slightly more unstable than M_2 in the *ab initio* work) around the xy axis again resulted in a non-zero frequency; the other two R_{25} frequencies increased to half that of the mode frozen in; and the related R–$M(x)$ and R–$M(y)$ RUMs also gained non-zero frequencies. The frequency of the $M_2(z)$ mode and R–$M(z)$ branch remained precisely zero, indicating that the $M_2(z)$ rotation may be considered as independent from these modes.

When both the R_{25} and M_2 modes were frozen into the cubic structure, none of the zero-frequency RUMs were retained: all were however still identifiable by their low frequencies. Figure 6.6 shows the lowest 12 (out of 24) phonon bands calculated by CRUSH in cubic and orthorhombic perovskite, using a cell containing four octahedra.

6.4.3 Analysis of *ab initio* results for $MgSiO_3$

Ab initio calculations presented in Chapter 5 were mostly performed with a twenty atom supercell, containing four formula units: this corresponds to allowing only distortions with Γ, X, M or R wavevector (which are mapped to the zone centre of the orthorhombic cell) to freeze into the cubic phase. RUMs with wavevectors between M and R, for example, are mapped to other parts of the new Brillouin Zone and are thus excluded from this discussion.

Phonons

When the phonons of cubic $MgSiO_3$ were calculated, the most unstable modes were found at R and M, consisting of rotations of the octahedra about the silicon atoms, corresponding exactly to the number and symmetry of the RUMs found from the CRUSH analysis. Other general results from the CRUSH model correspond to the *ab initio* work: when only one rotational mode was frozen in to equilibrium, not all other RUMs were coupled to it in the same way.

Structural information

The structural parameters obtained for the cubic and orthorhombic phases were given in Table 5.2. The orthorhombic phase has a volume per formula unit only 93% of that of the cubic phase, but the oxygen octahedra are larger by about 9% in volume. The angle between Si–OI and Si–OII bonds is 2° from perpendicular, so there is also a small degree of distortion. Furthermore, the single degenerate bondlength in cubic perovskite split into three different bondlengths in the orthorhombic phase, differing from the mean by up to 0.9%. The average Si–O bondlength increased by 2.9% from 1.721 Å to 1.770 Å. The orthorhombic relaxation of ions within the cubic cell described in Section 5.5.1 (i.e. allowing only pure phonon modes to freeze in) gave octahedra 18% larger than in the cubic phase, with approximately 1° distortion.

Since the unit cell has decreased but not by enough to keep the octahedra at the same size while they rotate, the results suggest that neither the phonon or rigid unit mode extremes are perfect descriptions of the $MgSiO_3$ perovskite phases. The distortion of bond angles within octahedra also demonstrates that the rigid unit model is not perfectly adhered to during the transition.

Orthorhombic structure under pressure

Karki [18, 107, 120] has determined the structure, bond lengths and angles of orthorhombic $MgSiO_3$ perovskite under pressures up to 140 GPa using the CASTEP code, using different pseudopotentials for Mg and O as described in Chapter 5. The extent to which RUM behaviour represents this system was determined by the author, using these results and from them calculating the volume of the octahedra: this analysis will now be described.

In the perfect cubic perovskite, SiO_6 octahedra occupy 1/6 of the total volume; if octahedra remain rigid, this fraction must increase with RUM rotation angle. When two rotational modes are involved, with rotation angles θ_1 and θ_2, this fraction increases to $1/(6\cos^2\theta_1\cos^2\theta_2)$. Figure 6.7 shows that if this fraction is calculated using Si–O–Si

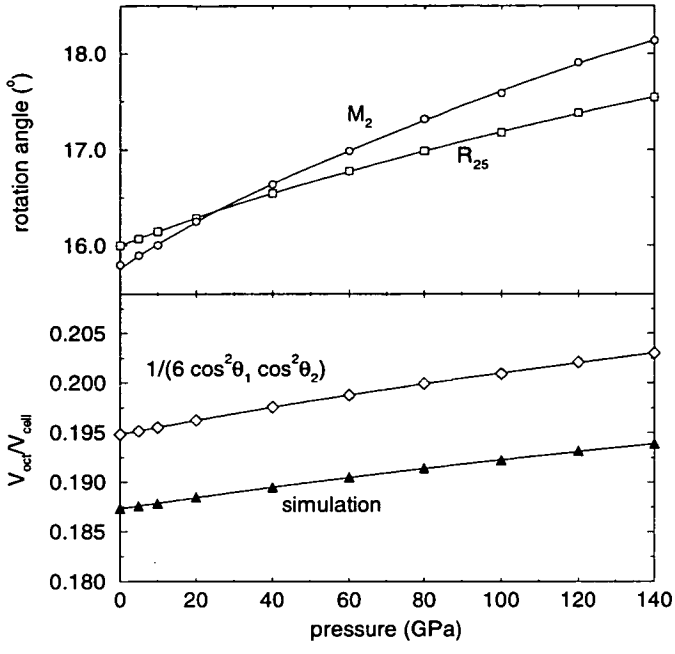


Figure 6.7. Rotation angles θ_1 and θ_2 for the two rotational modes contributing to orthorhombic perovskite (M_2 and R_{25}), and the fractional volume occupied by the SiO_6 octahedra as calculated from results of Karki [120], compared to the prediction for the fractional volume from θ_1 and θ_2 for perfect octahedra in the RUM limit.

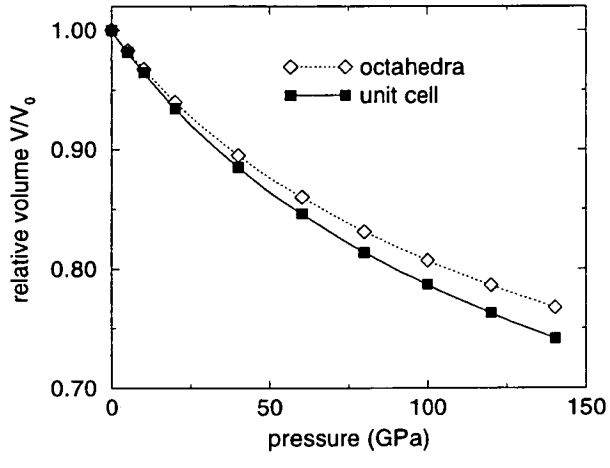


Figure 6.8. Volumes of the unit cell and SiO_6 octahedra of orthorhombic $MgSiO_6$ relative to their values at 0 GPa (calculated from results of Karki [120]).

bond angles from *ab initio* structural data, it does indeed increase with compression. The octahedra are in fact uniformly smaller, since this prediction is for perfect octahedra in the RUM limit. The deviation from the RUM ideal can also be seen in Figure 6.8, in which the octahedral volume can be seen to decrease almost as fast under pressure as the cell volume: in the RUM limit the octahedral volume would be constant. The octahedral volume reduces primarily by compression of the Si–O bonds: the O–Si–O bond angles do not change dramatically. The shape of the octahedra is thus well preserved even under compression. Karki also showed that the MgO_{12} coordination polyhedra are greatly distorted under pressure [120], confirming that the SiO_6 octahedra are the structures of interest.

It may appear that this compression of octahedra contradicts the expansion found at zero pressure with respect to the cubic phase (Section 6.4.3). The phonon extreme describes modes which are uncoupled to the strain; the RUM limit gives modes which are highly coupled to the strain. In the cubic–orthorhombic transition, in which a mode freezes in to break the symmetry, a RUM model would predict constant octahedral volume but a phonon picture (which must preserve the lattice parameters) an increased octahedral volume. During compression of the orthorhombic phase, under which lattice parameters must decrease, a RUM model would again preserve octahedral volume, forcing rotational modes to freeze in further. However, in a phonon model the rotational phonon amplitudes would remain constant, in the limit of zero phonon-strain coupling, so that the octahedral volume decreases with the cell volume. In both the cubic–orthorhombic transition and the compression of the orthorhombic phase, therefore, it is found that the observed behaviour lies between the RUM and phonon ideals, although there are different consequences in the two processes.

Phonons or rigid unit modes?

To further determine whether these rotations are best described in term of phonons or RUMs, the energy of the M_2 mode was investigated in both cases. To create a RUM,

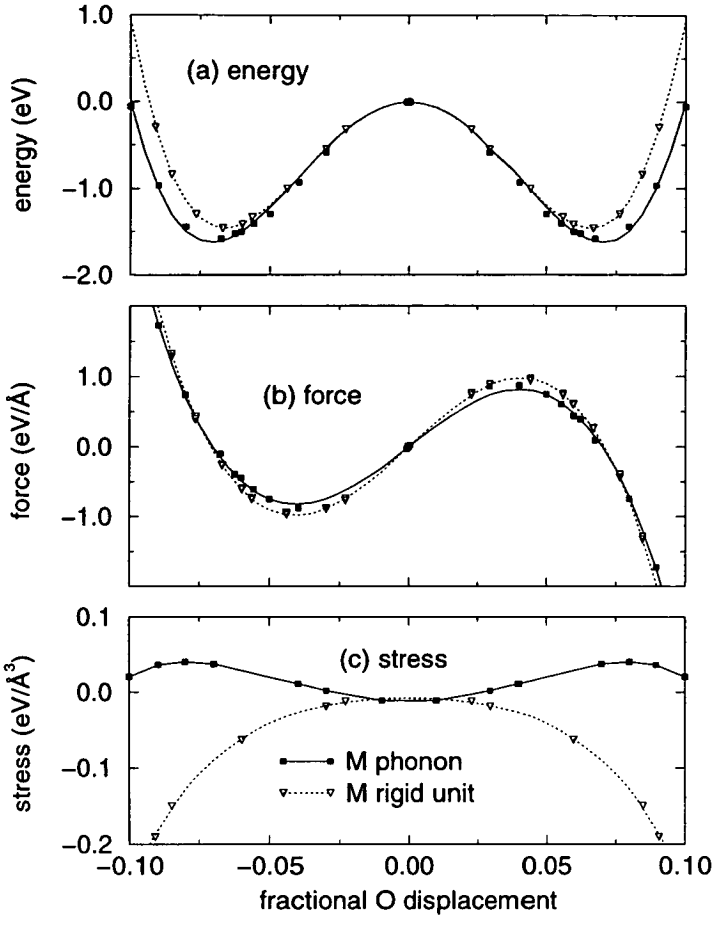


Figure 6.9. (a) Energies, (b) forces and (c) stresses for M_2 rotation of the octahedra in $MgSiO_3$ through different angles (labelled by change in fractional coordinate of oxygen), as both a phonon and a RUM. The energies are fitted to a quartic well, and the forces to a cubic. The results with constant cell size (phonon) are very similar to those with constant octahedral size (RUM), except for the stresses.

the unit cell was changed such that the size of the octahedra was constant while an M_2 rotation was frozen in, as shown in Figure 6.1. The results for the pure phonon and the RUM (phonon and strain) are shown in Figure 6.9. The RUM potential well had a minimum at approximately the same rotation angle as that for the phonon distortion, but was not quite as deep (as summarised in Table 5.6). The RUM is therefore not as good a description as the phonon rotation with fixed unit cell, in which the octahedra expand.

The stress on the unit cell was also calculated (with Pulay corrections) for the two

cases: the phonon rotation with constant unit cell had positive stress but the rigid unit mode with constant octahedra always had negative stress. This implies that the absolute minimum of energy with respect to rotation and strain lies somewhere between these two extremes. This is consistent with the equilibrium orthorhombic structure, in which the unit cell and octahedral size are intermediate between the requirements of phonons and rigid unit modes. In the case of relaxed ions in the cubic cell, the enlargement of the octahedra is much greater, since there can only be pure phonon modes.

The rigid unit mode approximation is useful not only in picturing the structure, but also to describe the changes in bondlengths and lattice parameters during a transition; this was also evident in the structural analysis in Section 6.4.3. The clear mechanism for a reduction in volume confirms that the orthorhombic phase is increasingly favoured relative to the cubic phase under pressure, and becomes more distorted. However, the RUM model is far from a perfect quantitative description: the structure does not exhibit purely RUM compression and the constant volume phonon distortions have slightly lower energy.

6.5 $BaZrO_3$ as a precursor to ZrW_2O_8

The cubic mineral ZrW_2O_8 has an open framework structure containing both octahedra and tetrahedra, but has a negative and isotropic thermal expansion [121]. Recently, the reduction in volume due to excitation of rigid unit modes was invoked as an explanation for this behaviour [14, 16]. Instead of simulating this complex structure, the rigid unit model was justified by investigating the behaviour of ZrO_6 octahedra in a perovskite phase: this Section describes this work in more detail.

A cation is required for charge neutrality: $BaZrO_3$ perovskite is known to be stable, so is a suitable candidate for study. However, the *s* valence electrons in heavy Group II atoms are only weakly bound, so that conventional pseudopotentials would require large

core radii, causing core-overlap problems in solid-state calculations. Furthermore, since the ground state of all Group II atoms contains only s valence electrons, highly excited states would be required to find potentials for other angular momentum components in a non-local pseudopotential. To overcome these problems, some of the electrons in the core may be treated as valence electrons, so that a smaller core may be used and other angular momentum components are present in the reference ground state [30]. For example, the $5s$ and $5p$ electrons are taken as valence in the potential Ba015Rg (see Appendix A for specifications of all pseudopotentials). This process does, however, increase the cutoff energy required, since the nucleus is then less well screened, and core electrons must be represented by plane waves. Such pseudopotentials are still not guaranteed to be as reliable as those for less problematic elements.

A single unit cell of $BaZrO_3$ cubic perovskite contains one ZrO_6 octahedron, and its phonon modes include some simple distortions which may be related to bond-bending and bond-stretching [14]. A basis set energy cutoff of 900 eV and a $4\times4\times4$ Monkhorst-Pack k-point grid were used, giving convergence of the total energy to within 0.1 eV per formula unit. The equilibrium structure was found at a=4.245 Å with the Ba013 pseudopotential and 4.253 Å with Ba015Rg, compared to 4.19 Å experimentally and 4.156 Å in a previous pseudopotential calculation [96]. This overestimation of the lattice parameter is an indication that the pseudopotential generation method is not yet perfected for barium [30].

The frequencies and eigenvectors of zone centre modes using the Ba015Rg pseudopotential are given in Table 6.1. The frequency of mode 4, in which the Ba ion 'rattles' in the 12-fold coordinated site between octahedra, varies considerably with the choice of Ba pseudopotential. The frequencies are volume-dependent, as shown in Figure 6.10 for pseudopotential Ba015Rg; the 'rattle' mode becomes unstable at higher volumes, as might be expected. Conversely, at smaller volumes (closer to that observed experimentally), even the mode calculated with Ba013 should stabilise.

However, we are primarily concerned with those modes involving mainly Zr and

		ω_1	ω_2	Ba	Zr	OI	OII	OIII
1	Γ_{15}	18.02	18.03	-0.02	-0.22	0.94	-0.17	-0.17
2	Γ_{15}	5.54	5.36	-0.33	0.74	0.03	-0.41	-0.41
3	Γ_{15}	0.00		0.70	0.57	0.24	0.24	0.24
4	Γ_{15}	0.01	$2.35i$	-0.63	0.26	0.23	0.49	0.49
5	Γ_{25}	3.44	2.35	0	0	-0.71	0.71	0

Table 6.1. Frequencies (in THz) and eigenvectors of the Γ modes of cubic $BaZrO_3$, at $a = 4.245$ Å. Displacements are given in mass-reduced coordinates; imaginary frequencies correspond to unstable modes. Frequencies ω_1 were obtained with Ba pseudopotential Ba015Rg, ω_2 with Ba013.

O (modes 1,2 and 5), which are affected to a much lesser extent by the choice of pseudopotential, so this variation is not of direct consequence. The Ba015Rg potential was chosen for subsequent calculation, since it gave a lattice parameter slightly nearer to the experimental value and no phonon instabilities. Nevertheless, the unstable Ba 'rattle' mode predicted by the Ba013 potential suggests that it may simulate Ba ions which are more isolated from the ZrO_6 octahedra, providing a more valid extrapolation from the behaviour of octahedra in $BaZrO_3$ to those in the open structure of ZrW_2O_8.

The frequencies of the other modes, which involve distortions of the octahedra, suggest that the ZrO_6 octahedra are not as stiff as SiO_6 (see Table 5.1), but the Γ_{25} mode, the softest stable mode, is not very much softer than that in $MgSiO_3$, so rigid unit modes may still be applicable.

A comparison of freezing in a phonon and a RUM for the M_2 rotational mode, equivalent to that described in Section 6.4.3, was performed for $BaZrO_3$ at a=4.245Å; the results are shown in Figure 6.11. There is a weak instability for this mode, but the decrease in energy is an order of magnitude smaller than in $MgSiO_3$. However, experimentally the cubic structure is stable [122]; this implies either that the *ab initio* calculated volume is too far from the experimental value, or that there is in fact a low-temperature phase transition ($T \approx \Delta E/k_B \approx 700$ K).

Despite the indications from the Γ modes, the RUM is slightly more stable than the

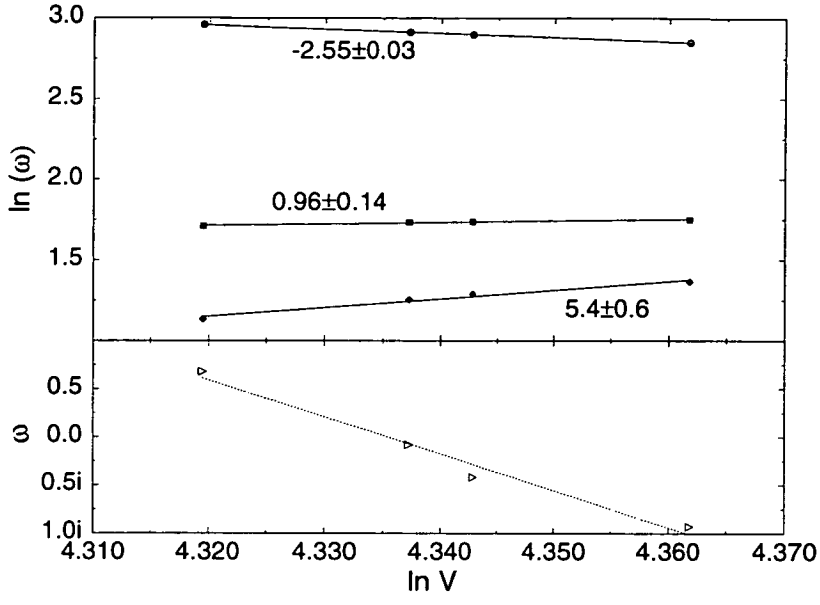


Figure 6.10. Variation of $BaZrO_3$ Γ phonon frequencies with volume; phonons which are stable across the entire volume range have $\ln(\omega)$ plotted against $\ln(V)$ and the Grüneisen parameters calculated; the frequency of the Ba 'rattle' mode is plotted directly since it changes from being stable to unstable.

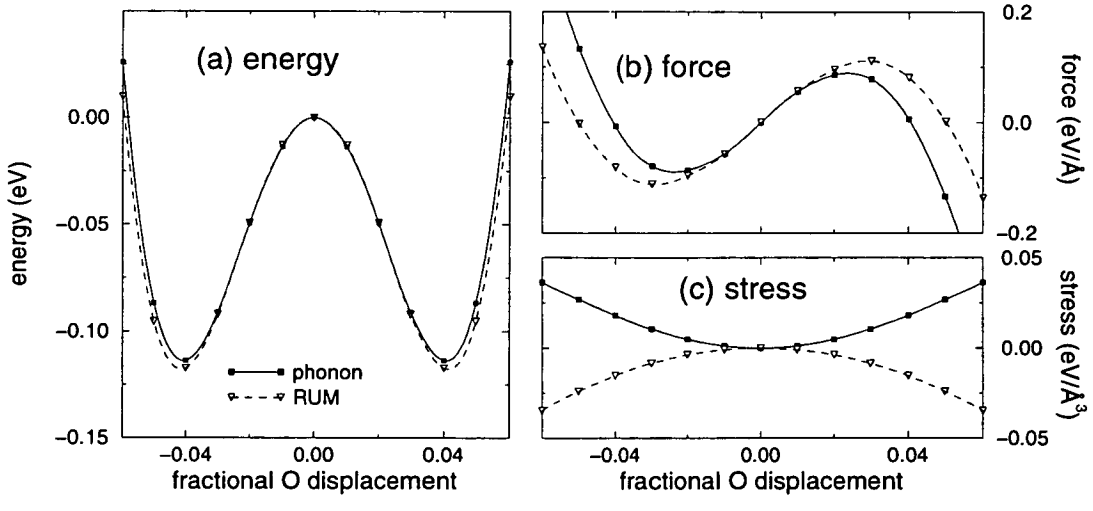


Figure 6.11. (a) Energies, (b) forces and (c) stresses relative to the undistorted structure, for M_2 rotation in $BaZrO_3$. Lines are a guide to the eye only: quartic fits to the energy did not produce good distinct fits to the two separate data sets. The rigid unit mode (RUM) gives slightly lower energy than the pure phonon, but again the stresses suggest an intermediate mode with lower energy.

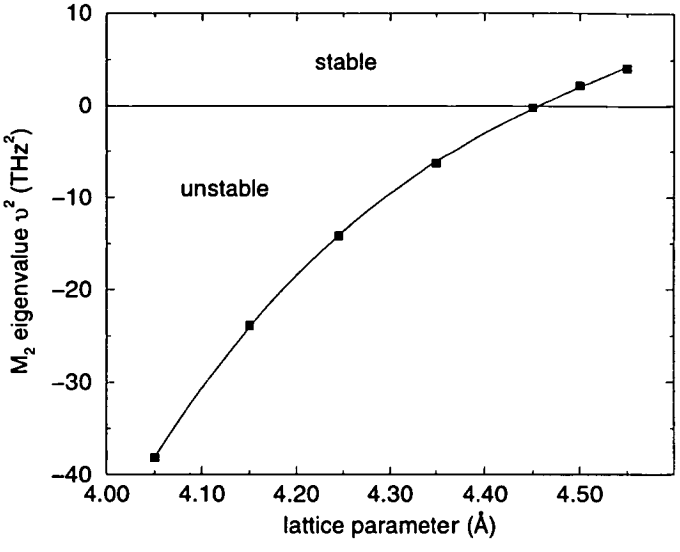


Figure 6.12. Square of frequency (proportional to negative eigenvalue) for the M_2 phonon mode in $BaZrO_3$ at different lattice parameters. This mode stabilises as the volume is increased, becoming stable at a=4.458 Å.

phonon, but again the stresses indicate that the lowest-energy mode would have mixed character. With a fractional oxygen displacement of 0.05, the a and b lattice vectors were relaxed while c was held constant. Zero stresses σ_{xx} (= σ_{yy}) occurred at 5.900 Å, compared to 6.003 Å for the undistorted structure and 5.887 Å for a RUM. This again demonstrates that the preferred mode is closer to a rigid unit mode. Figure 6.12 shows that the phonon mode is stable for lattice parameters above 4.46 Å (5% above the calculated equilibrium lattice parameter and 10% above the experimental value) and confirms that the frequencies of RUM-like modes are strongly coupled to volume.

It is obviously not strictly accurate to compare octahedra in perovskite to the more open structure of ZrW_2O_8 and so only qualitative conclusions should be drawn about the behaviour of the more complicated structure [14, 16]. However, if some rigid unit character is found in the perovskite, ZrO_6 octahedra in more complex structures should also display rigid unit mode behaviour. It is hoped that as ever more powerful computing facilities become available, phonon calculations on larger systems of up to 100 atoms, including heavy Group II and transition elements, will become routine.

6.6 Silica

Silica (SiO_2) undergoes several pressure-induced phase transitions, some of which are thought to be responsible for discontinuities in the mantle [123, 124]. As the pressure rises to 5 GPa, the rutile structure becomes stable, and contains six-fold coordinated silicon ions. At higher pressures, it transforms first to a $CaCl_2$ structure, then to a phase with space group $Pnc2$ and possibly to $Pa\bar{3}$ at pressures above 200 GPa. B.B. Karki [15] has determined the structure of these phases and the phonon modes of the rutile and $CaCl_2$ phases at various pressures using CASTEP with the unit-cell optimisation and phonon calculation techniques described in Chapter 3. Rigid unit mode analysis of his results is presented here, expanding on the conclusions presented in [15], with the aim of finding mechanisms for the various phase transitions and investigating whether they may involve rigid unit modes.

Structural data for rutile up to 16 GPa has been determined by single-crystal X-ray diffraction [115]: it was remarked that the SiO_6 octahedra have the largest bulk octahedral bulk modulus[2] of the rutile-type oxides, and that the structure does not change appreciably with pressure. As with $MgSiO_6$ perovskite, the Si–O bondlengths compressed under pressure, but the O–Si–O bond angles were not found to change significantly. However, this is a low pressure study compared to ambient conditions in the Earth's mantle, whereas *ab initio* techniques can probe further.

6.6.1 Rutile and $CaCl_2$ phases

The tetragonal rutile ($P4_2/mnm$) phase is shown in Figure 6.13, and can be described in terms of distorted oxygen octahedra, linked along z by edge sharing, and in the xy plane by corner sharing. There is only one ionic structural parameter, which places oxygen ions at (x,x,0). The octahedra are distorted in that there are two different

[2] i.e. the bulk modulus relating the volume of the octahedra to the pressure, rather than the unit cell volume

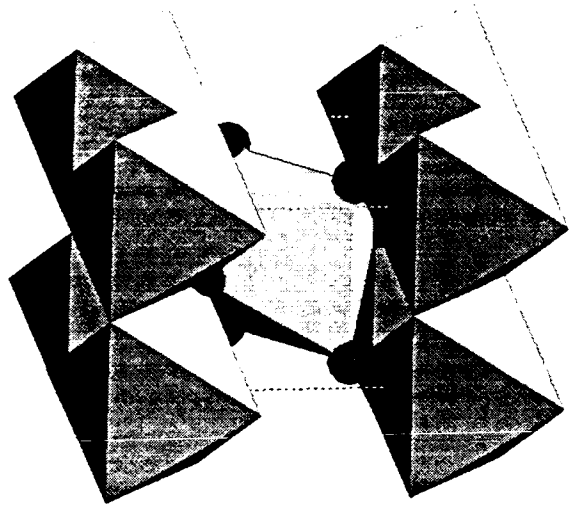


Figure 6.13. The tetragonal rutile structure of stishovite, with SiO_6 octahedra drawn as solid bodies: each oxygen links three octahedra.

Si–O bondlengths (with degeneracies given in brackets):

$$l_1(4) = \sqrt{2a^2\left(\frac{1}{2}-x\right)^2+\frac{c^2}{4}} \tag{6.5a}$$

$$l_2(2) = \sqrt{2}ax \tag{6.5b}$$

and the angles between the four equal l_1 bonds are not 90°, but are instead θ and $\pi-\theta$, where

$$\cos\theta = \frac{2\left(\frac{a}{2}-ax\right)^2-\frac{c}{4}}{2\left(\frac{a}{2}-ax\right)^2+\frac{c}{4}} \tag{6.6}$$

The orthorhombic $CaCl_2$ structure ($Pnnm$) structure is similar, but oxygens are instead found at (x,y,0); the structure is similar to that shown in Figure 6.14(b), but without the constraints described therein. It can be seen that the octahedra are rotated with respect to their orientations in the rutile phase, suggesting that a rigid unit mode picture should be considered.

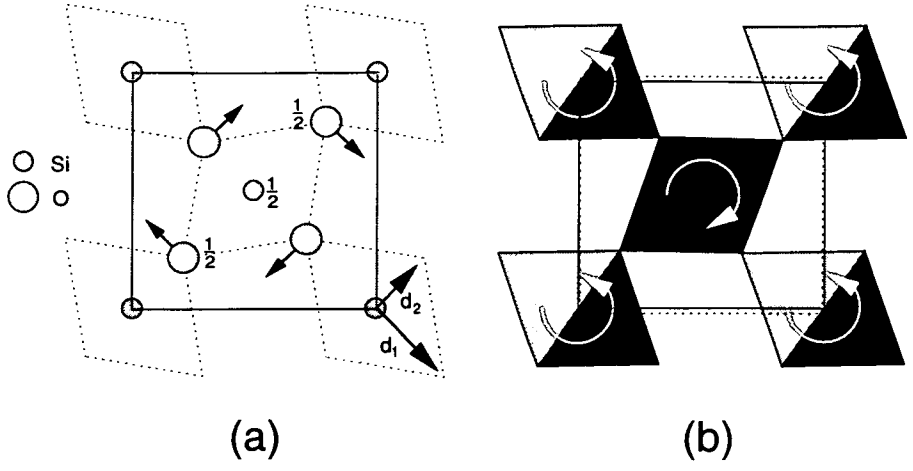


Figure 6.14. (a) Tetragonal SiO_2 (rutile) with outlines of the SiO_6 octahedra and the eigenvector of the B_{1g} mode, which rotates the octahedra; (b) orthorhombic ($CaCl_2$) phase, formed by assuming that the SiO_6 octahedra remain rigid units, and rotating them around the z axis as shown. The dotted cell outline shows the rutile unit cell. The *ab initio* structure does not differ greatly from this simplification. The lengths d_1 and d_2 label the rotation arms in the plane of the figure.

Indicators of mechanism from prediction of P_c

Karki *et al.* [15] identified the transition pressure for the rutile $\rightarrow$ $CaCl_2$ transition by examining the relative enthalpies, elastic moduli for orthorhombic strains and phonon modes of both phases at different pressures. The results from these three methods agreed very closely, to give $P_c = 47 \pm 1$ GPa. However, the results also displayed important pointers to the underlying mechanism.

The $c_{11} - c_{12}$ elastic modulus of the rutile phase was calculated by Karki by applying a small orthorhombic strain (breaking the $a = b$ lattice symmetry) to the unit cell, allowing the ions to re-relax, and then measuring the stress from *ab initio* calculations. It was found to decrease dramatically beyond 30 GPa, and become negative beyond 46 GPa, indicating an elastic instability, and thus a transition to an orthorhombic phase, in agreement with previous DFT studies [113]. It is significant that the modulus does not decrease with pressure unless the ions are allowed to re-relax after distorting the cell, indicating a strong coupling between ionic positions and elastic properties.

The calculated phonons of the rutile phase were found to be in good agreement with Raman spectroscopy. Raman techniques probe only a small region of the sample, so coupling to the strain is suppressed, and therefore only phonons are detected. Comparison between Raman results and *ab initio* phonon calculations is therefore a valid way of confirming that simulations are reliable, but it does not show whether phonons are actually the normal modes of the system.

All calculated phonon frequencies increased with pressure, except for the mode with the lowest frequency (B_{1g} symmetry) which decreased from 6.6 THz at 0 GPa to 3.8 THz at 60 GPa. Such softening is a strong indicator of a phonon-related phase transition, and the importance of this mode has already been identified [112, 113]. The frequencies of all phonons of the $CaCl_2$ phase were found to increase under pressure; when identified with their closest match in the rutile phase by eigenvector, in a method similar to that of Section 5.4, the lowest mode (A_g) was identified with the soft rutile B_{1g} mode. On extrapolation, the $CaCl_2$ A_g frequency crossed that of the soft rutile mode at 48 GPa (at 5.1 THz): the transition therefore occurs before the soft phonon frequency in rutile can decrease completely to zero. The calculated eigenvector of this mode is shown in Figure 6.14(a).

Further analysis

The calculation and analysis of pure phonon modes was discussed above, but the rutile–$CaCl_2$ transition is now analysed as part of this thesis in terms of RUMs rather than phonons, to see if they give a more accurate description of the dynamics and might explain observed features of the transition. The rutile soft-mode eigenvector clearly resembles a rotation of the SiO_6 octahedra around the z axis, corresponding to the displacements observed between the rutile and $CaCl_2$ phases, as expected for a soft-mode phase transition. However, unlike the perovskite M_2 mode which it superficially resembles, the different lengths of rotation arms in the xy plane, d_1 and d_2 say, require that the tetragonal symmetry is broken for a pure rigid unit mode, as shown

in Figure 6.14. The two-fold degenerate Si–O bondlength l_2 forms d_1, whereas d_2 is half the length of the O–O bond in the xy plane, and so is rather shorter (the ratio $d_1/d_2 = 2x/(1-2x) = 1.57$ at 0 GPa). To maintain rigid units at some rotation angle θ the lattice parameters must then take the form

$$a = 2\left[d_1 \sin\left(\frac{\pi}{4}+\theta\right) + d_2 \cos\left(\frac{\pi}{4}+\theta\right)\right] \tag{6.7a}$$

$$b = 2\left[d_1 \cos\left(\frac{\pi}{4}+\theta\right) + d_2 \sin\left(\frac{\pi}{4}+\theta\right)\right] \tag{6.7b}$$

so that

$$\left.\frac{\partial a}{\partial \theta}\right|_{\theta=0} = \sqrt{2}(d_1 - d_2) = -\left.\frac{\partial b}{\partial \theta}\right|_{\theta=0} \tag{6.8}$$

For small angles, one lattice parameter must, according to (6.8), increase slightly under a RUM, and the other decrease, as was observed in the *ab initio* calculations.

CRUSH analysis was performed for all four phases. The rutile structure does *not* have a zero-frequency mode, because the necessary coupling to the strain is of first order even around the high-symmetry phase, as shown in (6.8). Only if the octahedra are distorted so much that they appear square in projection along z ($d_1 = d_2$) is a zero-frequency rotational mode obtained at Γ; this requires $x = 0.25$, compared to 0.306 at 0 GPa and 0.300 at 140 GPa. This limit is thus approached only very slowly under increasing pressure so does not in itself wholly explain the onset of the transition. However, this was the only RUM found in this structure, even for $x = 0.25$: the edge-sharing between octahedra severely restricts the possible modes.

The RUM picture provides a natural explanation of the coupling between the rutile B_{1g} mode and the elastic modulus for orthorhombic distortion, since the B_{1g} mode provides a low-energy compression mechanism. Without ionic relaxation into the 'twisted' state when calculating $c_{11}-c_{12}$, the structure cannot access this compression mechanism, so the elastic modulus remains high; without allowing orthorhombic strain, the B_{1g} mode cannot soften to zero. If both strain and the rotational mode are accessible, $P\Delta V$ may provide a driving force for the transition to the $CaCl_2$ phase. This transition

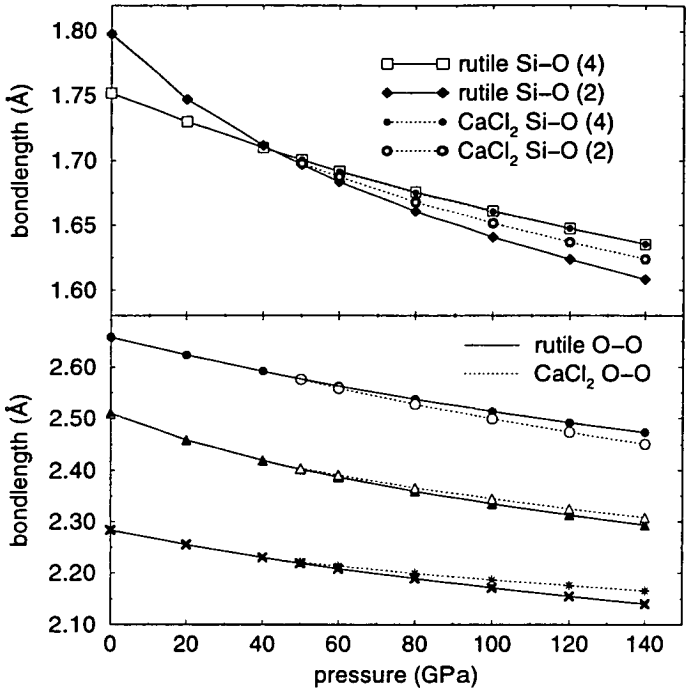


Figure 6.15. Si–O and O–O bondlengths within SiO_6 octahedra in rutile and $CaCl_2$ phases of stishovite under pressure.

is expected to be second-order [113], because both the lattice parameters and the ionic positions may increase continuously from zero by this mechanism.

Figure 6.15 shows the Si–O and O–O bondlengths in the rutile and $CaCl_2$ phases obtained from the CASTEP simulations. The spread of O–O bondlengths shows that both structures are quite far distorted from containing perfect octahedra: only very slight convergence towards equal bondlengths is observed. In contrast, the two Si–O bondlengths in rutile cross over, and are equal at 43 GPa, just below the predicted transition pressure. The $CaCl_2$ structure preserves their ordering, although their difference is smaller. In a pure rigid unit mode, however, these bondlengths would be unaffected by pressure, as would the bond angles within octahedra. The angle between the four-fold degenerate Si–O bonds is shown in Figure 6.16. The angle between this plane of four equal bonds and the remaining Si–O bonds is exactly 90° in rutile and

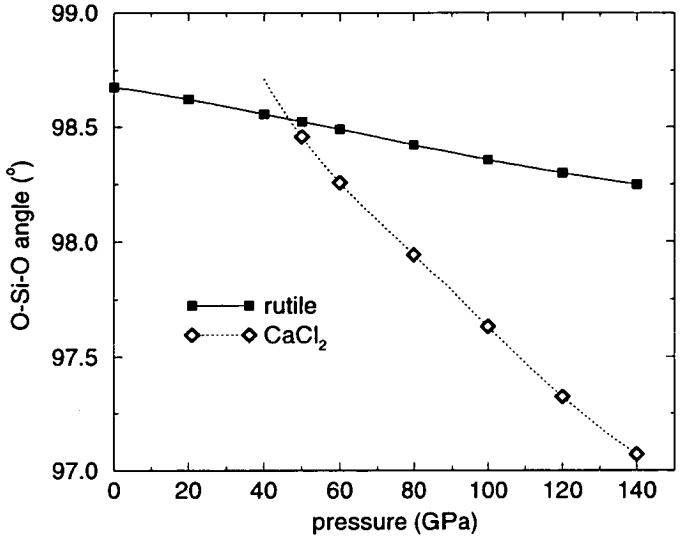


Figure 6.16. Angle between two Si–O bonds of equal length in rutile and $CaCl_2$ phases of stishovite. The ideal octahedral angle is 90°. Extrapolation gives $P_c = 47$ GPa.

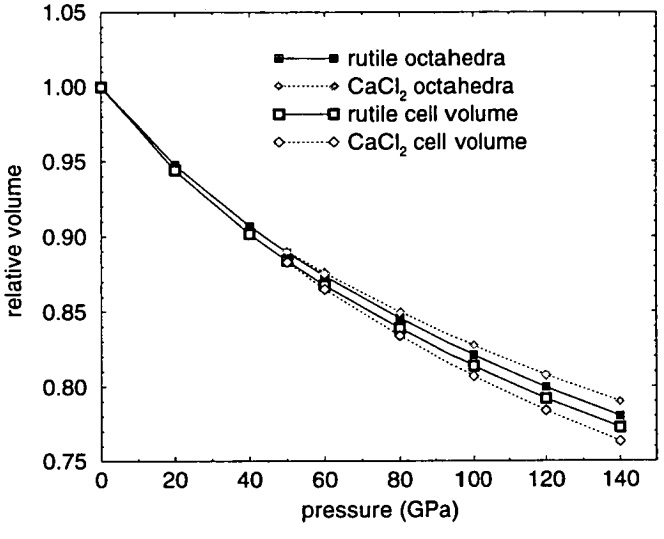


Figure 6.17. The compression of the unit cell volume and that of SiO_6 octahedra in rutile and $CaCl_2$ phases of stishovite under pressure.

changes little in $CaCl_2$. Although the $CaCl_2$ phase displays a greater decrease in distortion with pressure, the octahedra remain significantly distorted from an ideal form, and their shape is not perfectly preserved under pressure in either phase.

The cell and octahedral volumes, another indicator of the extent of RUM behaviour, are shown in Figure 6.17. The octahedral volume is preserved only slightly better than the cell volume, whereas in a pure rigid unit mode the octahedral volume would be expected to be independent of pressure. When compared to the behaviour shown for orthorhombic $MgSiO_3$ in Figure 6.8, it can be seen that the RUM model holds even less well in the rutile and $CaCl_2$ structures than in perovskite.

Although the rigid unit mode picture of this transition is appealing in terms of understanding the strain-B_{1g} mode coupling, in the light of the bondlengths and volume data it seems likely that it is not by any means an accurate description of the detailed structural behaviour. Instead, it is the marked crossover of the bondlengths in the rutile phase and increase in angular distortion in $CaCl_2$ relative to rutile that characterise this transition. However, even a small amount of RUM character, as observed here, is sufficient to cause some coupling with the strain, so the insights gained from studying the pure RUM extreme may usefully be applied to the transition.

6.6.2 *Pnc*2 and *Pa*$\bar{3}$ phases

The orthorhombic *Pnc*2 phase contains four formula units in the unit cell. The cell parameters suggest a simple cell doubling along the z axis compared to the rutile or $CaCl_2$ phases, but the chains of octahedra parallel to z have been lost, and more complicated edge-sharing exists: the shared edges are not all parallel. No rigid unit modes were found in this phase using `CRUSH`.

The cubic *Pa*$\bar{3}$ phase also contains four formula units, with face-centred silicon ions and oxygen ions at (x,x,x). There is only one bond length ($l = a\sqrt{1-4x+6x^2}/\sqrt{2}$) and bond angles of θ, $\pi-\theta$ where $\cos\theta = (2x-1)^2/2(1-4x+6x^2)$. However, $\theta \neq \pi/2$ for all x apart from the unphysical $x = 0.5$; Karki's *ab initio* calculations gave $x = 0.348$

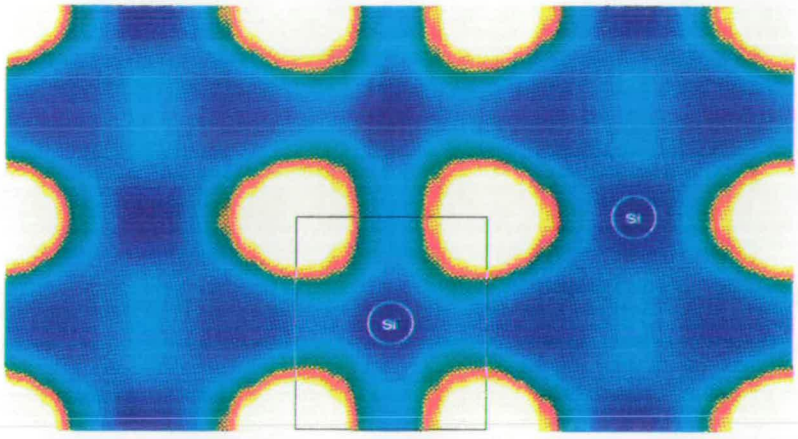


Figure 6.18. Contours of valence charge density (blue=low; red=high; areas inside core radius are excluded) across a (110) plane in rutile SiO_2, which passes through the centre of octahedra. The positions of Si atoms and the outline of one octahedron are shown, with the short edge shared with the octahedron directly above. Significant electron density can be seen across the short shared O–O edge.

at 140 GPa. Unlike the other phases investigated, there is no edge-sharing between octahedra: each oxygen is instead shared between vertices of three octahedra. No rigid unit modes were found in this phase.

These two phases, though they contain the same number of formula units, are structurally very dissimilar: even the silicon positions in the two phases may not be related by any simple transformation. Karki [15] predicted a transition from the $CaCl_2$ to $Pnc2$ at 95 GPa on the basis of enthalpy calculations, and that the $Pa\bar{3}$ phase is the most stable only above 220 GPa. However, these transitions must involve more drastic structural changes than the rutile $\rightarrow$ $CaCl_2$ transition, and therefore considerable hysteresis.

6.6.3 Covalent or ionic bonding?

Density-functional perturbation theory calculations on rutile [125] showed that interatomic force constants are significant over 11 shells for Si atoms and 17 for O atoms, so it is clear that the bonding must involve some ionic character; it was also shown

that dipole-dipole terms are important. However, experimental evidence suggests that rutile has more covalent bonding than most of the oxygen minerals [115]. In the rutile structure there are the right number of valence electrons for it to be, in principle, entirely covalent (although not all bonds are alike); in contrast, the perovskite structure requires cations to balance a charged octahedron.

Figure 6.18 shows the electronic charge density in a [110] plane through the centre of one set of octahedra in the rutile phase at 50 GPa. The bonding appears to be largely ionic, but there is some charge density between the closest oxygen pairs, along the shared edge. This is one of the shortest known non-bonded separations between oxygen atoms and has been proposed to have a slight shielding effect between the silicon atoms in neighbouring octahedra which increases the incompressibility parallel to the c axis [115].

6.7 Conclusion

Rigid unit modes are clearly involved to some extent in the structural phase transitions of several structures. It has been shown that in these cases the eigenvectors of vibrational modes calculated from first principles are in agreement with the RUM approximation, and are consistent with the atomic displacements observed experimentally. In all these structures, the RUM analysis helps to explain the observed structural behaviour. The capability of computational techniques to study structures other than the most stable allows for detailed investigation of such modes.

However, in all the cases studied here, the limit of perfectly rigid octahedra has not been reached, and in the case of $MgSiO_3$ the constant volume phonon description is closer to the full structural behaviour found from *ab initio* calculations. The ionicity of corner-sharing octahedra in the perovskite structure may explain why they are not perfectly rigid: ionic bonding may instead lead to a 'hard sphere' model of the oxygen ions being appropriate, in which the behaviour is determined by avoidance of

compression of the ions. Although the rutile structure is more covalent, not all bonds are equivalent, suggesting that some deformation is possible.

Chapter 7

Simple models of phase transitions

To be simple is the best thing in the world; to be modest is the next best thing. I am not so sure about being quiet.
G.K. Chesterton, *All Things Considered*

First-principles electronic structure calculations can provide much useful information about the nature of structural phase transitions: the relative energies of different phases, the atomic rearrangements necessary to relate them, and the pressures at which phases become metastable. Such studies have been discussed and illustrated in other Chapters of this thesis. However, these calculations are usually performed at zero temperature, because of the enormous computational effort required for *ab initio* molecular dynamics (MD): the electronic charge density must be converged with sufficient accuracy to calculate the Hellmann-Feynman forces after after every movement of the ions. The advent of parallel supercomputers has made feasible *ab initio* MD simulations at finite temperature [126], such as those described in Chapter 5. Nevertheless, the size and complexity of the systems which may be treated are still severely restricted.

Elementary models which apply a simple potential to only a few variables instead of the tens of thousands involved in *ab initio* calculations allow the treatment of much

larger systems. They are perhaps better suited to representing the disorder which is inherent in temperature-induced phase transitions. Such models have, of course, been in use for a long time, but the parameters are usually fitted to experimentally known features or scaled by the transition temperature, so they are used to investigate the progress of a transition rather than predict its onset. Some approaches to using the results of *ab initio* simulations to parametrise simple models, with the aim of predicting transition temperatures for structural phase transitions, are described in this Chapter.

7.1 Phonons as order parameter

It has been shown in previous Chapters that many structural phase transitions involve one or more phonon modes, often at the zone centre or the zone boundary. Such modes are unstable in one phase but stabilise after freezing in to some finite amplitude, forming a new phase with lower symmetry. The mean amplitude of the mode is thus zero in the high symmetry phase and finite in the lower symmetry phase, providing a natural order parameter. The order parameter may decrease continuously to zero, so these transitions are usually second-order. Such phase transitions are also frequently followed experimentally, by observing the relevant mode: the frequency decreases up to the transition temperature or pressure, not necessarily to zero, but increases thereafter by being thermally stabilised [8]. Not all phase transitions may occur in this way: those which involve more drastic atomic rearrangement are not addressed in this thesis.

Parametrisation of models from *ab initio* simulations has already been used for ferroelectric transitions in $BaTiO_3$ [101, 127]. The relevant phonon modes and variables were determined with the use of entire dispersion curves obtained from DFPT[1]. However, at the present time the phonon frequencies may be accurately calculated only at principal symmetry points of the Brillouin Zone, using the CASTEP and CETEP packages, so ways of parametrising these models using this limited information are sought.

[1] Density-functional perturbation theory, see Chapter 4

7.2 Simple estimations of T_c

The most fundamental quantity calculated by *ab initio* techniques is the energy difference between two alternative structures, be they different volumes of the same phase, different ionic structural parameters or different phases with the same chemical composition.

The simplest way of estimating the transition temperature between two phases is to assume that the energy difference ΔE is overcome simply by thermal energy:

$$k_B T_c = \Delta E \tag{7.1}$$

This approximation has its origins in the Boltzmann weighting for thermally activated processes, and can give a order of magnitude estimate for the transition temperature [86]. However, it is a significant oversimplification, since it assumes that every part of a sample is in the same state, ignoring thermal disorder. A somewhat more sophisticated method of estimating T_c is thus required.

7.3 Interacting double well model

We review here one of the simplest systems used to model phase transitions [1, 128, 129]. Consider a simple system with local quartic potentials, having an energy at each site as a function of the local variable x_i, and a non-local energy from coupling between sites via nearest neighbour harmonic interactions. The total energy is then

$$E = \sum_i (\alpha x_i^2 + \beta x_i^4) + \frac{1}{2} \sum_{ij'} J x_i x_{j'} \tag{7.2}$$

where the second sum is over nearest neighbours j'. The local (on-site) potential is $\alpha x^2 + \beta x^4$. If $\alpha < 0$ and $\beta > 0$, the local potential thus consists of a double well; if $\alpha > 0$ the local potential has only a single minimum at $x = 0$.

The interaction terms may be represented by a springs of strength $-J$ between each

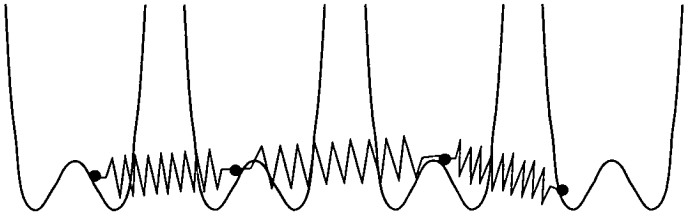


Figure 7.1. One-dimensional representation of local quartic double well potentials coupled by harmonic springs.

variable and each of its N_J neighbours (Figure 7.1), since

$$\frac{1}{2}\sum_{ij'} J' x_i x_{j'} = -\frac{1}{4}\sum_{ij'} J'(x_{j'} - x_i)^2 + \sum_i \frac{N_J J'}{2} x_i^2 \tag{7.3}$$

i.e. the on-site terms resulting from a spring-like interaction may be simply taken to be absorbed into the local potential. Such an interaction in one dimension would tend to keep all the x_i equally spaced, and therefore having the same value. The sign of the interaction parameter J determines whether the lowest energy state is one in which all x_i are equal (for $J < 0$), or whether they alternate in sign ($J > 0$).

7.3.1 Features of model

This simple model has several features which relate it to observable phase transitions. In the slightly more general case a model should be considered in which the interaction may depend on the direction, so that J is replaced by $J_{ij'}$.

Firstly, a characteristic dispersion curve may be obtained [119]. The force on one element is

$$\begin{aligned} F_i &= -\frac{\partial E}{\partial x_i} \\ &= 2\alpha x_i + 4\beta x_i^3 + \sum_{j'} J_{ij'} x_{j'}. \end{aligned} \tag{7.4}$$

However, in a phonon of wavevector $\mathbf{q}$ the displacements of each variable (usually

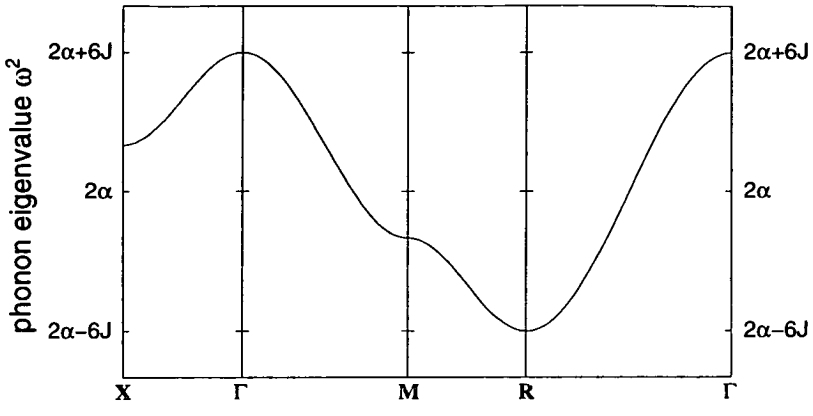


Figure 7.2. Dispersion curve from the simple model of local quartic potentials and harmonic interactions given by (7.2).

atomic positions) are given by $x = A \exp i\mathbf{q}\cdot\mathbf{r}$, so $x_{j'} = x_i \exp(i\mathbf{q}\cdot\Delta\mathbf{r}_{ij'})$, giving

$$F_i = -x_i(2\alpha + \sum_{j'} J_{ij'} \exp[i\mathbf{q}\cdot\Delta\mathbf{r}_{ij'}]) - 4\beta x_i^3 \tag{7.5}$$

which in the harmonic approximation must obey $F_i = -\omega^2 x_i$ and hence

$$w^2(\mathbf{q}) = 2\alpha + \sum_{j'} J_{ij'} \exp i\mathbf{q}\cdot\Delta\mathbf{r}_{ij'}. \tag{7.6}$$

We may scale $\mathbf{q}$ by choosing $|\Delta\mathbf{r}_{ij'}| = 1$ for nearest neighbours. Figure 7.2 shows the resulting dispersion curve for $J_{ij'} = J$. It can be seen that the local harmonic potential α determines the average frequency, whilst the interaction potential J determines the dispersion across the zone. This immediately suggests qualitative ways in which real dispersion curves may be analysed to elucidate information about interatomic interactions. If an isolated phonon band has a large dispersion, it may be assumed to relate to strong interatomic or intermolecular interactions.

At a wavevector $\mathbf{q}_c$ which minimises $\sum_{j'} J_{ij'} \exp(i\mathbf{q}\cdot\Delta\mathbf{r}_{ij'})$ to be $-K$ the frequency is

$$w(\mathbf{q}_c) = \sqrt{2\alpha - K} \tag{7.7}$$

and if $\alpha < 0$ or the interaction K is large enough then this phonon will be imaginary: indeed, there may be several regions of the Brillouin Zone with imaginary modes. As discussed in previously in this thesis, imaginary phonons correspond to structural instabilities. The wavevector $\mathbf{q}_c$ thus describes the periodicity of the strongest instability in the structure.

For example, if $J_{ij'} = J > 0$ as in Figure 7.2, then the most unstable frequency is at the R point of the Brillouin Zone, $\mathbf{q}_c = (\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$, with a value of $w(\mathbf{q}_c) = \sqrt{2\alpha + 6J}$; the x_i then have equal magnitude but alternate in sign between nearest neighbours. If $J_{ij'} = J < 0$, the instability corresponds to $\mathbf{q}_c = 0$, i.e. all x_i are equal. It may be noted that if all $|J_{ij'}|$ are equal, the convention $J_{ij'} = J < 0$ may always be used, by defining the sense of measurement of x to correspond with the appropriate displacement pattern at $\mathbf{q}_c$. From (7.3), this is analogous to having harmonic springs between neighbouring sites. It also redefines $\mathbf{q}_c$ to be the Γ point.

The lowest energy 'phase' is then one in which the displacements adopt this periodicity: this is equivalent to freezing in the phonon at $\mathbf{q}_c$, which is the most unstable in the system. The system may have an ordered instability even if $\alpha > 0$, that is, even if local wells do not have a unstable central maximum and there is thus no local instability. It is the inclusion of the interaction J that produces an overall instability.

If it is taken that $J < 0$, and so $\mathbf{q}_c = 0$, the energy at $\mathbf{q}_c$ is

$$E = \sum_i (\alpha x^2 + \beta x^4) + 3Jx^2 \tag{7.8}$$

giving a lowest energy state at

$$x_0^2 = \frac{-(\alpha + 3J)}{2\beta} \tag{7.9}$$

with energy per site

$$E_0 = -\frac{1}{4}\frac{(\alpha + 3J)^2}{\beta}. \tag{7.10}$$

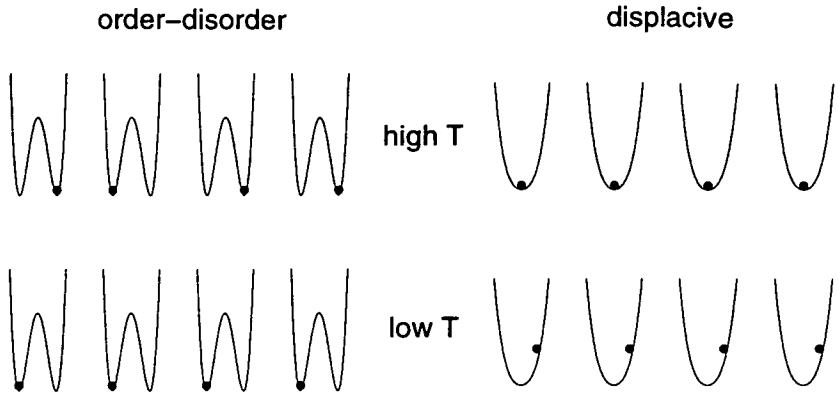


Figure 7.3. Representation of the time-averaged x_i in the low and high temperature phases of order-disorder and displacive regimes, in their local potential wells. Both low temperature phases have ordered arrangements of non-zero x_i, but the high temperature phase only $x = 0$ as the ensemble average whereas in the displacive limit each x_i oscillates around zero.

It may be deduced that the fraction of this energy which is due to interactions, $3Jx_0^2/E_0$, is then (still assuming $J < 0$)

$$\frac{E_J}{E_0} = \frac{6J}{\alpha + 3J} \tag{7.11}$$

and is hence independent of the anharmonicity β; a corresponding expression may be derived for $J > 0$. However, these results apply for any value of α or β.

The phase at $\mathbf{q}_c$ with all $x_i = x_0$ is associated with the low energy (low temperature) phase in frozen-phonon phase transitions. The highest energy phase would occur for some other $\mathbf{q}$. The most symmetrical phase will be that with $x_i = 0$, and is associated with the high-temperature phase.

As the temperature increases from zero, where the low symmetry phase is stable, eventually there is enough thermal energy for at least the average x to be zero over macroscopic regions of the sample. This can occur in two ways, as illustrated in Figure 7.3. Firstly, if the local potential has two deep minima, the local variables tend to fluctuate with time mostly inside only one of these wells, even at high temperature, but become disordered over the system so that the ensemble average is zero; this is the order-disorder type of transition. Alternatively, if the local potential is more like a

single well centred on $x = 0$, or only a weak double well, individual x_i will average to zero with time; this is known as the displacive regime since only small displacements need be made to order the phase. However, in both of these cases there must be some transition temperature T_c below which the ordered phase is stable.

For a deep double-well local potential, i.e. the order-disorder limit, we require α to be negative and stronger than the interaction J. This will give a dispersion curve which is entirely unstable. In contrast, the displacive limit requires $\alpha > 0$, in which case at least some parts of the Brillouin Zone have stable phonons. The crossover between these two limiting behaviours has been discussed by Evans *et al.* [128].

7.3.2 Molecular dynamics simulations

The transition temperatures of this simple model may be compared with the initial approximation that $k_B T_c = E_0$. The transition temperature was determined from molecular dynamics calculations in a system with a grid of 16×16×16 sites and periodic boundary conditions, using the massively parallel CM200 supercomputer. Different values of α/J were used, to give a range of behaviour between displacive and order-disorder extremes. Figure 7.4 shows the fraction $k_B T_c/E_0$ for some systems across this range. The spread of values arises from different values of β, typically from $\beta/\alpha = 0.04$–0.4.

It can clearly be seen that the transition temperature varies considerably from the simplistic value of unity from (7.1); the ratio $k_B T_c/E_0$ varies by over a factor of three across the range of α/J shown. Although (7.1) may provide an order-of-magnitude estimate of the transition temperature, it is clearly not reliable for a good determination of T_c.

However, the energy E_J may also be considered: this is due only to the interaction between neighbouring sites, i.e. the last term in (7.2). Figure 7.5 shows the transition temperature as a fraction of E_J for the same systems as in Figure 7.4. Although there is clearly still a systematic variation in this fraction, it is less dramatic than for $k_B T_c/E_0$,

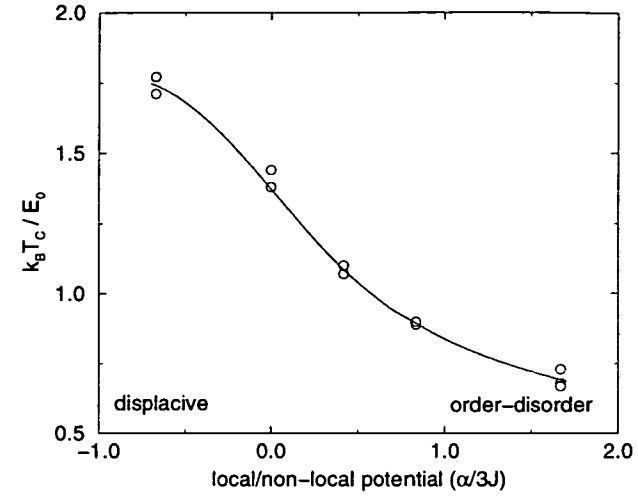


Figure 7.4. Transition temperature as a proportion of the total energy difference between two 'phases' of a simple double well model with nearest neighbour interaction, plotted against the ratio between the local and interaction harmonic potentials. Different values of the anharmonic local potential β were used, giving a spread of results around the general trend.

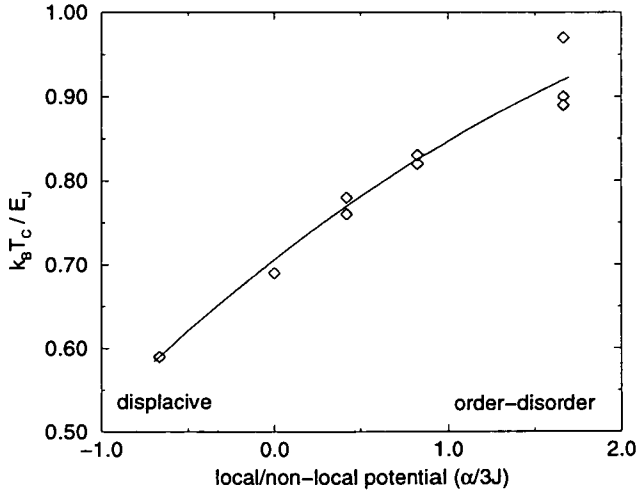


Figure 7.5. Transition temperature as a proportion of the interaction energy in a simple double well model, plotted against the ratio between the local and interaction harmonic potentials. The spread of temperatures is again due to the different anharmonicities.

increasing by a factor of only 1.5 over the same range of α/J as before. Although the effect of different values of β is particularly strong in the order-disorder limit, no clear trends were found; further work is necessary to interpret this variation other than as effective error bars on the results.

7.3.3 Choice of localised variables in real systems

In order to investigate long-range order and disorder in specific phase transitions, local variables must be formulated, which will act as the x_i in this model. These variables must then be affected by both local and interaction (non-local) potentials: in many cases the interactions may be assumed to be nearest-neighbour potentials only. The eigenvectors of any relevant phonon modes may help to determine suitable variables.

7.4 Magnesium silicate

If either the transition between the tetragonal and orthorhombic phases, or that between cubic and tetragonal phases, may be modelled by only one local coordinate, this would simplify the system greatly and the model described above may be applied. The work described in Chapter 5 showed that both these transitions were dominated by the freezing in a phonon consisting of octahedral rotations. This suggests that a local variable related to the rotation of each individual octahedron may be considered.

7.4.1 Choice of local variables and order parameter

Although octahedral rotations are suggested as the appropriate variables, at points away from M and R (of the cubic Brillouin Zone), there are no normal modes consisting purely of octahedral rotations. Using the simple rigid unit mode model discussed in Chapter 6, it was found that at other points on a dispersion branch containing a rotational phonon, there must be modes which involve distortions of the octahedra. The exact nature of the local variable thus cannot be simply an octahedral rotation angle,

since this does not describe the mode at other points of the zone. An alternative choice for a local variable could be the polarisation vector $\boldsymbol{\xi}$ (or eigenvector) of the branch containing the phonon of interest, which defines displacements over the primitive cell and in other cells by modulation according to the wavevector (Section 4.1).

However, in the real system, there are considerably more variables than in the rigid unit mode model: not only are there more ways of distorting the octahedra, but there are also Mg ions which participate in the normal modes and transitions. Phonon branches with the same symmetry will mix instead of crossing, so the eigenvector $\boldsymbol{\xi}$ may not be the same across an entire continuous branch. To consider the fluctuations involved in a transition at different wavevectors, it is probably more physical to consider a single continuous branch than a set of modes with the same eigenvector which is split over several branches. The choice of the continuous branch would also ensure that the lowest frequency phonons (or lowest energy fluctuations) at each wavevector, within the constraints of a given symmetry, are joined along a single branch. The local variable should thus be such that it represents all the modes on such a mixed branch, and will not be simply a polarisation vector $\boldsymbol{\xi}$.

Lattice Wannier functions have been proposed in the case of ferroelectric perovskites [130], to separate the periodic displacements of phonon eigenvectors into localised variables. These use a formalism similar to that of lattice dynamics, but the motion of a given ion depends on the phase factors $(\mathbf{q}\cdot\mathbf{r})$ of neighbouring unit cells, not just a single cell. This has the effect of changing the apparent polarisation vector across the branch, according to the mixing between branches that has occured. Such a variable thus describes a whole continuous branch as observed. They may be constructed to reproduce the observed eigenvectors at certain points of the Brillouin Zone, particularly those at the zone centre or boundary which participate in the relevant transition. However, the exact nature of the variable is not required to be known, as long as all the modes lie on one branch of the dispersion curve.

Chapter 4 showed that the method used in this thesis to calculate phonon modes

is restricted to finding modes only at selected points in the Brillouin Zone, so it is not possible to unambiguously identify which modes lie on the same branches. The phonons presented in Chapter 5 were instead compared with full dispersion curves obtained empirically [92], in terms of degeneracy and approximate ordering. This suggested that the mode at Γ which is related to the M_2 rotational mode is that dominated by Mg displacement, which is stable in the tetragonal phase. It is therefore assumed that both these modes lie on a single branch, which may be represented by a local variable as required for application of the model described earlier in this Chapter.

In addition to allowing application of the model, this assumption has interesting consequences for the nature of the motion during the transition. As a phase transition is approached from $T > T_c$, fluctuations are initially relatively short-range and localised, involving contributions from phonons at all wavevectors: the Mg modes assumed to lie on the same branch as the rotational mode must therefore be included. Fluctuations involving Mg would in any case have lower energy than distortions of the octahedra. As longer-range fluctuations occur, and long-range order increases, smaller sections of the Brillouin Zone are involved, so the nature of the ionic motion would be expected to change.

7.4.2 Monte Carlo simulations

The simple model in (7.2) was parametrised for the M_2 instability between the tetragonal and orthorhombic phases at zero pressure, using the frequency of this branch at Γ and M (2.59 and $5.63i$ THz respectively) and the energy difference between the phases (0.235 eV per formula unit)[2]. This parametrisation gave $\alpha = -6.27$ THz, $\beta = 269$ THz^2/eV, $J = 3.21$ THz, and $x_0 = 0.172$ (eV/THz)$^{1/2}$. Since energy is measured in eV, the Boltzmann constant takes the value $k_B = 8.617 \times 10^{-5}$ K/eV. The simple approximation of using only the total energy difference, according to (7.1), gives the

[2]These values were calculated using the more recent set of pseudopotentials which was found to give equilibrium structures closer to the LAPW values; the original set of potentials were used in the published work [10].

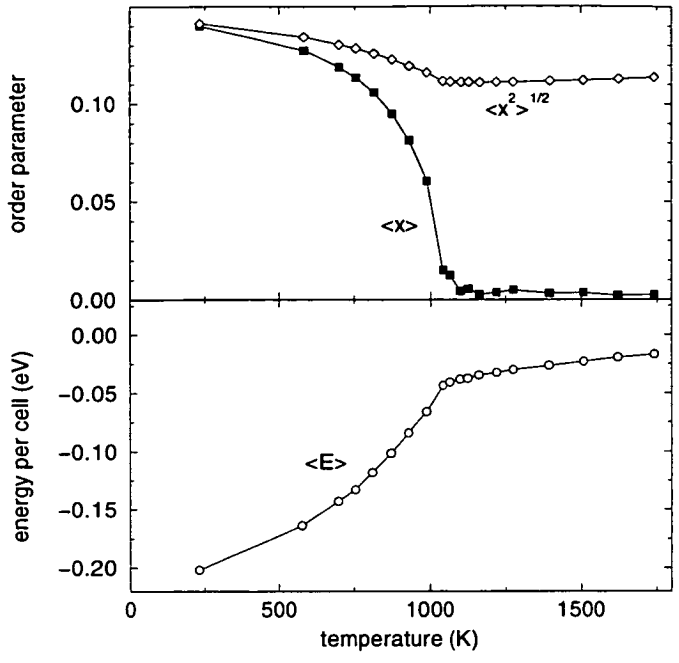


Figure 7.6. Monte Carlo simulation of the model given in (7.2) parametrised for the orthorhombic–tetragonal transition in $MgSiO_3$ perovskite. A 16×16×16 grid of elements was used, and the average energy E, order parameter $|\langle x \rangle|$ and r.m.s. order parameter $\langle x^2 \rangle^{1/2}$ averaged over 5000 steps after convergence are shown. Error bars are no larger than symbols.

temperature as $T_c = 2730$ K, which is above the melting temperature at zero pressure.

A value of $\alpha/3J$ of -0.65 sets this model between the order-disorder and displacive regimes. As noted in Chapter 5, the dispersion curve assumed to relate these two frequencies is not entirely unstable, reflecting the fact that the local potential is not particularly strong when compared to the interaction. A strong interaction is to be expected considering the rigid unit interpretation: any modes distorting the octahedra, as must be the case for modes at Γ on the same branch as rotational modes, will have much greater frequency than the rigid unit modes.

A Monte Carlo simulation of this three-dimensional, one variable system was performed ; the results are shown in Figure 7.6. The transition temperature was found to be $T_c = 1080 \pm 50$ K, which is much lower than suggested by the *ab initio* MD simulations of Chapter 5.

The ratio $k_B T_c/E_0$ was found to be 0.40; the ratio E_J/E_0 is 1.18 so $k_B T_c/E_J = 0.33$. In this case, the difference between using E_J and E_0 to estimate the transition temperature is not large, compared to the ratio of the transition temperature to either of these energies. This temperature is lower than found for model systems in Section 7.3.2; this is ascribed to the effects of the anharmonicity β, which is not yet fully understood.

7.5 Discussion

The application of a simple single-well model to the results of *ab initio* studies of soft-mode phase transitions has been discussed, with particular reference to the $MgSiO_3$ perovskite system, with the aim of estimating the transition temperature.

The results obtained from parametrising a single variable model do not appear to agree well with those from the *ab initio* molecular dynamics simulations of Chapter 5; the transition temperature obtained from the model is at least a factor of 1.8 smaller than suggested by the first-principles MD study. There are several factors which may contribute to this large discrepancy.

Firstly, the *ab initio* MD simulation comprises only four formula units (and so four 'local variables') with periodic boundary conditions, so cannot properly represent the long-range fluctuations inherent in temperature-induced phase transitions. The parametrised model uses a much larger grid, so would be expected to be more reliable in this respect.

The parametrised model also has limitations. Only one variable is assigned to each site in the model, by assuming that the mode containing octahedral rotation lies on an isolated branch of the dispersion curve. In reality, the dispersion curves involve a certain amount of degeneracy and thus one mode may lie on several branches, so this assumption is not always justified. More sophisticated parametrisation schemes from full first-principles dispersion curves have taken this into account [127].

Furthermore, it is known from both the static and dynamic calculations in Chapter 5

that four modes are involved in the transition between tetragonal and orthorhombic phases, so focussing on the most unstable may not be sufficient. Since the transition temperature appears to be closely related to the energy of interaction and thus to the interaction parameter J, the choice of the band and thus the dispersion which determines J may have a great effect on the predicted T_c. Using the lowest energy modes of the same symmetry, i.e. taking account of mixing, would be expected to give a low temperature, consistent with the additional degrees of freedom which allow mode mixing. Ignoring mixing and using branches with the same polarisation vectors would give a higher temperature. The true transition temperature may lie somewhere between these two extremes, if mixing between several branches is considered.

In addition, the ratio found in the model of transition temperature to either total energy or interaction energy between phases when parametrised from the *ab initio* data does not correspond well with that found for more general models. This may be due to differences in the anharmonicity of the model, the effects of which are not yet well understood; further work may reveal the reasons for this discrepancy.

7.6 Conclusions

The simple model described here reproduces some features of the soft-mode phase transitions which were deduced from *ab initio* calculations in Chapter 5. However, the values of the transition temperature obtained from first-principles molecular dynamics studies of a small cell, Monte Carlo simulations of a large system with the parametrised model, and estimates based on energy differences do not give consistent results. Each of these methods makes different approximations and is limited by different factors, so none may be taken as definitive.

We thus may only conclude that the true transition temperature is likely to lie between the values of 1080 K and 2700 K; the transition temperature could only be identified as being above 1800 K from the *ab initio* MD simulations. This does not

exclude the possibility of a phase transition to the tetragonal phase within the mantle, although equivalent studies at higher pressures should be performed. Further work should be able to more tightly bound the transition temperature and give a better understanding of the relationships between these methods.

Chapter 8

Zirconia

'And if, saith he, the Masters of Ships and Pilots will take the pains... diligently and faithfully to set down in severall columns, not onely the Rumb they goe on and the measure of the Ships way in degrees, & the observation of Latitude and variation of their compass; but alsoe their conjectures and reason of their correction they make of the aberrations they shall find... I doubt not but that there shall be in convenient time, brought to light many necessary precepts wch may tend to ye perfecting of Navigation...'

Isaac Newton, quoting Oughtred on Navigation
Vol. III of the *Correspondence*

8.1 Introduction

Zirconia (ZrO_2) is of great importance in the ceramics industry, having many technological applications including that of 'transformation toughening' when small inclusions of tetragonal zirconia are introduced into ceramics. Under the stresses induced by a crack tip, a transformation to a less dense monoclinic phase fills the empty space in the crack, thus inhibiting further propagation [131, 132]. Despite both theoretical and experimental study, its exact phase behaviour is still not fully known or understood [133, 134].

The stability of the various phases as determined experimentally has been reviewed elsewhere [135, 136, 137]. Cubic zirconia adopts the fluorite O_h^5 ($Fm3m$) structure,

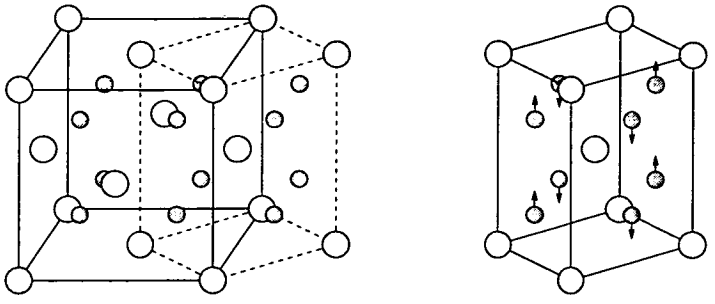


Figure 8.1. Cubic and tetragonal cells of zirconia. The dotted lines show the tetragonal unit cell within that of the cubic; the arrows indicate the oxygen distortion adopted in the tetragonal phase.

shown in Figure 8.1; experimentally, this phase is stable at ambient pressure above 2600 K. A related tetragonal unit cell may be formed, shown as dotted lines in Figure 8.1: within this a displacement of the oxygen ions generates the tetragonal D_{4h}^{15} ($P4_2/nmc$) phase, stable above 1400 K. Below this temperature, the monoclinic C_{2h}^5 ($P2_1/c$) phase is generally understood to be stable, although recent X-ray diffraction results [138] suggest the possibility that it is unstable. Some workers instead quote 2370 K for the cubic–tetragonal transition and 1170 K for the tetragonal–monoclinic phase [136]; higher pressures reduce the transition temperatures [138] so it is possible that these values originate from experiments under pressure. Under increasing pressure, the monoclinic phase transforms into a variety of orthorhombic phases, which are not addressed in this thesis. The tetragonal and cubic phases may be partially or fully stabilised by the addition of impurities, but these are not studied in this work since extremely large supercells would be required.

Theoretical studies have already been performed on various phases. The potential-induced breathing (PIB) model was used [132] to calculate equations of state for several phases, but significant discrepancies from experimental measurements were found, the monoclinic phase did not have the expected stability, and a negative bulk modulus was predicted for the tetragonal phase. These unsatisfactory results were attributed

to either the lack of non-spherical charge distribution in the PIB model, or the fact that real samples may have disorder in the oxygen positions and stabilising impurities. The Hartree-Fock method (see Section 2.1.2), which allows a more general charge distribution, was applied to the cubic and tetragonal phases [137] using effective core potentials (equivalent to pseudopotentials) and an atomic orbital basis set, but the complete energy surface was not explored due to the computational resources required. The FLAPW[1] method has also been used to study the stability of cubic and tetragonal phases [136, 139], but the forces and hence the phonon modes may not easily be found using this method. More recently, the stabilisation of cubic zirconia with CaO or MgO impurities has been studied with Hartree-Fock methods [140].

Study of the monoclinic phase is very computationally intensive, since there are 12 atoms in the unit cell, and the low symmetry dictates that a larger number of **k**-points are required. Neither a PIB model [132] nor the Hartree-Fock approach [140] gave the monoclinic phase as the most stable at zero pressure, even when the structure was optimised. FLAPW studies were not extended to the monoclinic phase because of the large computational resources required [139].

In this Chapter, density-functional calculations of the cubic, tetragonal and monoclinic phases of pure zirconia, using the CASTEP and CETEP codes, are presented, with the aim of identifying the mechanisms of transitions between the phases. Phonon calculations as described in Chapter 4 have already been proposed as being important [136, 139]. However, they have not been extensively performed by previous workers: without plane-wave basis sets, force calculations are very complicated, and simpler models (such as PIB) do not even always produce the expected equilibrium properties.

The oxygen pseudopotential found to give the best performance from the studies of $MgSiO_3$ (as described in Section 5.7 and Appendix A) was used. Zirconium pseudopotentials were generated using the Q_c tuning method [30]; two such potentials were tested, as described below. An energy cutoff of 1000 eV was used to ensure a good

[1] Full-potential linearised augmented plane-wave

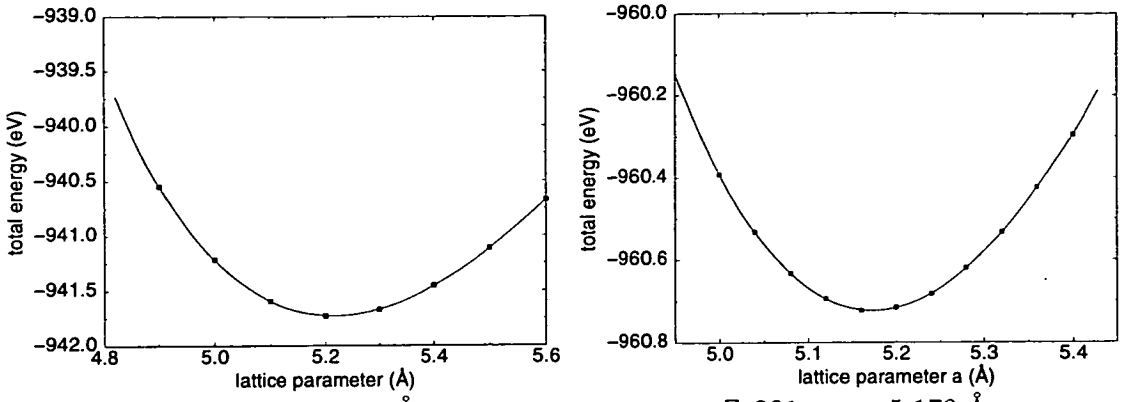
convergence of the total energy to within 0.02 eV of the fully converged value. The difference in lattice parameter between using Monkhorst-Pack k-point grids of 2×2×2 and 3×3×3 in the primitive cubic cell was less than 0.1%; the difference in energy was less than 0.1 eV. Other properties such as phonon frequencies were found to require higher densities of k-point grids. The local density approximation was used throughout.

8.2 Cubic phase

The structure of the cubic phase is determined by only the lattice parameter a, so the most efficient way to determine the equilibrium structure is from a simple set of total energies at different values of a, rather than via relaxation under the calculated stress. The results for both pseudopotentials are shown in Figure 8.2.

The equilibrium lattice parameters and bulk moduli using these pseudopotentials and a range of other methods which have been used are shown in Table 8.1. It should be noted that the neutron diffraction results at room temperature [135] are for an artificially stabilised phase, $Zr_{0.875}Mg_{0.125}O_{1.875}$, which found a_0=5.086 Å at 295 K [135], since the pure cubic phase is unstable under these conditions. The high-temperature neutron scattering results [138] give a much larger lattice parameter, due to thermal expansion. Experimental results may therefore not be compared directly to theoretical predictions, but of course act as a useful guide. The FLAPW results would be expected to be the most reliable of the theoretical results, since they are affected only by the local density approximation. In the light of these comparisons, the Zr001 potential was used for rest of the work presented here.

It is unusual for the lattice parameter to be overestimated by calculations using the local density approximation (LDA), which generally underestimates the volume. However, choice of pseudopotentials may have a non-negligible effect on lattice parameters, as found for the Mg and O pseudopotentials in $MgSiO_3$ and MgO (Chapter 5) which may be greater than the effects of the LDA.



Zr000: $a_0 = 5.214$ Å Zr001: $a_0 = 5.173$ Å

Figure 8.2. Murnaghan fit to total energy versus lattice parameter for cubic zirconia using two different pseudopotentials for zirconium.

8.3 Cubic elastic constants

The elastic constants of a phase may be deduced from *ab initio* simulations by introducing small strains and calculating the resulting stresses from first principles [12, 13]. A symmetric strain which takes the unit cell boxmatrix h to $h' = (I + \epsilon)h$ may be written

$$\epsilon = \begin{pmatrix} \epsilon_{xx} & \epsilon_{xy} & \epsilon_{xz} \\ \epsilon_{xy} & \epsilon_{xx} & \epsilon_{yz} \\ \epsilon_{xz} & \epsilon_{yz} & \epsilon_{zz} \end{pmatrix} \tag{8.1}$$

or in the Voigt notation, so that the elastic constants may be written as c_{ij}, the strain is $e = (\epsilon_{xx}, \epsilon_{yy}, \epsilon_{zz}, 2\epsilon_{xz}, 2\epsilon_{xz}, 2\epsilon_{xy})$. A cubic phase has only three elastic constants: c_{11}, c_{12} and c_{44}. A shear strain $e = (\epsilon_1, 0, 0, 2\epsilon_4, 0, 0)$ then results in a stress $\tau = (c_{11}\epsilon_1, c_{12}\epsilon_1, c_{12}\epsilon_1, 2c_{44}\epsilon_4, 0, 0)$. The elastic constants may therefore easily be deduced from the stress on a distorted structure calculated from first principles, or from the energy difference from the equilibrium structure. The results are shown in Table 8.1, together with those calculated in other studies [136, 137] and determined experimentally. The elastic moduli presented here are in good agreement with the PIB calculations, but there is a wide variation in values from different techniques. The

cubic ZrO_2	a_0	elastic constants (GPa)			
	(Å)	B	c_{11}	c_{12}	c_{44}
This work, Zr000	5.214	157			
This work, Zr001	5.173	266	519	147	183
FLAPW [136]	5.054	230±70	500±10	90±20	
Hartree-Fock [140]	5.154				
Hartree-Fock [137]	5.035	222	628	19	82
PIB [132]	5.101	288	560	153	180
Neutron expts [135]	5.090†				
Neutron, 2690 K [138]	5.269				
Elasticity expts [141]		194	417	82	47

† Extrapolated to zero temperature [140]

Table 8.1. Equilibrium lattice parameter a_0 and bulk modulus B for cubic ZrO_2, from both theoretical predictions and experimental data.

experimental results are extrapolated to correspond to pure zirconia, but were taken at temperatures above 300 K and may also be affected by disorder: this is expected to particularly affect the value of c_{44} via the internal strain on the oxygen positions away from sites of high symmetry [132].

8.3.1 Phonons in the cubic phase

The tetragonal supercell shown in Figure 8.1 is used for phonon calculations (with cubic structural parameters) since it corresponds to the unit cell of the tetragonal phase. Phonons at the Γ and R points of the cubic Brillouin Zone are then accessible, due to the restrictions of periodic boundary conditions (Chapter 4). A Monkhorst-Pack grid of 4×4×4 **k**-points generated between six and sixteen **k**-points in the reduced Brillouin Zone according to the displacement made. Four displacements were required: one zirconium ion and one oxygen ion were both moved along x and z by 0.02% of the cell dimension. This number of calculations was necessary because the shape of the supercell breaks the cubic symmetry, in terms of the symmetry elements available to construct the dynamical matrix. The results are given in Table 8.2: nearly degenerate triplets at Γ may still be observed, although only xy pairs are exactly degenerate.

mode rank		frequency ν (THz)	Zr A	Zr B	O AI	O AII	O BI	O BII
Γ								
2,3	xy	16.30	.00	.00	.50	-.50	.50	-.50
4	z	16.23	.00	.00	.50	-.50	.50	-.50
10,11	xy	7.47	-.36	-.36	.43	.43	.43	.43
12	z	7.46	-.36	-.36	.43	.43	.43	.43
15,16	xy	0.00170	.61	.61	.25	.25	.25	.25
17	z	0.00072	.61	.61	.25	.25	.25	.25
R								
1	z	19.36	.00	.00	-.50	.50	.50	-.50
5,6	xy	15.80	-.09	.09	-.49	-.49	.49	.49
7	z	10.59	.71	-.71	.00	.00	.00	.00
8,9	xy	9.752	.00	.00	.50	-.50	-.50	.50
13,14	xy	3.624	.70	-.70	.07	.07	-.07	-.07
18	z	5.556i	.00	.00	-.50	-.50	.50	.50

Table 8.2. Frequencies and eigenvectors of the phonons in cubic zirconia, at Γ and R, at $a = 5.214$ Å, ranked according to frequency. Eigenvectors are given in mass-reduced coordinates, along the direction shown in the second column (only x and y displacements form exact degenerate pairs due to supercell geometry).

The degeneracy splitting and non-zero frequency of the acoustic modes at Γ gives an indication of the errors arising from anharmonicity.

The eigenvectors of these modes are also obtained. One unstable mode is found, at R, consisting of the displacement pattern of oxygen ions shown in Figure 8.1, which on freezing in generates the tetragonal phase. The cubic–tetragonal system thus displays the potential for a soft-mode transition, in that the structural phase transition is manifested as an unstable mode in the higher-symmetry structure, together with a small change in lattice vectors. The highest frequency modes involve oxygen displacement only.

8.4 Strain instability from cubic to tetragonal

The transition to the tetragonal phase involves not only a distortion of the ions following the phonon described above, but also a small change in the strain. This may be better

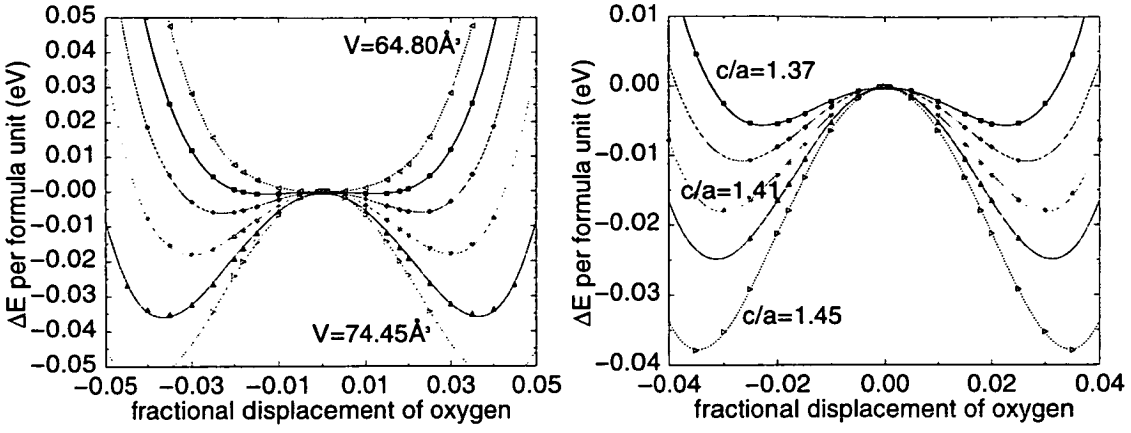


Figure 8.3. Energy difference on freezing in phonon as a function of phonon amplitude d for a variety of ratios c/a (left) and volumes (for two formula units) at $c/a = \sqrt{2}$ (right).

described by building up the energy surface as a function of the lattice parameters and the oxygen displacement. At a number of different values of c and a, displacements of the oxygen atoms were made, following the unstable phonon eigenvector, of different amplitudes d.

First indications of the effects of strain on the phonon instability may be seen from the results shown in Figure 8.3, in which $\Delta E = E(d) - E(0)$ is plotted for various c/a at constant volume and for various volumes at $c/a = \sqrt{2}$. Better fits were obtained than to FLAPW results [139]. It can be seen that the phonon instability decreases with decreasing c/a ratio, and with decreasing volume, as might be expected for a phonon with displacements along z. Below some value of c/a (which will be less than $\sqrt{2}$), the phonon becomes stable, as it is seen to do by the lowest volume. At these points, the 'tetragonal' phase would have zero phonon amplitude and is therefore indistinguishable from the strained cubic phase.

From a larger number of simulations, a polynomial least squares fit for $E(a, c, d)$ was made, of the form

$$E(a, c, d) = E_0(a, c) + \alpha(a, c)\, d^2 + \beta(a, c)\, d^4 \tag{8.2}$$

		E_0	$\alpha/10^2$	$\beta/10^4$
a^2	f_{20}	21.4610	7.77932	4.08268
ac	f_{11}	6.72990	19.7619	-116.306
c^2	f_{02}	4.19013	-11.9056	66.9145
a	f_{10}	-191.860	-164.064	570.224
c	f_{01}	-67.9866	47.7810	-272.959
1	f_{00}	-1394.61	192.775	-313.910

Table 8.3. Parameters of fit for $E(a, c, d)$ according to (8.2) and (8.3). Units are such that d is dimensionless (fractional coordinates), a and c are in Å and energy is in eV.

where $E_0(a, c)$, $\alpha(a, c)$ and $\beta(a, c)$ are all expanded as harmonic functions of a and c:

$$f(a, c) = a^2 f_{20} + acf_{11} + c^2 f_{02} + af_{10} + cf_{01} + f_{00} \tag{8.3}$$

This gives a total of eighteen parameters, which were fitted from 159 data points (with variance 2.0×10^{-4} and standard deviation 1.2×10^{-3}); the parameters are given in Table 8.3. The complete four-dimensional surface cannot easily be visualised, but a section at $a = 3.660$ is shown in Figure 8.4: the saddle point corresponding to the cubic phase and stable minima of the tetragonal configuration can be clearly seen. It is again shown that the phonon instability increases with c. The monoclinic phase is not manifested by a more stable minimum, simply because the symmetry has been constrained to be at least that of the tetragonal phase.

8.4.1 Properties of cubic and tetragonal phases

With d constrained to be zero, the minimum energy of the surface $E(a, c)$ was found, giving c/a within 0.01% of $\sqrt{2}$. This minimum thus represents a cubic phase, and has an equilibrium lattice parameter of 5.174 Å, compared to 5.173 Å obtained directly. The cubic elastic constants given in Table 8.1 were also obtained using this function. The bulk modulus was found to be $B = 271$ GPa (compared to 266 GPa directly). The parametrised function thus provides a satisfactory representation of the data around

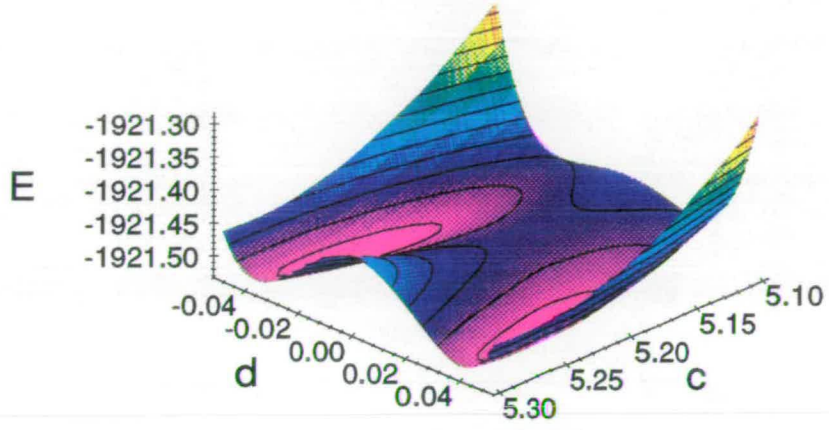


Figure 8.4. Energy as a function of lattice parameter c and oxygen distortion d, for $a = 3.660$ (cubic value). The central saddle point corresponds to the cubic phase; the two minima at non-zero d correspond to equivalent settings of the tetragonal phase.

the cubic phase.

When $\partial E/\partial d = 0$ is solved to give the harmonic well $E(a, c)$, the tetragonal minimum ($\partial E/\partial a = \partial E/\partial c = 0$) is found at $\{a = 3.662,\ b = 5.231\}$, and positional parameter $d = 0.034$. The volume is thus 1.3% larger than in the cubic phase. The energy difference between cubic and tetragonal phases is only 0.0425 eV, equivalent to 494 K. This is much lower than the difference between the temperatures at which the tetragonal and cubic phases become stable at zero pressure, but as has been seen in previous Chapters, the energy difference does not necessarily accurately determine the temperature required for the transition. This has also been discussed in FLAPW studies of this mode [139], but only for a single well of $E(d)$ at experimental values of a and c.

The energy surface $E(a, c)$ can also be used to deduce the optimum c/a ratio for a given volume, and hence the energy-volume (EV) curve for the tetragonal phase; even with the relatively simple parametrisation, only numerical solutions were possible. The value of d at each volume may then also be obtained. Figure 8.5 shows the EV curves for both the cubic and tetragonal phases, and the values of c/a and d. It may be seen that below a volume of Ω_c=66 Å^3, the tetragonal and cubic EV curves are indistinguishable:

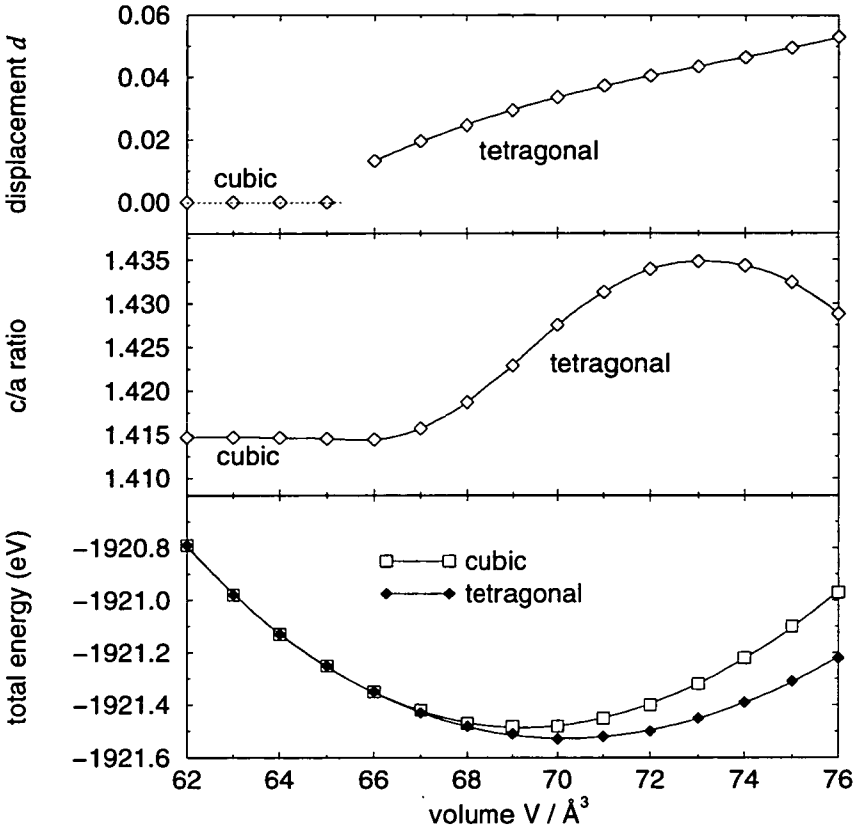


Figure 8.5. Total energies, c/a ratio of lattice parameters and oxygen displacement d in cubic and tetragonal phases of ZrO_2, deduced from the parametrised surface $E(a, c, d)$. At volumes below 66 Å^3, the tetragonal phase adopts the cubic structure.

this is because the soft phonon has stabilised, so that freezing in this mode no longer reduces the energy from that of the cubic phase. The value of d which minimises the energy function is imaginary beyond this point, so to correspond to observable phases, d is taken to be zero. Below this volume, the c/a ratio thus has the perfect cubic value of $\sqrt{2}$, and the structure is cubic.

The bulk modulus of the tetragonal phase was found to be 213 GPa, with $\partial B/\partial P = 1.7$. An experimental value of 173 GPa has been reported [132]. Calculations using the potential induced breathing (PIB) model [132] found a negative bulk modulus for the tetragonal phase at all pressures investigated; the tetragonal structure only stabilised

after it adopted the cubic structure (at around 84 Å^3). The PIB model does not, however, allow a non-spherical charge distribution. The volume of 66 Å^3 found here for the transition to the cubic phase occurs at a pressure of 14 GPa, according to the fitted polynomial equation of state. High pressure studies [138] have found that at temperatures at which the tetragonal phase is stable at zero pressure, increasing the pressure leads first to a transition to the cubic, as is found here. The phase diagram at high temperatures and pressure is not well-known, but studies of the monoclinic–orthorhombic phase boundaries [142] suggest that the higher temperature phase changes from tetragonal to cubic somewhere between 6 and 18 GPa. As the temperature increases towards 2700 K, the tetragonal–cubic transition pressure reduces to zero.

8.4.2 Phonons in the tetragonal phase

Phonons of a tetragonal structure (a=b=3.660Å, c=5.215Å) were calculated using the tetragonal primitive cell (two formula units), differing only from the cubic phonon supercell in c/a ratio and oxygen positions. The set of phonons obtained must therefore use the same eigenvector basis set as the cubic phonons, i.e. contain only linear combinations thereof. The frequencies and eigenvectors are given in Table 8.4; the closest eigenvector matches with the cubic modes are also given, obtained in the same way as in Chapter 5.

As would be expected in a soft-mode transition, the mode which is unstable in the cubic structure becomes stable, and is even of higher frequency than some other tetragonal phonons. This stabilisation is similar to that observed when freezing in unstable phonons in $MgSiO_3$ perovskite (Chapter 5), and corresponds to stable oscillation around the bottom of the double wells evident in Figure 8.3.

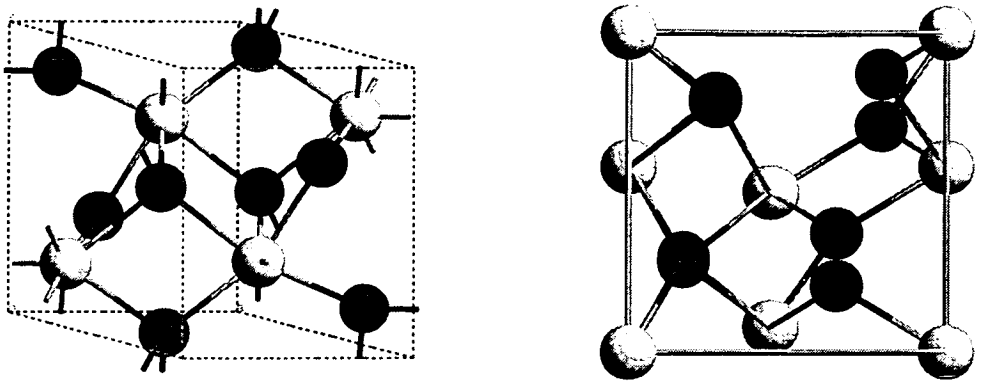
cubic		frequency	Zr		O			
match	mode	ν (THz)	A	B	AI	AII	BI	BII
1	z	19.58	.00	.00	.50	-.50	-.50	.50
2+3	xy	19.26	-.11	.11	-.03	.70	-.70	.03
4	z	17.38	.18	-.18	.48	-.48	.48	-.48
5+6	xy	13.57	-.01	.01	-.71	-.04	.03	.71
8+9	xy	13.52	-.23	-.23	-.11	.66	.66	-.11
12	z	10.06	-.36	-.36	.43	.43	.43	.43
7	z	9.785	-.68	.68	.12	-.12	.12	-.12
18	z	6.841	.00	.00	-.50	-.50	.50	.50
10+11	xy	4.950	-.28	-.28	.65	.01	.01	.65
13+14	xy	3.975	.70	-.70	-.02	.11	-.11	.02
15+16	xy	0.0806	.61	.61	.25	.25	.25	.25
17	z	0.0153	.61	.61	.25	.25	.25	.25

Table 8.4. Frequencies and eigenvectors of the phonons in tetragonal zirconia, at Γ and R, for a=b=3.660 Å, c=5.215 Å. Eigenvectors are given in mass-reduced coordinates, along the direction shown (x and y displacements form degenerate pairs). The first column gives the identity of the cubic phonon with the closest eigenvector.

8.5 Shear distortion to monoclinic

The monoclinic phase was described by Simha [143] as a shear distortion of the tetragonal phase, with a unit cell containing four formula units, which maintained the Zr ions on face centred sites. Relaxation of the structure from first principles was initially attempted from this assumption, but repeatedly tended to a tetragonal phase (zero shear distortion), suggesting that this monoclinic phase is not stable with respect to the tetragonal phase.

However, neutron powder diffraction studies [135] concluded that the monoclinic structure instead has space group $P2_1/c$, as shown in Figure 8.6. The standard setting of this cell is not face centred, but by a shift of origin a cell may be generated which has Zr ions at $(0,0,0)$, $(\frac{1}{2}-\eta_x, \frac{1}{2}, \eta_z)$, $(\frac{1}{2}-\eta_x, \eta_y, \frac{1}{2}+\eta_z)$ and $(0, \frac{1}{2}-\eta_y, \frac{1}{2})$, and which may therefore be seen to be related to a face-centred structure. The experimentally deduced structures both have all η_α ($\alpha = x, y, z$) smaller than 0.1; their values are, of course, linearly related to the fractional positions Zr_x, Zr_y and Zr_z. The oxygen ions



$P2_1/c$ unit cell alternative cell

Figure 8.6. Structure of monoclinic zirconia with $P2_1/c$ space group, showing seven-fold coordinated zirconium ions and three-fold (type OI) and four-fold (OII) coordinated oxygen ions. Sticks connecting ions are to show coordination only. The alternative cell, projected down the c axis, shows the deviation from a face-centred structure.

are similarly displaced from the ideal positions of the higher-symmetry structures, by up to 0.17 (Figure 8.6).

The cell-doubling with respect to the tetragonal phase means that only phonons at Γ and M of the tetragonal Brillouin Zone may contribute to the ionic structural parameters of this phase, in the way considered in Chapter 5. It can be seen from Figure 8.6 that the ionic displacements from a sheared tetragonal structure do not follow either a single zone-centre or zone-boundary mode of the tetragonal phase, so combinations of modes at both Γ and M must be present.

Several studies have suggested that the monoclinic phase may not in fact be stable with respect to the tetragonal and cubic phases. X-ray diffraction studies found that the monoclinic phase had a very low bulk modulus [138], from which a possible instability was proposed. A pressure of only 4 GPa was required for a phase transition to an orthorhombic phase at zero temperature [142]. Calculations using the PIB model found the monoclinic phase to be stable only at negative pressures [132]; twinning and symmetry breaking in real samples were proposed as stabilising mechanisms. The sign

of the energy difference between cubic and monoclinic phases has been found to depend on the theoretical technique used: Hartree-Fock calculations found the monoclinic phase to be less stable than the cubic by 5 meV, but semi-empirical and shell model methods found it to be 0.2–0.3 eV more stable [140].

8.5.1 Equilibrium monoclinic phase

Relaxation of this structure using the first-principles forces and stresses was very slow, partly because of the the large size of the system, and partly because of the very shallow energy surface explored. The stresses were relaxed to below 2×10^{-3} eV/Å^3, and the forces to below 0.05 eV/Å. A volume of 148.5 Å^3 was found, compared to 140.8 Å^3 from neutron scattering at 295 K [135]; the 5% overestimate in volume is slightly smaller than that in the cubic phase (7.5%). The ionic parameters found are given in Table 8.5, and have a r.m.s. deviation of 0.0044 from the mean experimental positions (i.e. those averaged between the neutron and X-ray results). The residual forces were largest on the O$_I$ ions, whose positions may be seen to deviate from experiment by the greatest amount; further relaxation would be likely to improve the agreement with experimental data but would be very time-consuming.

Each Zr is seven-fold coordinated by oxygen, as shown in Figure 8.6, compared to the eightfold coordination in the cubic and tetragonal phases: the type O$_I$ oxygen ions have only three close Zr neighbours rather than the more usual four. The mean Zr–O bondlength was found to be 1.9% longer than experimentally observed. This is consistent with the general overestimation of bondlengths found using these pseudopotentials.

The energy of the monoclinic phase was found to be 0.0107 eV per formula unit higher than the cubic phase. This is inconsistent with the order of phases found experimentally with temperature, but as discussed above, there has been evidence that this phase is not as stable as previously thought. However, the energy differences are very small, so further work to confirm that calculations are absolutely converged should

monoclinic ZrO_2		this work	experimental data (300 K) neutron [135]	X-ray [144]
a (Å)		5.239	5.151	5.145
b (Å)		5.336	5.212	5.208
c (Å)		5.383	5.317	5.311
β		99.65°	99.23°	99.23°
V (Å^3)		148.4	140.8	140.4
Zr	x	0.2792	0.2754	0.2758
	y	0.0429	0.0395	0.0411
	z	0.2103	0.2083	0.2082
	η_x	0.0584	0.0508	0.0516
	η_y	0.0858	0.0790	0.0822
	η_z	0.0796	0.0834	0.0836
O I	x	0.0744	0.0700	0.0703
	y	0.3430	0.3317	0.3359
	z	0.3380	0.3447	0.3406
O II	x	0.4483	0.4496	0.4423
	y	0.7583	0.7569	0.7549
	z	0.4833	0.4792	0.4789
⟨Zr–O⟩ (Å)		2.197	2.159	2.159

Table 8.5. Structural parameters for monoclinic ZrO_2 (baddeleyite), compared to previous neutron [135] and single-crystal X-ray [144] experimental data.

be performed. Furthermore, it may be that the experimentally observed phase has a doubled cell with respect to that used here, or has a small amount of stabilising disorder.

8.5.2 Shear distortions from tetragonal phase

The tetragonal phase, having point group D_{4h}, has six independent non-zero elastic moduli: c_{11}, c_{33}, c_{12}, c_{13}, c_{44} and c_{66}. Softening of both tetragonal c_{44} and c_{66} have been proposed as mechanisms for the tetragonal–monoclinic transition on the basis of crystallographic symmetry analysis [143, 145], although only the c_{44} possibility was discussed in earlier work [131]. The elastic moduli have been previously calculated with the PIB model [132], and coupling between the shear moduli and the ionic positions was discussed.

The elastic constants c_{11}, c_{33}, c_{12} and c_{13} must all be positive, since the equilibrium tetragonal phase is stable with respect to small changes in a and c, and the bulk modulus was found to be positive. In order to calculate the remaining moduli, c_{44} and c_{66}, the tetragonal phase was distorted with xy and xz shears respectively, and the moduli again deduced from the *ab initio* stresses. If there is any phonon-strain coupling (see, for example, Section 6.6.1), the ionic positions should be re-relaxed in the distorted phase, to correspond to the compensation which would occur in real samples.

Without relaxation of ions, both the shear distortions were stable: $c_{44} = 116$ GPa and $c_{66} = 156$ GPa. These results were obtained with shear strains of $\mathcal{O}(10^{-3})$, where errors in c_{ij} from anharmonicity were less than 1%. The ϵ_{xy} strain caused only small forces on the oxygen ions along z, coupling only to the amplitude of the oxygen mode already frozen in: relaxation of these ions had negligible effect on the calculated stress. However, the ϵ_{xz} strain induced large forces on the ions along x, showing considerable coupling between c_{44} and ionic positions [132]. The ions were thus relaxed until the forces were below 0.01 eV/Å^3, converging the energy to 10^{-7}eV, and reducing the stress so that c_{44}=48 GPa. This is much softer than the shear modulus of the cubic phase, reflecting the extra ionic degrees of freedom which are available. The reported experimental values are c_{44}=101 GPa and c_{66}=156 GPa [132].

The displacement of the ions from their cubic fractional positions is considerable in the monoclinic phase, so these results suggest that c_{44} softening will be more relevant in a transition to monoclinic; this modulus is much softer than the c_{66}, so would probably give the lowest energy barrier. However, no elastic instabilities (negative elastic moduli) have been found.

8.6 Conclusions and suggestions for further work

An unstable phonon in the high symmetry cubic phase is again found to be related to a structural phase transition, in this case to a tetragonal phase. The energy surface

representing cubic and tetragonal phases has been parametrised in terms of the strain and the phonon amplitude. Both the temperature- and pressure-induced mechanisms for the cubic-tetragonal transition may be identified from this energy surface.

The transition to the monoclinic phase involves both more complex ionic rearrangements and a strain including shear components. A soft shear modulus has been found in the tetragonal phase, which is coupled to significant ionic distortion. However, the zero-pressure monoclinic phase was not found to be more stable than the cubic and tetragonal phases, within the accuracy of the calculation, so no direct instability is expected. The monoclinic structure should be simulated at other pressures to further investigate its stability. None of the zone-centre phonons of the tetragonal phase were found to be dynamically unstable. Calculations of phonons at the M point (zone boundary) of the tetragonal Brillouin Zone would be needed to identify any phonon modes which contribute to this transition, and determine whether any of them have low frequency. If the tetragonal phase is stable with respect to both phonons and elastic moduli, then some other mechanism for a transition between monoclinic and tetragonal phases must be sought.

Other phases induced by increasing pressure, which are mostly orthorhombic, have not been studied here, and further calculations which include these should elucidate the mechanisms by which these occur.

Chapter 9

Summary

I never know, when I'm telling a story, whether to cut the thing down to plain facts or whether to...shove in a lot of atmosphere.
P.G. Wodehouse, *The inimitable Jeeves*

In the work described in this thesis, the widely used CASTEP and CETEP plane-wave density-functional theory codes have been extended and used to investigate the mechanisms and conditions required for a number of structural phase transitions. Full structural relaxation is now possible, with a first-order correction for the effects of using a finite number of plane waves in the wavefunction basis set; this also allows determination of crystal structures under external pressures.

The use of Hellmann-Feynman forces calculated from first principles to give the normal modes of a crystal has been described, and a practical scheme developed, which exploits any symmetry. Phonons at specific points of the Brillouin Zone may be calculated, according to the periodic boundary conditions employed. The precautions which should be taken during simulations to avoid excessive noise and error in the results have been discussed. Phonon frequencies calculated by this methods have generally been found to be accurate to within a few percent, but have the one significant restriction of not giving any splitting between longitudinal and transverse phonons at the zone centre. It is hoped to overcome this problem in the future.

Eigenvectors of the calculated phonons have been interpreted in $MgSiO_3$, ZrO_2 and other systems in terms of their role in structural phase transitions. In particular, unstable modes indicate a mechanism by which a distortion may lower the energy of a crystal, precipitating a transformation into a lower symmetry phase. The phase transitions of $MgSiO_3$ have been extensively analysed in terms of the contributions from different modes.

The phonon eigenvectors may also be used to identify modes which are candidates for rigid unit behaviour, and first-principles simulations have been performed to determine whether phonons at constant volume or rigid unit modes (RUMs) which decrease the volume are better descriptions of modes in $MgSiO_3$, $BaZrO_3$ and SiO_2. Although in none of these cases was the RUM model found to be more accurate than the phonon representation, the consequences of any extent of RUM behaviour were found to describe many of the structural changes occuring in these systems.

The most accurate method for estimation of transition temperatures from *ab initio* calculations is not yet determined. A series of first-principles molecular dynamics simulations of twenty atoms of $MgSiO_3$ suggested that the transition temperature for a tetragonal-orthorhombic transition is higher than the melting point, and thus that this transition would not occur in the mantle. However, such simulations are severely hampered by a finite system size. In contrast, a simple model parametrised from *ab initio* data gives a much lower transition temperature, but uses only a single variable, despite it being clear that several phonon modes are involved in the transition.

Throughout the work described in this thesis, it has been found that the understanding of the factors affecting the transferability of pseudopotentials is still incomplete, and pseudopotentials that perform well for one system may not produce acceptable results in others. The comparison between theoretical and experimental results that this necessitates is somewhat contrary to the spirit of *ab initio* calculations. However, it is hoped that identification of these problems may help to identify more rigorous tests for pseudopotentials in solid-state calculations.

Appendix A

Pseudopotential specifications

The generation and use of pseudopotentials has unfortunately been described as something of a black art [30]. There are many pseudopotentials in circulation whose generation or expected range of transferability is not precisely known. To clarify which pseudopotentials have been employed, and the issues behind their transferability, those used or investigated in this work are described here.

The pseudopotentials used were generated either by J.S. Lin [36, 146] or, using the Q_c tuning method, by M.H. Lee [30, 37], at the Theory of Condensed Matter Group at the University of Cambridge.

A.1 Description of terms

Several values are output from the `CASTEP` and `CETEP` codes during initialisation and are therefore useful for unambiguous identification of the potentials used during a simulation. `ICHARG`, or Z, is that charge which is left remaining on the core, equal to the number of valence electrons which are considered. For some species, especially those with d electrons, it is advantageous to use more than the highest shell of valence electrons. `PSCORE` is the core potential, i.e. the $G = 0$ term which remains after subtraction of the Coulomb terms; `PSCALE` is a scale factor for each component of

the pseudopotential. Zero values of PSCALE indicate components which are taken to be local. The electronic configurations and core radii used to generate the potentials are given where known. As explained in Section 2.3, angular momentum components which are not present in the ground state are usually generated from ionised excited states to ensure that these states are bound.

A.2 Silicon

Name	Z	PSCORE	PSCALE			generation
			$l=0$	l=1	l=2	
Si012	4	112.017	0	-0.8927	-4.5432	*sp*: $3s^2 3p^2$; *d*: $3s^1 3p^{0.75} 3d^{0.25}$ r_c=0.74 Å; M.H. Lee

A.3 Magnesium

Name	Z	PSCORE	PSCALE			generation
		0	$l=0$	$l=1$	$l=2$	
Set #1	2	141.84	0	-1.150	-0.559	J.S. Lin
Set #2	2	-48.921	3.29	2.18	0	J.S. Lin
Mg006 (#3)	2	116.658	0	0	0	$3s^2$, r_c=1.05 Å; M.H. Lee

Section 5.7 discusses the problems of generating pseudopotentials for Group II elements, and the results of using these three pseudopotentials for MgO and $MgSiO_3$. Some suggestions were made as to why the fully local potential Mg006 generates the most satisfactory results, despite being generated from only states of the neutral atom.

A.4 Oxygen

Name	Z	PSCORE	PSCALE			generation
		0	$l=0$	$l=1$	$l=2$	
Set #1	6	14.648	20.80	-9.697	0	J.S.Lin, for O^{2-}
O020c (#2)	6	4.571	31.63	0	0	*sp*: $2s^2 2p^4$; *d*: $2s^1 2p^{1.75} 3d^{0.25}$ r_c=0.95 Å; M.H. Lee
O020 (#3)	6	-28.347	37.47	4.71	2.44	*sp*: $2s^2 2p^4$; *d*: $2s^1 2p^{1.75} 3d^{0.25}$ r_c=0.95 Å; M.H. Lee

The use of these potentials for MgO and $MgSiO_3$ is also described in Section 5.7. The second and third potentials are based on the same electronic configuration, but use different Q_c tuning parameters [30]. The second potential was found to give the most satisfactory results in these cases.

A.5 Zirconium

Name	Z	PSCORE	PSCALE			generation
		0	$l=0$	$l=1$	$l=2$	
Zr000	4	-268.153	0	6.064	-10.269	M.H. Lee
Zr001	4	329.814	0	0.596	-41.415	*sd*: $4d^2 5s^2$; *p*: $4d^2 5s^{0.75} 5p^{0.25}$
						r_c=1.43 Å; M.H. Lee

The Zr001 potential was used in most of the work in Chapter 8. A Kerker construction was used for the s component.

A.6 Barium

Name	Z	PSCORE	PSCALE			generation
		0	$l=0$	$l=1$	$l=2$	
Ba013	10	-208.404	-2.461	0	0	M.H. Lee
Ba015Rg	10	183.483	1.963	0	1.2843	*spd*: $5s^2 5p^{5.75} 5d^{0.25} 6s^2$
						r_c=1.43 Å; M.H. Lee

A value of 10 for Z is due to the n=5 shell being included in the effective valence configuration, to overcome problems with the radius of the normal core electrons being too large, and to give better d components in the pseudopotential (see Chapter 6).

Appendix B

Symmetry of eigenvalues in diamond

It was commented in Section 4.5.1 that when a nearest-neighbour model for interatomic interactions is used in the diamond structure, a dispersion curve is obtained in which values of ν^2 fall in symmetric pairs around $\nu_0^2/2$, where ν_0 is the frequency of the optic modes at Γ. The dispersion curve with squared frequencies is shown in Figure B.1. This means that the Fourier transformed dynamical matrix has pairs of eigenvalues of the form $\lambda_0 \pm \chi$. It is attempted here to show why this is so.

B.1 Form of the Fourier transformed dynamical matrix

If it is assumed that only the self-interaction and nearest-neighbour interaction force constants are non-zero, only these elements need to be considered when forming the Fourier-transformed dynamical matrix $D_{\alpha\beta}(\kappa\kappa'|\mathbf{q})$ from (4.7). The atom label κ may take only two values, since there are only two atoms in the primitive unit cell.

The diamond structure has cubic symmetry, so the self-interaction force constant is purely diagonal:

$$\Phi_{\alpha\beta}\begin{pmatrix}0 & 0\\ \kappa & \kappa\end{pmatrix} = \Lambda\delta_{\alpha\beta} \tag{B.1}$$

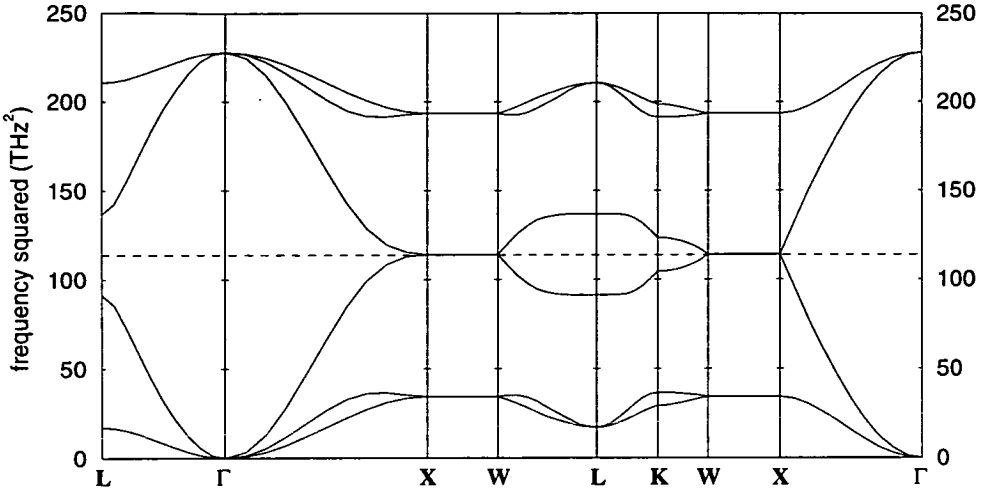


Figure B.1. Dispersion relations ν^2 (THz^2) against $\mathbf{q}$ for FCC silicon, assuming only nearest neighbour central interaction. Force constants are taken from *ab initio* calculations in a 64 atom supercell. The values of ν^2 are symmetric about the dotted line.

where Λ is a constant. There are four nearest neighbours, at the corners of a tetrahedron centred at the origin. The four neighbours lie in four different unit cells. Since the positions $\mathbf{x}_\kappa$ of the atoms within these cells may be freely chosen, the cell origins may be placed at each at the 'neighbour' atoms, and the vector connecting atom κ with one of its neighbours of type κ' written as $\mathbf{x}_l(\kappa\kappa')$; for convenience, cartesian axes for α, β and γ are used. There are then only two different numerical values of the force constant, so that for $\alpha = \beta$

$$\Phi_{\alpha\alpha}\begin{pmatrix} 0 & l \\ \kappa & \kappa' \end{pmatrix} = \Psi \tag{B.2}$$

where Ψ is a constant. The sign on the force constant for $\alpha \neq \beta$ is determined by the direction of $\mathbf{x}_l(\kappa\kappa')$ relative to α and β, so

$$\Phi_{\alpha\beta}\begin{pmatrix} 0 & l \\ \kappa & \kappa' \end{pmatrix} = \Theta[\mathbf{x}_l(\kappa\kappa')]_\alpha[\mathbf{x}_l(\kappa\kappa')]_\beta \tag{B.3}$$

where Θ is a constant, and $[\mathbf{x}_l(\kappa\kappa')]_\alpha$ gives the direction cosine: the magnitudes of all

$[\mathbf{x}_l]_\alpha$ are equal. It can immediately be seen that

$$\Phi_{\beta\alpha}\begin{pmatrix}0 & l\\ \kappa & \kappa'\end{pmatrix} = \Phi_{\alpha\beta}\begin{pmatrix}0 & l\\ \kappa & \kappa'\end{pmatrix}. \tag{B.4}$$

Furthermore, on exchanging κ and κ' a set of neighbours is found at inverted vectors, i.e. we can match pairs such that $\mathbf{x}_l(\kappa\kappa') = -\mathbf{x}_{l'}(\kappa'\kappa)$, so we also have

$$\Phi_{\alpha\beta}\begin{pmatrix}0 & l'\\ \kappa' & \kappa\end{pmatrix} = \Phi_{\alpha\beta}\begin{pmatrix}0 & l\\ \kappa & \kappa'\end{pmatrix}. \tag{B.5}$$

The Fourier transformed dynamical matrix is given by (4.7); either $[\mathbf{x}_l - \mathbf{x}_0]$ or $\mathbf{x}\begin{pmatrix}l\\ \kappa'\end{pmatrix} - \mathbf{x}\begin{pmatrix}0\\ \kappa\end{pmatrix}$ may be used, simply by accepting a phase factor into the eigenvector. Here we take

$$D_{\alpha\beta}(\kappa\kappa'|\mathbf{q}) = \sum_l \Phi_{\alpha\beta}\begin{pmatrix}0 & l\\ \kappa & \kappa'\end{pmatrix}\exp(i\mathbf{q}.[\mathbf{x}_l(\kappa\kappa')]) \tag{B.6}$$

and it is assumed that $\mathbf{x}_0 = 0$. For $\kappa = \kappa'$, only the self-interaction term is required:

$$D_{\alpha\beta}(\kappa\kappa|\mathbf{q}) = \Lambda\delta_{\alpha\beta} \tag{B.7}$$

For $\kappa \neq \kappa'$, but $\alpha = \beta$, we have

$$D_{\alpha\alpha}(\kappa\kappa'|\mathbf{q}) = \Psi\sum_l \exp(i\mathbf{q}.\mathbf{x}_l(\kappa\kappa')) = [D_{\alpha\alpha}(\kappa'\kappa|\mathbf{q})]^* \tag{B.8}$$

because the pairs may be matched according to (B.5). Similarly, for $\alpha \neq \beta$

$$D_{\alpha\beta}(\kappa\kappa'|\mathbf{q}) = \Theta\sum_l [\mathbf{x}_l(\kappa\kappa')]_\alpha[\mathbf{x}_l(\kappa\kappa')]_\beta\exp(i\mathbf{q}.\mathbf{x}_l(\kappa\kappa')) \tag{B.9}$$

so that

$$D_{\beta\alpha}(\kappa\kappa'|\mathbf{q}) = D_{\alpha\beta}(\kappa\kappa'|\mathbf{q}) \tag{B.10}$$

and also

$$D_{\alpha\beta}(\kappa'\kappa|\mathbf{q}) = [D_{\alpha\beta}(\kappa\kappa'|\mathbf{q})]^*. \tag{B.11}$$

However, for general $\mathbf{q}$ it may be noted that $D_{\alpha\beta}(\kappa'\kappa|\mathbf{q}) \neq D_{\alpha\gamma}(\kappa'\kappa|\mathbf{q})$.

B.2 Implications for eigenvalues

We thus find that $\mathbf{D}(\mathbf{q})$ now takes the form

$$\mathbf{D}(\mathbf{q}) = \left(\begin{array}{ccc|ccc} \Lambda & 0 & 0 & A^* & c^* & b^* \\ 0 & \Lambda & 0 & c^* & A^* & a^* \\ 0 & 0 & \Lambda & b^* & a^* & A^* \\ \hline A & c & b & \Lambda & 0 & 0 \\ c & A & a & 0 & \Lambda & 0 \\ b & a & A & 0 & 0 & \Lambda \end{array}\right) \tag{B.12}$$

where A, a, b, and c are all in general complex constants. However, the submatrix

$$\mathbf{D}(\kappa\kappa'|\mathbf{q}) = \left(\begin{array}{ccc} A & c & b \\ c & A & a \\ b & a & A \end{array}\right)$$

may be diagonalised in a similarity transformation which does not affect that for $\kappa = \kappa'$. It is symmetric but complex, and in general will have complex eigenvalues, say χ_1, χ_2 and χ_3. The eigenvalues of (B.12) must therefore be identical to those of

$$\left(\begin{array}{ccc|ccc} \Lambda & 0 & 0 & \chi_1^* & 0 & 0 \\ 0 & \Lambda & 0 & 0 & \chi_2^* & 0 \\ 0 & 0 & \Lambda & 0 & 0 & \chi_3^* \\ \hline \chi_1 & 0 & 0 & \Lambda & 0 & 0 \\ 0 & \chi_2 & 0 & 0 & \Lambda & 0 \\ 0 & 0 & \chi_3 & 0 & 0 & \Lambda \end{array}\right)$$

and this may be further transformed into an Hermitian matrix with 2×2 blocks along the diagonal, so that it may be seen that its eigenvalues are

$$\lambda = \Lambda \pm |\chi_1|, \Lambda \pm |\chi_2|, \Lambda \pm |\chi_3| \tag{B.13}$$

as observed.

In the case that $\mathbf{q} = 0$, it is found that $a = b = c = 0$ and $A = -\Lambda$, so that the sum rule of (4.16) is satisfied; then $\chi_1 = \chi_2 = \chi_3 = -\Lambda$ and the overall eigenvalues are therefore triply degenerate, either 2Λ or zero.

Appendix C

Rotation-strain coupling

The coupling between rotation of polyhedra and changes in the lattice parameters was discussed in Section 6.3. However, the CRUSH package [99] for detecting rigid unit modes (RUMs) omits this coupling, with the results that RUMs that are expected to have zero frequencies may instead be calculated to have finite frequencies. This Appendix presents a specific example of a system in which neglecting the strain coupling produces a stable mode, but including the neglected terms recovers a rigid unit mode with zero frequency. The result for the RUM may be trivially obtained directly, but it may be helpful to clarify the definitions of a phonon and a RUM used in this thesis.

In the discussion in Chapter 6, the full second derivative of the energy $E(\theta, a)$ with respect to the amplitude θ of a simple rotational mode in the perovskite structure was given as

$$\frac{d^2E}{d\theta^2} = \frac{\partial^2 E}{\partial \theta^2} + 2\frac{\partial^2 E}{\partial \theta\, \partial a}\frac{da}{d\theta} + \frac{\partial^2 E}{\partial a^2}\left(\frac{da}{d\theta}\right)^2 + \frac{\partial E}{\partial a}\frac{d^2a}{d\theta^2} \tag{6.4}$$

and it is from this full derivative that the frequency of the mode should be calculated, using $\omega^2 \propto d^2E/d\theta^2$. Coupling with the strain is included via $a(\theta)$, i.e. any variation of the lattice parameter a with the rotation. In a pure harmonic phonon mode, the lattice parameter is constant, so both first and second derivatives of $a(\theta)$ are zero. Therefore only the first term on the right-hand side, the partial second derivative, contributes to

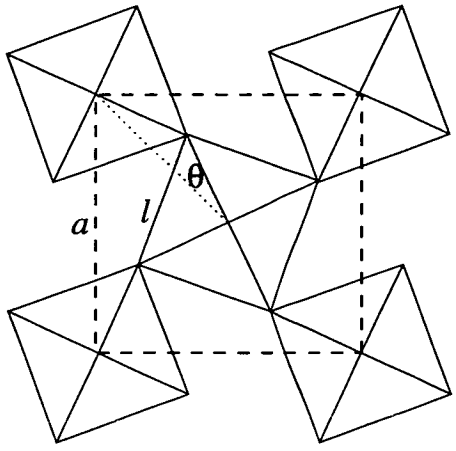


Figure C.1. Variables used to describe the M_2 rotational mode in the perovskite structure: the bondlength l, the cell parameter a and the rotation angle θ.

the frequency of a phonon mode. Since CRUSH does not allow strain coupling, this is also the term corresponding to frequencies calculated thereby.

We examine here each of the terms in (6.4) for a simple model of the perovskite M_2 rotational mode discussed previously. Only the rotational mode drawn in Figure C.1 (equivalent to Figure 6.1) is considered, in terms of the amplitude θ, and the lattice parameter a perpendicular to the rotation axis. Bond angles within the octahedra are still constrained to their ideal values but, as in CRUSH, rotation around the linking oxygen ions costs zero energy and therefore may occur freely. No other distortions are allowed as degrees of freedom. Instead of the CRUSH 'split atom' model for the energy, we use a more realistic bond-stretching model. Unlike in CRUSH, the O–O bondlengths are free to change, with some spring constant, whilst the octahedra are constrained to remain corner-sharing. This model is qualitatively equivalent to that of CRUSH (and will be shown to be identical for small oscillations around the undistorted phase), but is somewhat more physical and includes coupling to the strain in a natural way.

The lattice parameter a is related to the length l of an O–O bond by

$$a = 2l\cos\theta \tag{C.1}$$

and if the energy is defined as only depending on the stretching of these bonds from their ideal value l_0, this gives (to within some arbitrary factor)

$$E = 4(l - l_0)^2 \tag{C.2}$$

$$= \left(\frac{a}{\cos\theta} - a_0\right)^2 \tag{C.3}$$

where $a_0 = 2l_0$, thus defining $E(\theta, a)$. In the CRUSH split-atom methodology, in which the bondlengths are fixed to be l_0, changing the lattice parameter would simply affect the distance separating the atoms, giving an energy of the form

$$E_{\text{CRUSH}} = k_C\left(\sqrt{2}l_0(1 - \cos\theta) + \frac{a - a_0}{\sqrt{2}}\right)^2$$

$$= \frac{k_C}{2}(a - a_0\cos\theta)^2. \tag{C.4}$$

We thus see that this differs from (C.3) by only a factor of $\cos^2\theta$, so will be exactly equal at $\theta = 0$ and behave similarly elsewhere. However, the bond-stretching model is retained to demonstrate that the conclusions reached are not an artefact of the CRUSH split-atom proposition. The three second derivatives of the energy may then be derived from (C.3):

$$\frac{\partial^2 E}{\partial\theta^2} = -2\frac{a\left(-3a + 2a\cos^2\theta + 2a_0\cos\theta - a_0\cos^3\theta\right)}{\cos^4\theta}, \tag{C.5a}$$

$$\frac{\partial^2 E}{\partial\theta\,\partial a} = 2\frac{\sin\theta(2a - a_0\cos\theta)}{\cos^3\theta}, \tag{C.5b}$$

$$\text{and } \frac{\partial^2 E}{\partial a^2} = 2\frac{1}{\cos^2\theta}. \tag{C.5c}$$

The cross derivative $\partial^2 E/\partial\theta\,\partial a$ is negative for $\theta < 0$ and positive for $\theta > 0$. The other two second derivatives are positive for all θ, except for $\partial^2 E/\partial\theta^2$ at $\theta = 0$, where it is zero. Although modes are usually only calculated in equilibrium structures, for which

$\partial E/\partial a = 0$, this term may be written in the general case as

$$\frac{\partial E}{\partial a} = 2\frac{a - a_0\cos\theta}{\cos^2\theta}. \tag{C.6}$$

We will be interested either in rigid unit modes, for which $a(\theta)$ must be determined such that the bondlengths are maintained, or pure phonons, in which $a(\theta) = a_0$. If the mode is taken to be a phonon, and a_0 is set to unity for simplicity, the frequency will be calculated from only the first term, since all derivatives of a are zero:

$$\frac{\partial^2 E_P}{\partial\theta^2} = -2\frac{(-3 + 2\cos^2\theta + 2\cos\theta - \cos^3\theta)}{\cos^4\theta}. \tag{C.7}$$

This is only zero at $\theta = 0$, as shown in Figure C.2. This is the result found from CRUSH (Figure 6.4): introducing a finite amplitude of a rotational mode produces a non-zero frequency.

However, in the rigid unit mode case, the rotation-strain coupling is required to take the form

$$\begin{aligned} l \equiv \frac{a}{2\cos\theta} &= l_0 \equiv \frac{a_0}{2} \\ \Rightarrow a_R &= a_0\cos(\theta) \end{aligned} \tag{C.8}$$

as has been quoted several times in Chapter 6. Substituting (C.8) into (C.5), and again using $a_0 = 1$, gives

$$\frac{\partial^2 E_R}{\partial\theta^2} = 2\frac{1 - \cos^2\theta}{\cos^2\theta}, \tag{C.9a}$$

$$\frac{\partial^2 E_R}{\partial\theta\,\partial a} = 2\frac{\sin\theta}{\cos^2\theta}, \tag{C.9b}$$

$$\text{and } \frac{\partial^2 E_R}{\partial a^2} = 2\frac{1}{\cos^2\theta}. \tag{C.9c}$$

The corresponding term to that used to calculate the phonon is the first, given in (C.9a), and is compared with the phonon result in Figure C.2. It may be seen to have

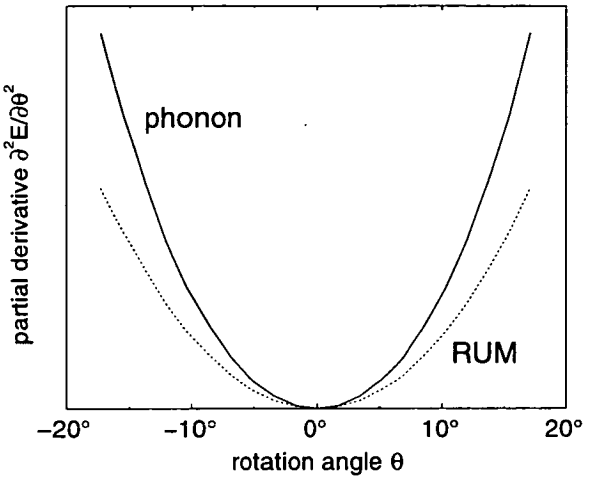


Figure C.2. Partial second derivative of energy with respect to rotation angle in the phonon (C.9a) and RUM (C.7) cases. In the phonon case, this gives the frequency; in the RUM case, this term is balanced by others to give a zero frequency. The units of energy are arbitrary in this model.

similar behaviour to the phonon term, although it has smaller magnitude.

To calculate the full second derivative, the first and second derivatives of $a_R(\theta)$ are also required. In this model, they are trivially $-\sin\theta$ and $-\cos\theta$ respectively. It can be seen that $dE_R/da = 0$ from (C.6) and (C.8), so the final term of (6.4) may be neglected. It is then found that in the case of a rigid unit mode the first three terms are related:

$$\begin{aligned} \frac{\partial^2 E_R}{\partial\theta^2} &= 2\frac{1-\cos^2\theta}{\cos^2(\theta)} \\ &= -\frac{1}{2}\frac{\partial^2 E_R}{\partial\theta\,\partial a}\frac{da_R}{d\theta} = \frac{\partial^2 E_R}{\partial a^2}\left(\frac{da_R}{d\theta}\right)^2 \end{aligned} \tag{C.10}$$

and so the full second derivative of (6.4) is zero for all θ for a rigid unit mode. This is the result anticipated from assigning zero energy to bond bending terms. It is the term linear in $da/d\theta$, and involving the cross derivative of E_R, which cancels out the other two.

We have thus demonstrated that for non-zero mode amplitudes of this simple system, the mode only has zero frequency when strain coupling is included, as has been stated in this thesis and elsewhere [99].

Bibliography

[1] A.D. Bruce. *Adv. Phys.*, 29:111–217, 1980.

[2] A.F. Voter. *MRS Bulletin*, 21:17–19, 1996.

[3] P. Hohenberg and W. Kohn. *Phys. Rev.*, 136:B864–871, 1964.

[4] W. Kohn and L.J. Sham. *Phys. Rev.*, 140:A1133–1138, 1965.

[5] R.O. Jones and O. Gunnarsson. *Rev. Mod. Phys.*, 61:689–745, 1989.

[6] W. Cochran. *Phys. Rev. Lett.*, 3:412–414, 1959.

[7] W. Cochran. *The Dynamics of Atoms in Crystals*, volume 3 of *The Structure and properties of solids*. Edward Arnold, 1973.

[8] E.J. Samuelsen, E. Andersen, and J. Feder, editors. *Structural Phase Transitions and Soft Modes*. Universitetsforlaget, 1971.

[9] R.M. Wentzcovitch. *Phys. Rev. B*, 50:10358–10361, 1994.

[10] M.C. Warren and G.J. Ackland. *Phys. Chem. Minerals*, 23:107–118, 1996.

[11] H.C. Hsueh, M.C. Warren, H. Vass, G.J. Ackland, S.J. Clark, and J. Crain. *Phys. Rev. B*, 53:14806–14817, 1996.

[12] B.B. Karki, L. Stixrude, S.J. Clark, M.C. Warren, G.J. Ackland, and J Crain. *Am. Mineral.*, 82:51–60, 1997.

[13] B.B. Karki, S.J. Clark, M.C. Warren, H.C. Hsueh, G.J. Ackland, and J. Crain. *J. Phys.: Cond. Mat.*, 9:375–380, 1997.

[14] A.K.A. Pryde, K.D. Hammonds, M.T. Dove, V. Heine, J.D. Gale, and M.C. Warren. *J. Phys: Cond. Mat.*, 8:10973–10982, 1996.

[15] B.B. Karki, M.C. Warren, L. Stixrude, G.J. Ackland, and J. Crain. *Phys. Rev. B*, 55:3465–3471, 1997.

[16] A.K.A. Pryde, K.D. Hammonds, M.T. Dove, V. Heine, J.D. Gale, and M.C. Warren. *Phase Transitions*. In press.

[17] G.J. Ackland and M.C. Warren. *Phase Transitions*. In press.

[18] B.B. Karki, L. Stixrude, S.J. Clark, M.C. Warren, G.J. Ackland, and J. Crain. *Am. Mineral.* In press.

[19] M.C. Payne, M.P. Teter, D.C. Allan, T.A. Arias, and J.D. Joannopoulos. *Rev. Mod. Phys.*, 64:1045–1097, 1992.

[20] T. Oguchi and T. Sasaki. *Prog. Theor. Phys. Supp.*, 103:93–117, 1991.

[21] M. Schluter and L.J. Sham. *Physics Today*, 35:36–43, 1982.

[22] J.P. Perdew and Alex Zunger. *Phys. Rev. B*, 23:5048–5079, 1981.

[23] D.M. Ceperley and B.J. Alder. *Phys. Rev. Lett.*, 45:566–569, 1980.

[24] A.A. Maradudin, E.W. Montroll, G.H. Weiss, and I.P. Ipatova. *The Theory of Lattice Dynamics in the Harmonic Approximation*, volume 3 of *Solid State Physics Supplement.* Academic Press, second edition, 1971.

[25] D.J. Chadi and M.L. Cohen. *Phys. Rev. B*, 8:5747–5753, 1973.

[26] H.J. Monkhorst and J.D. Pack. *Phys. Rev. B*, 13:5188–5192, 1976.

[27] P.J.H. Denteneer and W. van Haeringen. *J. Phys. C.:Solid State Phys.*, 18:4127–4142, 1985.

[28] S.J. Clark, C.J. Adam, G.J. Ackland, J. White, and J. Crain. *Liq. Cryst.*, 1996. In press.

[29] E. Holzschuh. *Phys. Rev. B*, 28:7346–7348, 1983.

[30] M-H. Lee. *Advanced Pseudopotentials.* PhD thesis, The University of Cambridge, 1995.

[31] G.P. Francis and M.C. Payne. *J. Phys.: Condens. Matter*, 2:4395–4404, 1990.

[32] V. Heine. *Solid State Physics*, 24:1–36, 1970.

[33] D.R. Hamann, M. Schlüter, and C. Chiang. *Phys. Rev. Lett.*, 43:1494–1497, 1979.

[34] L. Kleinman and D.M. Bylander. *Phys. Rev. Lett.*, 48:1425–1428, 1982.

[35] A.M. Rappe, K.M. Rabe, E. Kaxiras, and J.D. Joannopoulos. *Phys. Rev. B*, 41:1227–1230, 1990.

[36] J.S. Lin, A. Qteish, M.C. Payne, and Heine V. *Phys. Rev. B*, 47:4174–4180, 1993.

[37] M-H. Lee, J.S. Lin, M.C. Payne, V. Heine, V. Milman, and S. Crampin. Unpublished.

[38] G.B. Bachelet, D.R. Hamann, and M. Schlüter. *Phys. Rev. B*, 26:4199–4228, 1982.

[39] H-C. Hsueh, J. Maclean, G.Y. Guo, M-H. Lee, S.J. Clark, G.J. Ackland, and J. Crain. *Phys. Rev. B*, 51:12216–12222, 1995.

[40] P.P. Ewald. *Ann. Phys. (Leipzig)*, 64:253, 1921.

[41] M.T. Yin and M.L. Cohen. *Phys. Rev. B*, 26:5668–5687, 1982.

[42] M.T. Yin and M.L. Cohen. *Phys. Rev. B*, 26:3259–3273, 1982.

[43] K. Kunc and R.M. Martin. *Phys. Rev. B*, 24:2311–2314, 1981.

[44] W. Cochran and R.A. Cowley. *Handb. Phys.*, 25/2a:59–156, 1967.

[45] A.A. Maradudin and G.H Weiss. *Phys. Rev.*, 123:1968–1976, 1961.

[46] T.H.K. Barron. *Phys. Rev.*, 123:1995–1998, 1961.

[47] H.B. Rosenstock. *Phys. Rev.*, 121:416–424, 1961.

[48] N.W. Ashcroft and N.D. Mermin. *Solid State Physics*. Saunders College, Philadelphia, 1976.

[49] R. Car and M. Parrinello. *Phys. Rev. Lett.*, 55:2471–2474, 1985.

[50] X. Gonze, R. Stumpf, and M. Scheffler. *Phys. Rev. B*, 44:8503–8513, 1991.

[51] M.P. Teter, M.C. Payne, and D.C. Allan. *Phys. Rev. B*, 40:12255–12263, 1989.

[52] L.J. Clarke, I. Stich, and M.C. Payne. *Comp. Phys. Comms.*, 72:14–28, 1992.

[53] F.D. Murnaghan. *Proc. Natl. Acad. Sci. USA*, 30:244, 1944.

[54] R. Feynman. *Phys. Rev.*, 56:340–343, 1939.

[55] M. Scheffler, J.P. Vigneron, and G.B. Bachelet. *Phys. Rev. B*, 31:6541–6551, 1985.

[56] A.C. Hurley. *Proc. Roy. Soc. A*, 226:170–192, 1954.

[57] P. Pulay. *Mol. Phys.*, 17:197–204, 1969.

[58] M.T. Yin and M.L. Cohen. *Phys. Rev. Lett.*, 45:1004–1007, 1980.

[59] O.M. Nielsen and R.M. Martin. *Phys. Rev. Lett.*, 50:697–700, 1983.

[60] O.H. Nielsen and R.M. Martin. *Phys. Rev. B*, 32:3780–3791, 1985.

[61] O.H. Nielsen and R.M. Martin. *Phys. Rev. B*, 32:3792–3805, 1985.

[62] P. Gomes Dacosta, O.H. Nielsen, and K Kunc. *J. Phys. C*, 19:3163–3172, 1986.

[63] S. Froyen and M.L. Cohen. *J. Phys. C*, 19:2623–2632, 1986.

[64] J. Crain, R.O. Piltz, G.J. Ackland, S.J. Clark, M.C. Payne, V. Milman, J.S. Lin, P.D. Hatton, and Y.H. Nam. *Phys. Rev. B*, 50:8389–8401, 1994.

[65] H.C. Andersen. *J. Chem. Phys*, 72:2384–2393, 1980.

[66] M. Parrinello and A. Rahman. *Phys. Rev. Lett.*, 45:1196–1199, 1980.

[67] R.M. Wentzcovitch. *Phys. Rev. B*, 44:2358–2359, 1991.

[68] G.J. Ackland. PhD thesis, University of Oxford, 1987. Chapter 3.

[69] H.C. Hsueh, H. Vass, S.J. Clark, and J. Crain. *Europhys. Lett.*, 31:151–155, 1995.

[70] S. Wei and M.Y. Chou. *Phys. Rev. Lett.*, 69:2799–2802, 1992.

[71] G.P. Srivastava and K. Kunc. *J. Phys. C: Solid State Phys.*, 21:5087–5106, 1988.

[72] K. Kunc and P. Gomes Dacosta. *Phys. Rev. B*, 32:2010–2021, 1985.

[73] H. Wendel and R.M. Martin. *Phys. Rev. B*, 19:5251–5263, 1979.

[74] R.E. Cohen and H. Krakauer. *Phys. Rev. B*, 42:6416–6423, 1990.

[75] A.V. Postnikov, T. Neumann, and G. Borstel. *Phys. Rev. B*, 50:758–763, 1994.

[76] X. Gonze, D.C. Allan, and M.P. Teter. *Phys. Rev. Lett.*, 68:3603–3606, 1992.

[77] W. Zhong, R.D. King-Smith, and D. Vanderbilt. *Phys. Rev. Lett.*, 72:3618–3621, 1994.

[78] S. Baroni, P. Giannozzi, and A. Testa. *Phys. Rev. Lett.*, 58:1861–1864, 1987.

[79] P. Giannozzi, S. de Gironcoli, P. Pavone, and S. Baroni. *Phys. Rev. B*, 43:7231–7242, 1991.

[80] H. Bilz and W. Kress. *Phonon dispersion relations in insulators*, volume 10 of *Springer Series in Solid-State Sciences*. Springer-Verlag, 1979.

[81] P.N. Keating. *Phys. Rev.*, 145:637–645, 1966.

[82] D. Weaire and R. Alben. *Phys. Rev. Lett.*, 29:1505–1508, 1972.

[83] R. Alben, D. Weaire, Jnr Smith, J.E., and M.H. Brodsky. *Phys. Rev. B*, 11:2271–2296, 1975.

[84] C.R.A. Catlow and G.D. Price. *Nature*, 347:243–248, 1990.

[85] A. Navrotsky and D.J. Weidner, editors. *Perovskite: A Structure of Great Interest to Geophysics and Materials Science*. American Geophysical Union, 1989.

[86] L. Stixrude and R.E. Cohen. *Nature*, 364:613–616, 1993.

[87] H. Kawakatsu and F.L. Niu. *Nature*, 371:301–305, 1994.

[88] M.S.T. Bukowinski and G.H. Wolf. Equation of state and possible critical phase transitions in $MgSiO_3$ perovskite at lower-mantle conditions. In S. Ghose, J.M.D. Coey, and E. Salje, editors, *Structural and Magnetic Phase Transitions in Minerals*, pages 91–112. Springer-Verlag, New York, 1988.

[89] R. Jeanloz and S. Morris. *Ann. Rev. Earth Planet. Sci.*, 14:377–415, 1986.

[90] M.E. Lines and A.M. Glass. *Principles and Applications of Ferroelectrics and related materials.* Clarendon Press, Oxford, 1977.

[91] R.J. Hemley and R.E. Cohen. *Annu. Rev. Earth Planet. Sci*, 20:553–600, 1992.

[92] R.J. Hemley, M.D. Jackson, and R.G. Gordon. *Phys. Chem. Minerals*, 14:2–12, 1987.

[93] H.K. Mao, R.J. Hemley, Y. Fei, J.F. Shu, L.C. Chen, A.P. Jephcoat, Y. Wu, and W.A. Bassett. *J. Geophys. Res.*, 96:8069–8079, 1991.

[94] N.L. Ross and R.M. Hazen. *Phys. Chem. Minerals*, 17:228–237, 1990.

[95] Y. Wang, F. Guyot, and R.C. Liebermann. *J. Geophys. Res.*, 97:12327–12347, 1992.

[96] R.D. King-Smith and D. Vanderbilt. *Phys. Rev. B*, 49:5828–5844, 1994.

[97] G.P. Kerker. *J. Phys. C*, 13:L189–94, 1980.

[98] T. Yagi, H-K. Mao, and P.M. Bell. *Phys. Chem. Minerals*, 3:97–110, 1978.

[99] A.P. Giddy, M.T. Dove, G.S. Pawley, and V. Heine. *Acta. Cryst.*, A49:697–703, 1993.

[100] R.M. Wentzcovitch, J.L. Martins, and G.D. Price. *Phys. Rev. Lett.*, 70:3947–3950, 1993.

[101] W. Zhong, D. Vanderbilt, and K.M. Rabe. *Phys. Rev. Lett.*, 73:1861–1864, 1994.

[102] M. Matsui. *Phys. Chem. Minerals*, 16:234–238, 1988.

[103] H. Akbarzadeh, S.J. Clark, and G.J. Ackland. *J. Phys.: Condens. Matter*, 5:8065–8074, 1993.

[104] G. Kresse, J. Furthmüller, and J. Hafner. *Phys. Rev. B*, 50:13181–13185, 1994.

[105] R.D. Shannon. *Acta. Cryst.*, A32:751–767, 1976.

[106] M.J. Mehl, R.E. Cohen, and H. Krakauer. *J. Geophys. Res.*, 93:8009–8022, 1988.

[107] B.B. Karki, L. Stixrude, S.J. Clark, M.C. Warren, G.J. Ackland, and J. Crain. In preparation.

[108] H. Boysen, B. Dorner, F. Frey, and H. Grimms. *J. Phys. C: Solid St. Phys*, 13:6127–6146, 1980.

[109] M. Vallade, B. Berge, and G. Dolino. *J. Phys. (Paris) I*, 2:1481–1495, 1992.

[110] M.T. Dove, V. Heine, and K.D. Hammonds. *Min. Magazine*, 59:629–639, 1995.

[111] A.M. Glazer. *Acta. Cryst. B*, 28:3384–3392, 1972.

[112] R.J. Hemley, H-K. Mao, and E.C.T. Chao. *Phys. Chem. Minerals*, 13:285–290, 1986.

[113] C. Lee and X. Gonze. *J. Phys.: Condens. Matter*, 7:3693–3698, 1995.

[114] K. Robinson, G.V. Gibbs, and P.H. Ribbe. *Science*, 172:567–570, 1971.

[115] N.L. Ross, J-F. Shu, R.M. Hazen, and T. Gasparik. *Am. Mineral.*, 75:739–747, 1990.

[116] I.P. Swainson and M.T. Dove. *Phys. Chem. Minerals*, 22:61–65, 1885.

[117] K.D. Hammonds, M.T. Dove, A.P. Giddy, and V. Heine. *Am. Mineral*, 79:1207–1209, 1994.

[118] G.S. Pawley. *Phys. Status Solidi*, 49:475–478, 1972.

[119] P. Sollich, V. Heine, and M.T. Dove. *J. Phys.: Condens. Matter*, 6:3171–3196, 1994.

[120] B.B. Karki. Private communication.

[121] T.A. Mary, J.S.O. Evans, T. Vogt, and A.W. Sleight. *Science*, 272:90–92, 1996.

[122] R.W.G Wyckoff. *Crystal Structures*. Interscience Publishers, New York, second edition, 1963.

[123] L.R. Johnson. *Bull. Seismol. Soc. Am.*, 59:973–1008, 1990.

[124] M. Weber and J.D. Davis. *Geo. J. Inter.*, 102:231–255, 1990.

[125] C. Lee and X. Gonze. *Phys. Rev. Lett.*, 72:1686–1689, 1994.

[126] F. Kirchoff, J.M. Holender, and M.J. Gillan. *Phys. Rev. B*, 54:190–202, 1996.

[127] K.M. Rabe and U.V. Waghmare. *J. Phys. Chem. Solids*, 57:1397–1403, 1996.

[128] S. Padlewski, A.K. Evans, C. Ayling, and V. Heine. *J. Phys.: Condens. Matter*, 4:4895–4908, 1992.

[129] V. Heine, X. Chen, S. Dattagupta, M.T. Dove, A. Evans, A.P. Giddy, S. Marais, S. Padlewski, E. Salje, and F.S. Tautz. *Ferroelectrics*, 128:255–264, 1992.

[130] K.M. Rabe and U.V. Waghmare. *Phys. Rev. B*, 52:13236–13246, 1995.

[131] N. Simha and L. Truskinovsky. *Acta. metall. mater.*, 42:3827–3836, 1994.

[132] R.E. Cohen, M.E. Mehl, and L.L. Boyer. *Physica B*, 150:1–9, 1988.

[133] S. Block, J.A.H. Da Jornada, and G.J. Piermarini. *J. Am. Ceramic. Soc.*, 68:497–499, 1985.

[134] M. Yashima, K. Ohtake, M. Kakihana, H. Arashi, and M. Yoshimura. *J. Phys. Chem. Solids*, 57:17–24, 1996.

[135] C.J. Howard, R.J. Hill, and B.E. Reichert. *Acta. Cryst.*, B44:116–120, 1988.

[136] H.J.F. Jansen. *Phys. Rev. B*, 43:7267–7278, 1991.

[137] R. Orlando, C. Pisani, C. Roetti, and E. Stefanovich. *Phys. Rev. B*, 45:592–601, 1992.

[138] J.M. Leger, P.E. Tomaszewski, A. Atouf, and A.S. Pereira. *Phys. Rev. B*, 47:14075–14083, 1993.

[139] H.J.F. Jansen and J.A. Gardner. *Physica B*, 150:10–18, 1988.

[140] E.V. Stefanovich, A.L. Schluger, and C.R.A. Catlow. *Phys. Rev. B*, 49:11560–11571, 1994.

[141] H.M. Kandil, J.D. Greiner, and J.F. Smith. *J. Am. Ceram. Soc.*, 67:341–346, 1984.

[142] O. Ohtaka, T. Yamanaka, S. Kume, E. Ito, and A. Navrotsky. *J. Am. Ceram. Soc.*, 74:505–509, 1991.

[143] N.K. Simha. *Mechanics of the tetragonal to monoclinic transition of zirconia and transformation toughening.* PhD thesis, University of Minnesota, 1994.

[144] D.K. Smith and H.W. Newkirk. *Acta. Cryst.*, 18:983–991, 1965.

[145] N.K. Simha. *J. Phys. IV, Colloq. (France)*, 5:1121–1126, 1995.

[146] J.S. Lin, A. Qteish, M.C. Payne, and V. Heine. *Phys. Rev. B*, 47:4174–4180, 1993.