Ab fere initio Equations of Mechanical State

D C Swift

A thesis submitted in fulfilment of the requirements
for the degree of Doctor of Philosophy
to the
University of Edinburgh
2000

Abstract

This thesis describes the development and application of models to predict the equation of mechanical state of materials from first principles, concentrating on the regime of strong shock waves. Most effort was devoted to crystalline solids, though extensions to the fluid phase and higher temperatures are proposed. Equations of state and phase diagrams were predicted for aluminium, silicon and beryllium.

The method used is based on quantum mechanical treatments of the electrons in the solid and of the phonon modes. The importance of anharmonic effects (phonon-phonon interactions) was investigated, but was not included rigorously because it did not appear to contribute significantly. With fully *ab initio* methods, the equation of state and phase diagram could be predicted to a few percent in mass density, the discrepancy being caused mainly by the use of the local density approximation in predicting electron states.

The accuracy of the equation of state could be improved considerably by adjusting the internal energy to reproduce the observed mass density at STP. The resulting *ab fere initio* equation of state could essentially reproduce the observed states on the shock Hugoniot to within the scatter in the experimental data.

Because these equations of state are built on firm quantum mechanical and thermodynamic principles, they should allow properties to be predicted accurately away from the principal Hugoniot, unlike traditional empirical equations of state. Accurate temperatures are important in the development of models of material strength (elasticity and plasticity) based on microstructural phenomena. As an illustration of the versatility of the equations of state, hydrocode simulations were made of the splitting of a shock wave in silicon, caused by the phase change. The splitting appears to be in reasonable agreement with laser-driven shock experiments.

Acknowledgements

I would like to thank the individuals and organisations below for their help.

Graeme Ackland (University of Edinburgh) supervised this project, provided help with the practicalities of electron ground state and atomic restoring force calculations, and essential advice through many useful discussions. Justin Wark (University of Oxford) and Malcolm McMahon (University of Edinburgh) acted as examiners and suggested valuable improvements.

The 'CASTEP' electron ground state program was provided courtesy of Graeme Ackland, Mike Payne (University of Cambridge) and the 'U.K. Car-Parrinello Consortium'.

Stewart Clark (University of Edinburgh) and Michele Warren (Universities of Edinburgh and Cambridge) provided much patient and helpful advice on the use of symmetry operations in reducing the number of force calculations required. Their help included the use of computer programs.

Art Ruoff (Cornell University) kindly made available his diamond anvil cell measurements on the aluminium isotherm.

Bard Bennett (Los Alamos National Laboratory) provided a copy of the INFERNO atom-in-jellium code, along with advice on its use.

Allan Hauer (Los Alamos National Laboratory) made available the results of the transient X-ray diffraction experiments on the 'TRIDENT' laser. He and Justin Wark (University of Oxford) provided many useful discussions.

Ian Gray (AWE Aldermaston) gave encouragement and advice, and arranged for funding to support the development of the Dirac-based electron ground state program.

Intensive calculations were performed on computers at AWE Aldermaston, the University of Edinburgh and Fluid Gravity Engineering Ltd.

This work was partly funded by the U.K. Ministry of Defence through AWE Aldermaston and by Fluid Gravity Engineering Ltd.

Contents

Abstract i

Acknowledgements iii

Glossary xiii

1 Introduction 1

1.1 The equation of mechanical state 2
1.2 Why calculate equations of state? 3
1.3 Scope of this work 3
1.4 New aspects 4

2 Theoretical equations of state 7

2.1 Structure of an equation of state 8
  2.1.1 Internal energy of a solid phase 9
  2.1.2 Internal energy of a fluid phase 14
  2.1.3 Completing a single-phase equation of state 21
  2.1.4 Polymorphism 22
  2.1.5 Non-equilibrium polymorphism in macroscopic simulations 31
  2.1.6 Macroscopic model of non-equilibrium vibrations 34
2.2 Electron ground states 37
  2.2.1 Dirac band structure 38
  2.2.2 CASTEP 57
  2.2.3 INFERNO 59
2.3 Thermal motion of the atoms 61
  2.3.1 Momentum-space methods: phonons 61
  2.3.2 Atomic motion 73
2.4 Thermal excitation of the electrons 75

2.4.1 Electron band structure 75
2.4.2 Electron-thermal energy 76
2.4.3 Low temperature behaviour 78
2.4.4 Quantum free electron model 78
2.5 Correcting inadequacies in *ab initio* calculations 81
2.6 Summary of method for predicting equations of state 84

3 Application to a selection of materials **89**
3.1 Aluminium 90
3.1.1 INFERNO isotherms 90
3.1.2 CASTEP cold curve 92
3.1.3 Lattice-thermal contribution 98
3.1.4 Electron-thermal contribution 108
3.1.5 Equations of state 111
3.1.6 Comparison with mechanical data 112
3.2 Silicon 116
3.2.1 Diamond structure 116
3.2.2 Body-centred tetragonal structure 124
3.2.3 Diamond/BCT equation of state 134
3.3 Beryllium 137
3.3.1 HCP 137
3.3.2 BCC 141
3.3.3 FCC 143
3.3.4 Equilibrium phase diagram 143
3.3.5 Density at STP 144
3.3.6 Shock Hugoniot 145

4 Use in hydrodynamic simulations **153**
4.1 1D hydrocode 154
4.1.1 Continuum equations 154
4.1.2 Discrete representation 155
4.1.3 Numerical integration 156
4.1.4 Treatment of shock waves 158
4.1.5 Time step constraints 158
4.1.6 Software implementation 159
4.1.7 Validation 160

4.2 Double shock structure in silicon 164
  4.2.1 Stability of shock waves in polymorphic materials 164
  4.2.2 Equilibrium / isotropic predictions for silicon 166
  4.2.3 Comparison with transient X-ray diffraction data 167

5 Conclusions **172**

A Units **176**

B Crystallographic structures **178**
B.1 Introduction 178
B.2 Structures 178

C Interatomic potentials **180**
C.1 Introduction 180
C.2 Forms of potential 181
  C.2.1 Inverse power 181
  C.2.2 Morse 181
  C.2.3 Finnis-Sinclair 182
C.3 Interatomic pair potentials from bulk properties 183
  C.3.1 Structure factors 186
  C.3.2 Using derivatives of the specific energy 187
  C.3.3 Orthogonalised fitting functions 189
  C.3.4 Conventions for structure factors 190

D Direct simulation of ensembles of atoms **193**
D.1 Introduction 193
D.2 'Molecular' dynamics 194
D.3 Monte-Carlo 196
D.4 Faster calculations 198

E Crystallographic diffraction patterns **200**
E.1 Introduction 200
E.2 Scattering from a single atom 201
E.3 Coherent scattering from a crystal 201
E.4 Software implementation 203

F Evaluating against experiment 204

F.1 Introduction 204
F.2 Phonons 204
F.3 Fermi surface 204
F.4 Equilibrium density 204
F.5 Isotherms 205
F.6 Shock Hugoniots 205

G Empirical equations of state **209**
G.1 Introduction 209
G.2 Grüneisen 209
G.3 Murnaghan 210
G.4 Rose 210
G.5 Bushman et al 211
G.6 Schulte and Holzapfel 211

H Gram-Schmidt Orthogonalisation **212**
H.1 Introduction 212
H.2 Scalar product and orthogonal functions 212
H.3 Gram – Schmidt method 213
H.4 Modified method 214

Bibliography **216**

List of Figures

2.1 Schematic of the 0 K isotherm indicating regions of instability. . . 13
2.2 Equilibrium mixture of two non-interacting phases. 25
2.3 Directionality of energy barrier to phase transition. The shape of the curve alters with ρ and T. 35
2.4 Schematic of the liquid – vapour region. 36
2.5 Simple scheme for correcting for the effect of perturbations in the position of image atoms. 67
2.6 Statistical noise in density of phonon states for different numbers and distributions of $\vec{k}$ points. 69
2.7 Potential function allowing a major phase change. 70
2.8 Potential function allowing a minor perturbation to the structure. 71

3.1 INFERNO cold curves for aluminium.. 92
3.2 Predicted ground state energy for aluminium. 94
3.3 Predicted Fermi energy for aluminium, compared with variation for free electrons. 95
3.4 Grüneisen $\gamma(v)$ deduced for aluminium. 96
3.5 Grüneisen $\gamma(v)$ deduced from different models for aluminium. . . 97
3.6 Density of phonon states in aluminium, from neutron scattering measurements.. 98
3.7 Transferability of inverse power potentials. 100
3.8 Phonon density of states for aluminium, deduced from an inverse power potential. 101
3.9 $e(v)$ curves for aluminium from Morse fit to LDA FCC. 103
3.10 Phonon density of states for aluminium deduced from Morse potentials. 104
3.11 $e(v)$ curves from Finnis-Sinclair potentials fitted to FCC ground state energies. 107

3.12 $e(v)$ curves for aluminium from Finnis-Sinclair potentials fitted to FCC and BCC ground state energies. 108
3.13 Phonon density of states for aluminium deduced from an empirical Finnis-Sinclair potential. 109
3.14 Phonon density of states for aluminium deduced from atomic perturbation in the *ab initio* ground state calculation. 110
3.15 Density of electron energy levels in aluminium deduced from the band structure, over a range of values of the FCC lattice parameter. 111
3.16 Electron-thermal energy predicted from band structure. 112
3.17 293 K isotherm for aluminium. 113
3.18 Shock speed – particle speed relation for aluminium. 114
3.19 Shock pressure – density relation for aluminium. 115
3.20 Calculated frozen-ion cold curve. 117
3.21 Phonon densities of states (displacement of 0.01 a). 118
3.22 Variation of lattice-thermal energy with temperature (displacement by 0.01 a). 119
3.23 Variation of the lattice force constant or stiffness with displacement, for different values of the diamond cubic lattice parameter a. 120
3.24 Sensitivity of the density of phonon states to the atomic displacement. ($a = 5.5$ Å) 122
3.25 Variation of lattice-thermal energy with temperature for different displacements. ($a = 5.5$ Å) 123
3.26 Comparison between frozen-ion cold curves, isotherms and Hugoniot for silicon in the diamond structure. 125
3.27 Comparison between frozen-ion cold curves, isotherms and Hugoniot for silicon in the diamond structure (detail near $p = 0$). . . . 126
3.28 Cold curves for BCT Si, with a constant c/a ratio. 129
3.29 Summary of Si BCT ground state energies. 131
3.30 Predicted variation of c/a with specific volume v for BCT Si. . . . 132
3.31 de/dv for BCT Si from σ and polynomial fits to $e(v)$. 133
3.32 Predicted diamond/BCT phase boundaries compared with accepted results for silicon. 135
3.33 Hugoniot curves (pressure – density plane) compared with experimental data. 136

3.34 Predicted stresses on the cold curve for HCP Be with $c/a = 1.567$. 138
3.35 Predicted variation of c/a with a for HCP Be. 139
3.36 Predicted variation of ground state energy with a for HCP Be. . . 140
3.37 Predicted variation of ground state energy with mass density for HCP Be. 141
3.38 Density of electron energy levels calculated for HCP Be. Vertical bars show the Fermi energy for each density of levels. 142
3.39 Electron-thermal energy from band structure for HCP Be. 143
3.40 Electron-thermal energy from band structure for HCP Be. 144
3.41 Sensitivity of ground state energy to plane wave cutoff and k-point density for BCC Be. 145
3.42 Predicted stresses on the cold curve for BCC Be. 146
3.43 Density of electron energy levels calculated for BCC Be. Vertical bars show the Fermi energy for each density of levels. 147
3.44 Difference between chemical potential and Fermi energy (chemical potential at $T = 0$) for BCC Be. 148
3.45 Electron-thermal energy from band structure for BCC Be. 149
3.46 Electron-thermal energy from band structure for BCC Be. 149
3.47 Comparison between cold curve energies for HCP, BCC and FCC Be. 150
3.48 Phase diagram for Be. 151
3.49 Shock Hugoniot for Be (pressure – density space). 152
3.50 Shock Hugoniot for Be (shock speed – particle speed space). . . . 152

4.1 Staggered mesh scheme for defining continuum fields. 156
4.2 Sod's problem: snapshot during propagation of shock wave. 161
4.3 Sod's problem: particle velocity at time 0.2. 161
4.4 Sod's problem: mass density at time 0.2. 162
4.5 Sod's problem: specific internal energy at time 0.2. 162
4.6 Sod's problem: pressure at time 0.2. 163
4.7 Sod's problem: bulk sound speed at time 0.2. 163
4.8 Shock splitting on a non-monotonic Hugoniot. 164
4.9 Range of shock strengths which may split. 165
4.10 Hugoniots for silicon (pressure – specific volume plane). 166
4.11 Schematic of transient X-ray diffraction experiments at 'TRIDENT'. 168
4.12 Sample result from transient X-ray diffraction experiment. 169

4.13 Transient X-ray diffraction traces from 'TRIDENT' experiments on silicon. 170
4.14 Example pressure profiles from hydrocode simulations of shocks in silicon with the diamond/BCT equation of state. 171

D.1 Decomposition of the region of simulation into cells of the range of the interatomic potential. 198

F.1 Idealised shock wave. 206
F.2 1D shock wave produced by a piston. 207

List of Tables

2.1 Binding energies. 57

3.1 Important INFERNO option settings. 91
3.2 Densities at $p = 0$. 92
3.3 Inverse power parameters for aluminium. 99
3.4 Morse parameters for aluminium. 102
3.5 Finnis-Sinclair parameters for aluminium, fitted to FCC cold curve. 105
3.6 Finnis-Sinclair parameters for aluminium, fitted to FCC and BCC cold curves simultaneously. 106
3.7 Free electron fitting parameters for predicted density of electron energy levels in aluminium. 110
3.8 Estimated atomic displacements in Å. 121

4.1 Shock splitting in silicon. 166

A.1 Units and conversions. 177

C.1 Comparison of original and modified structure factor scheme.s . . 187
C.2 Values of $F_j^{[c]}$ for $f_j(r) = r^{-(j+3)}$. 192

Glossary

Terms

EOS	equation of (mechanical) state
GGA	generalised gradient approximation; a correction to the LDA (qv)
Hugoniot	Locus of states which can be reached by the passage of a single shock wave into material of a given starting state.
LDA	local density approximation to the exchange – correlation energy of the electrons
MC	Monte-Carlo
MD	molecular dynamics
principal Hugoniot	Hugoniot for material starting at STP.
STP	standard temperature and pressure, 293 K and 0.1013 MPa

Symbols

e	specific internal energy
f	specific free energy
$\hbar$	Planck's constant $/2\pi$
k_B	Boltzmann's constant
m_e	mass of electron
m_n	mass of nucleon
p	pressure
s	specific entropy
t	time
$\vec{u}$	particle (material) velocity
u or u_p	particle speed
u_s	shock speed
v	specific volume ($= 1/\rho$)
ρ	mass density

Chapter 1

Introduction

1.1 The equation of mechanical state

In a mechanical context, the equation of state (EOS) is a relation between the thermodynamic variables, used to close the equations of motion in a model of the behaviour of materials at the macroscopic (human) scale.

For example, many continuum mechanics problems can be solved – or at least simulated – by considering the evolution in time of variables linked by the conservation of mass, momentum and energy. In 3D, these comprise five scalar equations involving six dependent quantities: mass density ρ, components of momentum $[\vec{p}]_i$, density of internal energy E and pressure p. The EOS may be expressed as an additional relation between ρ, E and p, allowing the system of equations to be closed and a solution found. In practice, it is more common to use the specific internal energy e rather than E.

EOS of the form $p(\rho, e)$ have been and are widely used in mechanical simulations. Since this form of EOS contains only mechanical quantities, it can be deduced from mechanical measurements. These are relatively straightforward to obtain from shock wave experiments – not easy to do accurately, though! – so many such EOS have been published for a wide range of substances. [1]

However, some physical processes are better expressed in terms of the local temperature T rather than e. Examples include chemical reaction rates, viscosity, dislocation activation energies, phase transitions and of course heat flow. The simple mechanical EOS is not thermodynamically complete – it does not express a thermodynamic potential in terms of its natural variables [2] – and so cannot be used to determine the full range of thermodynamic variables. One type of 'product' from the work described here is the thermodynamically complete EOS, from which T and other useful quantities can be found, for a variety of substances.

With computers of the speed which is readily available nowadays, it is possible to predict the EOS more or less from first principles. A scheme has been developed which applies – in principle – to any element or compound (encompassing alloys etc) and which can treat different polymorphic phases.

There are many technological applications of the EOS. Shock waves are found in several branches of mechanical engineering, such as the design of explosive munitions and defences against them, the interaction of space debris with satellites, spacecraft and asteroids, and the effect of high power radiation sources (e.g. lasers and particle beams) with targets. EOS are also used in geology, so determine the equilibrium structure of a planet.

1.2 Why calculate equations of state?

Experiments probing states at densities and temperatures far from those observed at STP can be prohibitively expensive. Calculating the EOS rather than measuring it is an attractive proposition, provided the underlying physics is understood and modelled with sufficient accuracy.

If the physics is tractable enough for the EOS to be expressed in an algebraic form, then this form can be used as a fitting function to interpolate between a small number of accurate experiments. This approach is not followed here; I have preferred to include the maximum amount of physical content which still allows the EOS to be determined in a reasonable time.

Other uses of theoretical EOS are in estimating parameters which cannot easily be measured experimentally (such as the temperature deep inside a shocked sample), and in predicting the states obtained in candidate experimental geometries (allowing design to be performed calculationally rather than by making an entire series of iterative development experiments). Furthermore, theoretical EOS can be forced to be a self-consistent representation of the underlying physics – for example, enforcing thermodynamic consistency.

Conversely, if a physical model of properties is constructed from assumed constituents of a more fundamental nature and then compared with experiments, the fundamental physics is tested. This use of a composite model is widespread in physics, including the quark model of the hadron spectrum, predictions of particle scattering and hyperfine binding energies based on quantum field theory, and gravitational models of orbital trajectories and periods. The present work can be applied as the same process applied to physics at the atomic scale, testing against macroscopic mechanical properties. I do not wish to overstate the importance of this – atomic structures can be probed by far more sensitive techniques such as spectroscopy – but the prediction of an accurate EOS from first principles does demonstrate the universality of the fundamental physics, and the EOS is a reasonable test of our understanding and models of many-body interactions between electrons.

1.3 Scope of this work

EOS can be deduced for densities from zero (atoms arbitrarily far apart) to that of a neutron star, and temperatures from absolute zero to values sufficiently large for atoms to be essentially stripped of all their electrons or even nuclei to evaporate. This range is far too great to be considered here, so I have restricted the states to those which might reasonably be relevant in a 'typical' shock wave experiment driven by a few kg of high explosive.

In compression and heating, the limits chosen are up to $\sim 1\,\mathrm{TPa}$ (10 million atmospheres) pressure, and $\sim 10^4\,\mathrm{K}$ ($1\,\mathrm{eV}$) in temperature. The pressure range is covered by compressions of a factor $\sim$ 2 to 5 in density (up to about 2 in interatomic spacing). The EOS are extended down to $T = 0$, but with a resolution of the order of a few degrees or tens of degrees K. For some substances, the expansion regime is considered out to the perfect gas regime, but relatively little work was devoted to fluid models, so this region of the EOS may be less accurate.

The substances considered were elements and stoichiometric compounds. The latter were constrained to have a similar number density of each species, because the computational effort required to predict the EOS increases with the number of atoms needed to represent a compound.

The states and substances chosen restrict the generality of the physical 'building blocks' required. It was assumed that the atoms were not ionised, but the outer electron states were allowed to change under different compressions and polymorphic structures.

1.4 New aspects

Many parts of the work presented here include the application of well-developed theoretical techniques from atomic and microstructural physics to shock wave physics. A large body of empirical and anecdotal data exists for the behaviour of materials under shock loading, for which complete and predictive descriptions have not yet been found. Reasons include the transient nature of the states obtained, spatial variations, the difficulty of measuring the extreme states produced (including a lack of standards and references), the cost of each experiment (limiting the scope to repeat experiments and determine sensitivities) and the difficulty of isolating the effect of competing processes with different length and time scales.

Other groups are very active in the same field, so my work complements and overlaps results from other researchers rather than being a unique investigation of a new field. It is generally true that each group prefers a different variant of the physical models and numerical schemes used to calculate contributions to the EOS. As far as I know, no other group uses quite the same method of calculating a complete EOS. To this extent at least, my results validate the underlying physics and, taken in conjunction with other people's work, can provide a measure of the uncertainty with which the basic physics is modelled. Some aspects of the work described here were at the cutting edge of technology when first reported, e.g. *ab initio* phonons for silicon.

In my own view, a major attraction of the scheme described here for generating EOS is its generality. Once a reasonable model has been obtained for atoms of an element, it can be used to determine the EOS of the element itself, including alternative polymorphic phases, and the EOS of compounds of different elements, including alloys. As computing power increases, EOS can be predicted for more complicated substances (e.g. with more atoms in the chemical formula) and with greater accuracy. The EOS scheme is thus flexible, general and extensible, whereas many previous methods have been more specific to particular substances.

To be of more use in the context of shock wave experiments, and to describe the work in a less disjointed fashion, I have attempted to summarise the relevant physics where appropriate. These descriptions do not reflect original physics, but the wording is my own and is intended to offer a slant from the shock wave point of view. General references are included for these descriptions.

Otherwise, physics and derivations are my own unless attributed by local references.

In my opinion, the following significant original work was performed in the course of this study:

- Developing algorithms based on the Dirac equation to allow the ground state energy and electron energy levels to be predicted for one or more atoms, isolated or in a lattice.
- Use of the QM codes cited to calculate energy, stress, restoring force and band structures for the materials, structures and densities listed.
- Deriving equations required to deduce EOS from the QM results. This is based on accepted physics, but the equations were re-derived in the most convenient form for EOS.
- Proposing and developing the time-dependent model of phase transitions.

- Proposing and developing the model of vacancy concentration for states away from $p = 0$.
- Proposing and developing the high density fluid model.
- Proposing and developing the low density models for solid and fluid.
- Considering time-dependent effects in the EOS.
- Developing the operator split model for phase transitions and equilibration.
- Proposing the model for non-equilibrium excitation of modes during rapid deformation of a material.
- Deriving equations and algorithms for fitting various empirical equations to data, e.g. the interatomic potentials.
- Deriving numerical schemes for MC and MD simulations. (This has been done before, but the schemes used here were derived independently.)
- Deriving equations to construct phase boundaries and an equilibrium EOS from a set of free energy surfaces tabulated over v and T.
- Re-expressing the equations for X-ray diffraction from crystals to be suitable for predicting powder patterns numerically given the positions of a set of atoms from MC or MD simulations.
- Preparing EOS and phase boundaries for the materials listed.
- Comparing the EOS with experimental data, including the prediction of shock Hugoniots from each EOS.
- Investigating different methods for adjusting theoretical EOS to reproduce experimental data.
- Predicting the splitting of shock waves into multiple-wave structures as in the diamond/BCT EOS for silicon, and comparing with experimental data.
- Evaluating different forms of interatomic potential for EOS use.
- Writing software to perform everything but the electron ground state calculation. The programs used to predict phonon frequencies from forces were based work by other people (see Acknowledgements).

Chapter 2

Theoretical equations of state

2.1 Structure of an equation of state

The objective of this work is to predict the isotropic equation of state – or scalar compressibility – of macroscopic objects. 'Macroscopic' is a subjective scale, but in the present context can be taken to mean 'everyday-sized', say $\sim$ millimetres to metres in extent.

All matter is composed ultimately of fields of fundamental particles. The scales which need to be taken into account when predicting properties depend on the temperatures or energies of interest. These define an effective 'condensation scale' below which no details need be taken into account, or what might be regarded as 'good quantum numbers'.

The present work is concerned with predicting properties for use in hydrodynamic systems, driven by energies typically available from chemical systems. Plasmas are barely considered, and atomic nuclei are assumed to be unaffected by any mechanical deformations. Thus the substructure of the nucleus can be ignored, and the effect of nuclei represented by a point mass and electromagnetic field. (In most cases, we can also assume that the inner electrons of an atom move with the nucleus.)

The EOS of a substance can be deduced from the variation of specific internal energy e with specific volume v and temperature T. Given $e(\rho, T)$, the laws of thermodynamics can be used to deduce the specific entropy s and hence the specific free energy f and pressure p.

The internal energy can be split into the cold curve $e_c(v)$ (zero temperature frozen-ion isotherm) and thermal contributions from the ions $e_i(v, T)$ and electrons $e_e(v, T)$ In this prescription, isotherm measurements would not tend to e_c as $T \to 0$ since the ion-thermal contribution includes the finite zero-point vibrational energy.

In the procedure developed in the present work, we attempt to retain the most physically complete model of the processes underlying each contribution to the EOS, within reason and the realities of computational power. Numerical calculations of the electronic wavefunctions play an essential part; the limitations and approximations used in the calculation of the wavefunctions, such as the spatial and energy resolution of the basis functions, are well-understood and straightforward to control. This is in contrast to other approaches to the prediction of EOS, where quantum mechanics may be used to predict effective interatomic potentials (IAPs) directly [3, 4, 5, 6, 7, 8]. The IAPs are generally easier to use to deduce the

EOS (except for the electron-thermal contribution), but the choice of functional form for the potential introduces approximations which are difficult to quantify.

2.1.1 Internal energy of a solid phase

In crystalline phases, each ion vibrates about a fixed average position. The energy can be split conveniently into contributions from the frozen-ion cold curve $e_c(v)$, the thermal motion of the atoms $e_l(v,T)$ and the thermal excitation of the electrons $e_e(v,T)$:

$$e(v,T) = e_c(v) + e_l(v,T) + e_e(v,T). \tag{2.1}$$

The cold curve energy for a phase at some volume is the energy of the systems with the electrons in their ground state with respect to the ions. The lattice-thermal and electron-thermal energies are the result of perturbations of the ion positions and electron occupation numbers from their ground state values.

Other contributions

Strictly, electron-phonon interactions should be taken into account. Displacement of the atoms from their equilibrium positions alters the electron energy levels, and excitation of the electrons from their ground state alters the stiffness of the lattice with respect to displacement of the atoms.

The effect of atomic displacement on the electron energy levels is largely taken into account in the calculation of the lattice-thermal energy, as discussed below. The variation in lattice-thermal properties with thermal excitations of the electrons was neglected.

At sufficiently high temperatures, the excitation of the electrons means that the separation into cold curve and electron-thermal contribution is not accurate. At the temperatures considered here, these interactions mainly affect the thermal and electrical conductivity; the contribution to the EOS is much smaller.

Vacancy population

For any temperature and density, there is some equilibrium concentration of vacancies in the lattice of atoms. The proportion of vacancies f_v can be calculated given the vacancy formation energy E_v:

$$f_v(T) = e^{-E_v/k_BT}. \tag{2.2}$$

Given a population of vacancies, contributions can be calculated to the internal energy of the lattice and also to the entropy, since there is a configurational entropy associated with the distribution of vacancies over possible sites.

The vacancy formation energy can be estimated from the mean binding energy E_b of atoms in the solid, since to a first approximation (order of magnitude) the energy required to form a vacancy is just the energy required to remove an atom from the lattice.

If the EOS is defined such that the cold curve energy $e_c \to 0$ as $\rho \to 0$ then $E_b = e_c A$ where A is the mass of an atom. (This calculation is considerably more complicated if the material is not an element.) Since e_c is a function of ρ it can be seen that the E_v also varies with ρ.

This use of the cold curve energy is over-simplistic because

- The crystal lattice deforms around a vacancy, altering the electron distribution and hence the effective formation energy.
- Vacancies can influence the electron energy levels and the phonon modes, altering these contributions to the equation of state. These effects depend on the vacancy concentration.
- It can be energetically favourable for several vacancies to merge, forming a microvoid.

A more rigorous treatment might consider the effect of the vacancy concentration on the other contributions to the EOS, and the variation of E_v with ρ, T and f_v.

In order to calculate the vacancy formation process properly, consideration must be given to the fate of the atom which is removed. This is crucial for calculations in compressed states. Vacancy formation can be treated in either of the following ways:

- An atom is removed, reducing the mass in the volume under consideration, and thus reducing the density for the same nearest-neighbour spacing. (In other words, the mean density is reduced for the same density in the unperturbed lattice.) The bulk density corresponding to a crystal density of ρ is $\rho(1 - f_v)$.

- An atom is moved from its initial position to a new region of lattice, and the whole lattice is then compressed to fit into the same volume, reducing the nearest-neighbour spacing for the same density. (In other words, the mean density is held constant whilst increasing the density in the unperturbed lattice.) The crystal density corresponding to a bulk density of ρ is $\rho/(1-f_v)$.

In either approach, f_v is varied iteratively to find the equilibrium population as a function of ρ and T, given the vacancy formation energy E_v as a function of ρ and T. If E_v is calculated perfectly, both approaches would give the same result for a sufficiently large system. If E_v is somewhat in error, the approaches should give an indication of the likely uncertainty in f_v.

As the material state changes, it takes a finite time for the equilibrium vacancy population to form. In principle, vacancies can be created and destroyed by thermal and mechanical disturbances of the perfect lattice. However, the energy barriers to such processes are much higher than for processes involving existing defects, such as dislocations [9], grain boundaries and other vacancies. The formation rate for vacancies from an atom on 'site i' can be approximated by a family of Arrhenius forms,

$$\dot{f}_{vi} = \int_{\delta S} F_i(\vec{r}) e^{-E_{vi}(\vec{r})/k_B T^*} \, d\vec{r}, \tag{2.3}$$

where the integral is over displacement directions $\vec{r}$ of the atom on the surface of a unit sphere δS. The energy barrier E_{vi} varies with density, and the attempt rate F_i with density and temperature. In a lattice without defects, F_i can be calculated from the density of phonon states. However, since E_v is lowest in the vicinity of existing defects, it is the corresponding values of F which should be used; these are not straightforward to calculate. Near some kinds of defect, e.g. a void or grain boundary, $F(\vec{r})$ is likely to be less than in the perfect lattice. Near interstitials, atoms are closer together and $F(\vec{r})$ is likely to be greater than in the perfect lattice. Near a dislocation, the average density is lower but atoms may be closer together in some directions, so $F(\vec{r})$ is likely to be similar to its value in the perfect lattice – although the displacement direction(s) most likely to cause a vacancy could be considered in more detail. A simple estimate can be made by taking an average phonon frequency from the density of states.

The Arrhenius factor is based on an 'effective temperature' T^* rather than simply the local temperature to take account of shock waves. Shocks change the material velocity in a short time. A finite time elapses before the kinetic modes reach thermodynamic equilibrium, so energy can be concentrated in certain translational or vibrational modes before equilibrating. Thus T^* can be significantly greater than T for a short period following the passage of a shock. For modes parallel with the shock, a plausible relation is $T^*_{\mathrm{max}} \sim 3T$. This superheating process is thought to explain some phenomena in the shock initiation of chemical reactions [10]. The T^* for any particular process may not be the maximum value; it depends on the specific pathway involved and may in fact be less than T, for example if driven by one of the modes normal to the direction of the shock.

The vacancy formation rate may be large during a phase transition, since large atomic rearrangements may occur and free energy may be available.

As well as contributing to the EOS, the vacancy concentration could be used to estimate the melt locus by considering contours of constant f_v in (ρ, T) space.

Behaviour at low densities

It is possible to imagine – and calculate – the effect of expanding a solid isotropically beyond the equilibrium ($p = 0$) density. First, the solid experiences a tensile stress. Equilibrium positions can still be found for the atoms, and physical phonons calculated. As the density decreases, the stress initially increases. As the density tends to zero, the stress eventually falls to zero. There is thus a maximum stress, corresponding to the maximum gradient of $e(v)$. For densities below this value, the material is mechanically unstable: it is energetically favourable for a region of constant density ρ to split into smaller regions of density ρ_0 (the $p = 0$ value) and empty space. As the density decreases, a growing number of phonon modes will be imaginary as the crystal becomes unstable with respect to the corresponding displacements of the atoms. (Fig. 2.1.)

In fact, for any density below ρ_0 the material is unstable with respect to non-isotropic (shear) deformations. If the stress contains any non-isotropic component (which is almost always the case for real materials in 'engineering' problems) then the material may deform plastically, eventually leading to fracture. Cracks may open and voids grow even in the case of an isotropic tensile stress. These processes are energetically favourable in a sample of sufficient size because the strain energy is proportional to the sample volume V, whereas the surface energy is proportional to $V^{2/3}$.

These failure modes are time-dependent. In fact, sufficiently small stresses

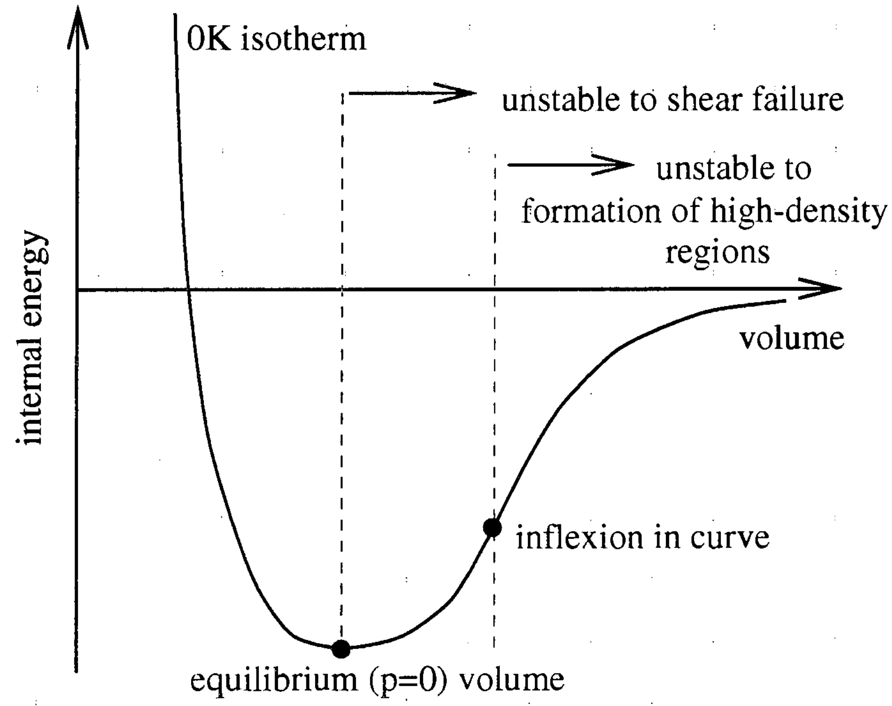
0K isotherm
unstable to shear failure
unstable to formation of high-density regions
internal energy
volume
inflexion in curve
equilibrium (p=0) volume

Figure 2.1. Schematic of the 0 K isotherm indicating regions of instability.

may result in very slow rates of plastic deformation, such as creep and metal fatigue. Damage and failure models form a large area of research in their own right [11]. An approximate model suitable for extending EOS to the cold expansion region in a reasonable way would be to approach $p = 0$ according to one or more relaxation time scales. A separate time scale should be chosen for the shear failure compared with the failure by density instability. The time scale can be made a function of density to allow this flexibility; if also a function of temperature then it can be used to represent thermally-activated processes such as the motion of dislocations. This simple model includes no further dependence on internal state, such as a damage parameter or porosity. Porosity is probably the next most important improvement.

Rigorously, the instability to density perturbations should involve an equilibration process. In reality, the energy depends on the size distribution of high density pieces.

Strictly, when the density instability takes effect and empty regions nucleate, the remaining regions of solid are in equilibrium with a vapour of atoms. This contribution to the EOS is not considered further in this work, because

- Equilibration of cold material with vapour is likely to take place over long time scales.
- The effect is almost certainly negligible in the response of most materials of current engineering interest for loading at high rates such as in a shock wave.

2.1.2 Internal energy of a fluid phase

In fluid phases, there may be no static reference configuration of ions on which to base the energy contributions at each volume. Some possible approaches and models in this more complicated case are discussed below. Attention is concentrated on methods which are likely to be valid over a wide range of densities. Virial expansions and corrections to the ideal gas EOS (like van der Waals') are omitted, since these are known to have a restricted range of validity [2].

In the present work, fluids were considered only to extend the range of a solid compression EOS in an approximate way for convenience when applying the EOS to problems in which the material happened to re-expand. Some methods are easier to use in conjunction with the solid EOS calculations than others; for example, effective interatomic potentials can be predicted at least approximately with some ease. A careful application of the more involved techniques for high density fluids, such as integral equations and liquid perturbation theory [12, 13, 14], are beyond the scope of this work.

Ion positions

Depending on the fluid model used, it may be possible to define a 'typical' configuration of ions, as a function of temperature as well as volume.

An example is if the fluid is viewed as a disordered system. When the average interatomic separation is similar to that in a solid, the configuration could be represented to some finite accuracy as a periodic array of large lattice cells, each of which contains a random arrangement of atoms. The electron ground state, density of energy levels and atomic forces could then be found exactly as for a crystalline solid. Care would be required to include the configurational entropy properly, as this is more important than for near-perfect crystals. A significant number of vibrational modes are likely to be imaginary; these should be treated as free kinetic modes.

The arrangement of atoms is unlikely to be truly random, particularly in a dense fluid. However, it may be necessary to consider configurations consisting of a large number of atoms in order to reproduce the local structure faithfully enough to be representative of the fluid, at each volume and temperature. With a large lattice cell, the calculation of each ground state energy is also more demanding than in a simple crystalline structure.

It is unlikely in general that a sufficiently accurate static 'typical' configuration can be found. Fluids 'really' consist of a population of atoms whose configuration is in a dynamic state which maintains an average pressure and volume at a given temperature. The 'natural' state of a fluid is therefore an average over many configurations, each with its own ground state and thermal energies. It is possible to sample configurations using Monte-Carlo or molecular dynamic methods,[2] and to determine the electron states with respect to the instantaneous configuration of the ions. However, this approach is computationally expensive, particularly when repeated for a table of densities and temperatures.

Quasisolid

A much larger set of configurations for ion positions can be considered given an effective interatomic potential (IAP) for the material. The force on each atom and the total energy can be calculated far more quickly using an IAP than from an electron ground state calculation. [12]

IAPs can be deduced from more rigorous treatments of the lattice in a variety of ways, such as adjusting parameters in an algebraic IAP to reproduce states from rigorous calculations, as used elsewhere in the present work.

Given an IAP, the ion-thermal contribution for atoms in a fluid could be estimated by generating a large lattice cell containing a random distribution of atoms, and calculating a cold curve and phonons as for the lattice cell in a crystalline solid. Some phonon modes are likely to be imaginary; these should be associated with free translational modes.

Some degree of adjustment of the atom positions is necessary to start from equilibrium positions with respect to the random configuration. The adjustment should be carried out in a heavily-damped fashion, otherwise the atoms may gather enough momentum to rearrange themselves into a solid structure.

If the IAP has a finite range l then each simulation need not consider more atoms than are contained in a cube of side $2l$. In this case, it is advisable to perform a number of simulations with different random configurations. It may be more convenient to choose a single, much larger configuration and accept the penalty in unnecessary phonon calculations. The size is limited by the time required to calculate the phonon modes for a system of a given number of atoms N, which is determined by the time needed to find the eigenvalues of a $3N \times 3N$ matrix – orders of magnitude less than the electron ground state calculations. The accuracy of a given set of simulations can be tested statistically, to see whether the fluid state has converged to a constant value as additional configurations are added.

This 'quasisolid' approach is only possible for fluid at a reasonably high density, because it assumes that a significant proportion of the thermal energy is in collective vibrational modes. As with the solid in tension, this assumption breaks down when the expanded material becomes formally unstable, i.e. at a density corresponding to the inflexion in the cold curve. Low density fluids are considered below.

Local densities of states

In a crystalline material, the density of electron energy levels $G(E)$ can be calculated for the material as a whole. Making some assumptions about the range of interactions between atoms, the total density of levels can be expressed as a sum over *local* densities of levels $g_i(E)$ associated with each atom i:[15]

$$G(E) = \sum_i g_i(E).$$

The local density of levels might be parameterised in terms of the environment around each atom, e.g. the coordination number, orientation of nearest atoms, pair correlation function etc. Depending on the complexity, determining $g_i(E)$ from ground states of the whole system might require these to be found for a range of structures at each density.

The local density of levels is defined formally as

$$g_i(E) \equiv \int |\langle i|\Psi_k\rangle|^2 \, \delta(E - E_k) \, dE_k$$

where Ψ_k is the eigenstate of the system with energy E_k [15]. The weighting factors $\langle i|\Psi_k\rangle$ contain the information on and assumptions for using the local environment.

Once the $g_i(E)$ have been found, the energy of a non-crystalline configuration of atoms such as a fluid can be estimated quite quickly, by choosing the $g_i(E)$ which is appropriate for each. This should be far more efficient than calculating the electron ground state from scratch. However, since the implicit basis functions do not cover the whole system, they may not be mutually orthogonal. This inconsistency can affect the accuracy of the predicted band structure, and hence the estimates of forces between atoms.

An analogous procedure can be followed for the density of phonon states. Recasting the problem in terms of the local density of states appears to avoid the need to express interatomic interactions in terms of an effective interatomic potential. Related assumptions are made, such as the range at which atoms interact, but the physical framework appears to be more rigorous and easier to extend if higher-order terms are needed. However, the vibrational modes deduced are not 'normal modes' in the usual sense. Integrated properties such as the atom-thermal contribution should be predicted with reasonable accuracy, but it may not be safe to draw conclusions about detailed properties such as the split between stable and unstable vibrational modes.

Direct numerical simulation

As with non-linear effects in the solid lattice, the properties of a fluid can be investigated by direct numerical simulation of an ensemble of atoms. [12] These techniques are discussed in more detail below.

Limiting forms of the ion-thermal energy

According to classical statistical thermodynamics [2], each quadratic mode (including harmonic potentials, translational and rotational motion) contributes $\frac{1}{2}k_B$ to the total heat capacity. In quantum mechanics, modes are 'quenched' as the temperature T falls below their characteristic temperature, $\hbar\omega_i/k_B$, but all such modes contribute $\frac{1}{2}\hbar\omega_i$ to the total zero-point energy.

This picture can be complicated to apply in practice.

- Characteristic frequencies ω_i vary with density. Fluids exhibit an inhomogeneous local density $\rho(\vec{r})$. The average density could be used to estimate an average frequency $\bar{\omega}_i$, but if $d\omega_i/d\rho$ is non-linear (which is true in general) then the mean heat capacity per vibrational mode departs from k_B.

- Composite particles may behave as a 'point with mass' at low temperatures and densities, but dissociate or excite internal modes at high temperatures and/or densities. For example, molecules may start to rotate and atoms to ionise. This can be a less obvious case of mode quenching.
- Modes may not be quadratic, so the heat capacity may vary from $\frac{1}{2}k_B$. The apparatus of perturbation expansions can be used to correct for small anharmonicities, but such corrections are not in general guaranteed to converge to a well-defined answer.

The crystallite model of the fluid state [16] requires modes and contributions to the zero-point energy to be counted for different crystallite size distributions. Ionic contributions can be considered in terms of translation, rotation and vibration of each crystallite. It should be possible to reach the same result – at least to some limiting bounds – by considering atoms rather than crystallites, ignoring ionisation. Different viewpoints reflect the same physical processes.

For example, consider the kinetic degrees of freedom of a crystallite consisting of 2 atoms (or a diatomic molecule). Let us assume that individual atoms have no rotational degrees of freedom (the moments of inertia being very small so these modes are quenched) and that they remain un-ionised and hence may be treated as having no structure (no dissociation or internal degrees of freedom). An ensemble of N atoms has $3N$ translational degrees of freedom, giving a heat capacity of $3Nk_B/2$. The crystallite has 3 translational degrees of freedom, 2 rotational (rotation about the axis is suppressed because the moment of inertia is small) and one vibrational. Therefore, it makes no difference whether the degrees of freedom are counted by considering the atoms as separate or in a crystallite. The strength of interaction matters, because the vibrational mode may be quenched at sufficiently low temperatures. If not, the interatomic potential contributes an additional degree of freedom, giving *approximately* $\frac{1}{2}k_B$ of heat capacity for a potential which is *approximately* quadratic.

Modifications to the free electron model

The electron-thermal energy can be estimated *crudely* by assuming that conduction electrons are identical non-interacting fermions in a constant background potential. [2] As a simple alternative to band or bonding theory, with no guarantees of accuracy, free electron theory could be modified in several ways to add some more flexibility and maybe even plausible systematic trends by measures such as:

- Model electron interactions, exchange and correlation, by correcting the energy with simple functions of electron density.
- Considering the outermost electrons, associate them with an outer limit in an isolated atom ($\sim$ the 'atomic radius') and inner limit (representing a region from which they are excluded by the presence of the inner electrons). As atoms approach, the region of overlap between the outer electron shells from each atom defines a total volume for the conduction band. As atoms move even closer, the exclusion regions add an extra short-range repulsion. The electronic contributions to the energy of different crystallite populations could be modelled in this simple geometrical way.
- This model could be extended quite easily to inner electron shells. When these start to overlap, more electrons are made available for conduction. This process is sometimes called an electron phase transition or Mott transition, though the latter can also refer specifically to insulator / conductor transitions.
- More sophisticated versions of this approach, with electron states based on spherical harmonics around each atom rather than waves in a box, have been used previously in EOS theory.[17, 18]

Behaviour at low densities

The fluid configuration at low densities is dominated by the clustering of atoms into small groups, each moving independently in free space of a fixed volume. The density in each cluster is close to the $p = 0$ equilibrium density, so long as the overall density is low enough for atoms in each cluster to re-equilibrate between inter-cluster collisions.

The energy can be predicted as a function of temperature for a cluster of a known configuration. For small clusters, the normal modes can be calculated in the same way as the phonon modes in an infinite solid. For large clusters whose internal structure looks like that of a solid, the solid's density of phonon states can be applied, with a cutoff for wavelengths greater than the size of the cluster.

In principle, the electron energy levels should be calculated for each cluster configuration; in practice it may be reasonable to consider only the energy levels in an isolated atom, those in an infinite solid (again discarding wavelengths greater than the cluster size) and possibly those for 2 or 3 atoms close together.

As well as their internal energy, each cluster possesses translational kinetic energy, and rotational energy if the temperature is high enough to excite these modes.

In principle, the energy of the system at a given ρ and T should be calculated by considering a large set of alternative distributions of atoms into clusters of different types. In practice, it is possible to characterise the population in terms of a distribution over clusters of different size.

Various ways can be devised to predict the cluster size distribution as a function of density and temperature [16]. The method proposed here is to consider the partition function for a microcanonical ensemble. Consider a set of possible types of cluster, each having a mean binding energy E_i per atom and containing N_i atoms. E_{bi} includes the translational and rotational energy, and is a function of T. The probability that an atom will be found in a state of energy E_i is

$$f(E_i) \propto e^{-E_i/k_BT} \equiv \alpha e^{-E_i/k_BT} \tag{2.4}$$

for some constant α. If the number of clusters of size N_i is n_i then

$$f(E_i) = \frac{n_i N_i}{N}, \tag{2.5}$$

where N is the total number of atoms. Also,

$$\sum_i n_i N_i = N, \tag{2.6}$$

so

$$\sum_i f(E_i) = 1 \quad \Rightarrow \quad \alpha = \left[\sum_i e^{-E_i/k_BT}\right]^{-1}. \tag{2.7}$$

The distribution can be interpolated smoothly over large clusters (over 1000 atoms or so), and the distribution over smaller clusters considered explicitly if necessary. (There is evidence that small clusters tend to have 'magic numbers' of atoms, so the smooth distribution is liable to break down for small numbers.) It is probably reasonable to assume that clusters minimise their surface area;

protruberances are more likely to be removed in collisions between clusters. Given estimates of $E_i(T)$ for a 'representative' set of clusters, the $f(E_i)$ can be calculated as a function of T, giving $e(T)$. The validity of a given set can be tested by extending the set and determining whether a converged result has been obtained.

It may be possible to include the effect of interatomic potentials in influencing neighbouring clusters; this would add extra density-dependent interactions.

Cluster distributions as calculated above are an equilibrium population. As the density and temperature of a vapour changes, a finite time is needed for the population to reach equilibrium. It is useful to estimate the time scale for equilibration. Unless significantly shorter than the hydrodynamic time scales of interest, it may not be accurate to use a unique EOS. The time scale for equilibration can be estimated as a function of density and temperature from the mean free path of the clusters. To make this calculation more representative, the 'starting' distribution could be used as an alternative to the equilibrium distribution.

2.1.3 Completing a single-phase equation of state

Given the total specific energy $e(T)$ along each isochore, the specific entropy s was found by integration of the second law of thermodynamics ($de = T ds - p dv$):

$$s(T) = \int_0^T \frac{dT'}{T'} \frac{\partial e}{\partial T'}. \tag{2.8}$$

Since e was represented by a table of values at discrete (v, T) points, the table was interpolated by fitting a quadratic to each set of 3 adjacent $e(T)$ points on each isochore:

$$e = \alpha_0 + \alpha_1 T + \alpha_2 T^2, \tag{2.9}$$

giving

$$\frac{\partial e}{\partial T} = \alpha_1 + 2\alpha_2 T \tag{2.10}$$

and hence a difference in entropy between successive temperature ordinates T_i and T_{i+1} of

$$\Delta s = \int_{T_i}^{T_{i+1}} \left(\frac{\alpha_1}{T} + 2\alpha_2 \right) dT = [\alpha_1 \ln T + 2\alpha_2 T]_{T_i}^{T_{i+1}} . \tag{2.11}$$

This formula breaks down for $T = 0$, for which α_1 should equal zero. In practice α_1 was very close to zero in the first interval, so the logarithmic term was ignored there. Temperature intervals internal to each isochore were spanned by two such polynomial fits (with the third point at the next lower or higher temperature). The entropy contribution was averaged between these two values.

The specific free energy f was calculated at each (v, T) point simply by use of the formula $f = e - Ts$.

The pressure p was calculated by differentiating the free energy f:

$$p = -\frac{\partial f}{\partial v}. \tag{2.12}$$

To calculate p at each (v, T) point, f was fitted at the nearest three specific volumes along the isotherm with a quadratic

$$f = \beta_0 + \beta_1 v + \beta_2 v^2, \tag{2.13}$$

giving

$$p = -(\beta_1 + 2\beta_2 v). \tag{2.14}$$

The EOS were generated in an ASCII version of SESAME 301 format [19]. This consists of rectangular tables of pressure p (GPa) and specific internal energy e (MJ/kg) as functions of temperature T (K) and density ρ (Mg/m^3).

2.1.4 Polymorphism

All atomic matter exists in several polymorphic phases, depending principally on the density and temperature. The simplest materials may exist in solid, liquid or vapour states, and at higher temperatures the atoms ionise to form plasma. In more complicated materials, the solid and liquid phases may be subdivided into other phases, e.g. the different crystalline structures.

Given an equation of state for each phase of a polymorphic substance, it is possible to predict the overall equation of state allowing phase transitions to take place. However, the polymorphic equation of state is not unique. Phase transitions do not take place instantaneously, so the polymorphic equation of state is really time-dependent.

Because phases may persist outwith the range of states where they are observed in equilibrium, the EOS for each phase should be generated to span a wider range of states.

The approach adopted here is to deduce the EOS for polymorphic materials from a thermodynamically complete EOS for each structure. Thus the poly-phase EOS is itself thermodynamically complete, and phase boundaries are predicted as functions of pressure and temperature. This is a considerable advance over calculations of phase boundaries as a function of pressure or compression at $T = 0$ [20, 21, 22, 23, 24]. As well as neglecting the variation with temperature, calculations at $T = 0$ in the frozen-ion model also neglect the contribution from the zero-point energy of the phonons. In addition, a careful inclusion of the thermal contributions can provide insights into the mechanism whereby a phase transition takes place, and hence provide estimates of the rate of a phase transition [25].

Equilibrium mixtures of ideal phases

The general condition for equilibrium is that the availability

$$a = e - \sum_i \alpha_i \beta_i \tag{2.15}$$

is minimised. α and β are state parameters whose product has dimensions of specific energy, one of which is a property of the surroundings of the system under consideration (i.e. a constraint) and the other of which is a free parameter of the equilibrium state. The availability is the work which can be extracted from the system with respect to its surroundings. Examples are $\alpha\beta = T_R s$ and $\alpha\beta = -p_R v$, where the parameter with subscript R refers to the surrounding reservoir (temperature, pressure) and the other parameter may be varied in determining the equilibrium state. [2]

There is no obviously unique way of choosing the $\{\alpha\beta\}$ for a general-purpose equation of state – it depends upon the intended application. In the present work, we consider constraints appropriate for continuum mechanics. The definition of the environment is a source of confusion. In many algorithms used to model continuum mechanics, the equation of state is used to evaluate the pressure p given the density ρ and specific internal energy e. However, ρ and e are not necessarily constraints in the sense required to define the availability. The convention adopted here is to tabulate the state variables over density ρ and temperature T, i.e. to calculate e and p at a fixed ρ and T. The energy associated with these parameters is the (specific) free energy,

$$f \equiv e - Ts, \quad df = -sdT - pdv. \tag{2.16}$$

Note that the Gibbs free energy,

$$g \equiv e - Ts + pv, \tag{2.17}$$

is not the appropriate potential to use because the specific volume of a mixture of phases is not free to change to obtain an equilibrium with respect to the environmental pressure.

It can be argued that the real physical process is a dynamic propagation of waves driven by differences in temperature T and pressure p, so these are the natural constraints for the polymorphic equation of state at any point. A difficulty with this line of reasoning is that in a dynamic situation, it is not possible to identify a 'reservoir' of constant T and p, and infinite heat capacity and volume, surrounding each element of material for which the equation of state must be evaluated. The reason for this is that thermodynamic variables are ill-defined over sufficiently small regions of space and time. To use a thermodynamically-based equation of state, we assume implicitly that the intervals of space and time over which the physical processes underlying the equation of state take place are much smaller than the intervals used to represent the materials in our calculations.

Consider a hypothetical material consisting of two phases A and B (Fig. 2.2). Over most of state space, f is minimised if the material is pure A or pure B. However, at intermediate volumes, f can be reduced if the material splits into a mixture of A and B. If the regions of pure A and B in the mixture do not interact then the free energy is the tangent between the curves for the pure phases.

For materials usually met in mechanical applications, a necessary condition for different phases not to interact over a finite length of time is that they must be in pressure and temperature equilibrium. Depending on the material, other

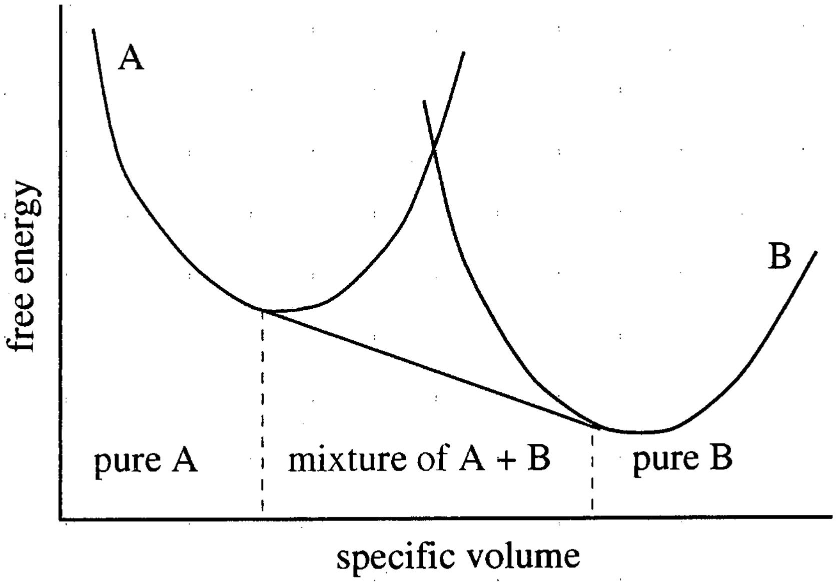

Figure 2.2. Equilibrium mixture of two non-interacting phases.

constraints may be found. For example, each phase may have a well-defined strain, chemical potential, magnetic dipole moment and so on. In principle, each of these constraints should be included in finding the equilibrium state. In addition, the regions of different phase may interact with each other, for example inducing a surface tension at the interface.

The interface effects such as surface tension depend on the scale of the regions of each phase. Initially we assume that the domain of interest can be made arbitrarily large compared with the regions of pure phase, so that the contribution of the interfaces is negligible in comparison with the availability of the bulk material. In this case the common tangent construction can be used to find the proportions of A and B giving the lowest availability.

An idealised equilibrium polymorphic equation of state $\{e,p\}(\rho,T)$ can be generated from a set of single-phase equations of state $\{\{e,p\}_i(\rho,T)\}$ for each (ρ,T) state as follows:

1. Calculate the $\{f_i(\rho,T)\}$ e.g. applying the second law of thermodynamics [2] to each $\{e_i(\rho,T)\}$ in turn to determine $\{s_i(\rho,T)\}$ as for a single phase equation of state.

2. Consider the set of isotherms $\{f_i(\rho)\}$ at the desired temperature T, constructing the isotherm of minimum f between them:

(a) For each pair of isotherms $\{i,j\}$, construct the compound isotherm

$$f_{ij}(v) = \min(f_i(v), f_j(v), \tilde{f}_{ij}(v)), \tag{2.18}$$

where $\tilde{f}_{ij}(v)$ is the common tangent between the isotherm of phases i and j, defined only between the points where it intersects its parent isotherms.

(b) Calculate the minimum of all the $\{f_{ij}(v)\}$, and hence the proportion λ_i of each phase in the mixture. This is done using the lever rule: $\lambda_i = 0$ except for those phases on the lowest of the $\{f_{ij}(v)\}$, where the λ are found from the position of the desired volume v with respect to the volumes $\tilde{v}_{\{i,j\}}$ at which the common tangent intersects the pure-phase isotherms:

$$\begin{aligned} \lambda_i &= \min[1, \max[0, (v - \tilde{v}_j)/(\tilde{v}_i - \tilde{v}_j)]] \\ \lambda_j &= 1 - \lambda_i. \end{aligned}$$

Since the EOS for each phase is generated separately and then combined to find the equilibrium EOS, it contains a complete description of the phase changes including the order and latent heat. The procedure is equally valid for compounds as well as elements.

Care is needed when constructing the complete EOS from the equilibrium free energy surface that the discontinuities in specific internal energy are treated properly. This can be done only approximately with a rectangular $e(\rho, T)$ table, and problems may occur with the interpolation scheme. One area for future development is the structure of EOS tables where phase transitions are important. One way is to base the EOS on a set of isochores where the temperatures tabulated can be different between isochores; in this way, the tabulation can be chosen to include a point on each phase boundary. Another way of generalising, described below, is to use the set of single-phase EOS explicitly and add a (time-dependent) model of polymorphism to transfer between the alternative EOS.

It has been argued that the common tangent construction for 2-phase equations of state means that the shock Hugoniot should have a region of constant pressure. This is only true in the rather specific circumstance where the slope of the common tangent remains constant with temperature; any other dependence will give a varying pressure across the 2-phase region. A region of constant pressure will of course occur during isothermal compression.

Equilibrium mixtures of real phases

In general, the different phases present in a mixture may interact. When this happens, the availability of the mixture may not be given accurately by the common tangent construction.

If the chemical potentials of the phases are not equal then electrons will transfer between the regions of different phase until the potentials equalise. This alters the structure of the Coulomb potential of the material: it may be electrically neutral over a longer length scale. The electron-thermal energy may be slightly different from the pure phases, since the energy levels will be filled in a slightly different way.

Some other effects can be considered, such as atomic disorder, the scattering of electrons and phonons at the boundaries between the phases, different magnetic moments (electrons + nuclei), and differences in stress-strain behaviour. However, these effects also apply to polycrystalline preparations of pure phases. In the present work, effects of this magnitude are ignored.

As with most 'real' preparations of pure phases, mixtures are prepared in a finite time so the regions of each phase may be much smaller than the whole system. The entropy of mixing may make a significant contribution to the total energy.

Phase change dynamics

Phase changes occur at a finite rate, which may depend on temperature, compression and the strain rate. The time required for tens of percent of a material to change phase may be significant compared with typical time scales for objects ~mm in size to deform under the influence of a shock driver, particularly if the material is a stoichiometric compound rather than an element.

For a region of material to undergo a polymorphic phase change, atoms must rearrange. Usually, the initial structure is stable, and each atom oscillates about its equilibrium position. In order to reach the alternative structure, atoms must cross a potential energy barrier. The rate at which this happens depends on the barrier height, the temperature, the frequency at which the atom oscillates in the potential well, and interference effects in the oscillations of neighbouring atoms. In the early stages in the formation of an alternative structure, the new positions are usually less stable than the old. Atoms are quite likely to jump back to their old position. Once atoms start to nucleate a region of the new structure, the effective depth of the potential wells at the new positions becomes deeper. (For example, the wavefunctions of electrons in conduction bands can include more contributions at longer wavelengths, so the ground state energy decreases.) This phase changes occur by a process of nucleation and growth.

It is possible in principle to find phase changes which do not involve an energy barrier of this type. As the lattice deforms, it might be energetically favourable for the electrons to rearrange in a Mott transition. Electron transitions take place far faster than the nuclear motion, so the atoms could find themselves suddenly in positions of static instability. In this situation, the phase change could take place far more quickly.

Again in principle, phase changes might occur through intermediate structures. If a large amount of atomic rearrangement is required, it may be that some metastable structures occur with significant lifetimes, bridging the gap between the initial and the alternate structures.

When the deformation history is driven by shock waves, some other effects may matter. Shock waves have a characteristic direction of propagation: in the shocked state, the material is hotter and moves with respect to its undisturbed state, so the average velocity of the atoms must change and the magnitude of variations about the average must increase. Shock waves have a finite thickness, reflecting the length scale over which large amplitude variations in density (or velocity, etc) can propagate without scattering. Intuitively, disturbances parallel to the direction of propagation of the shock should move most quickly – interatomic forces are usually greatest when atoms approach along the line between their centres. Thus when a region of material is first affected by a shock wave, the longitudinal modes of vibration are likely to be excited first. (This may not always be the case. In some materials, angular forces between electronic orbitals of higher angular momentum may cause some transverse modes to be excited first, over some range of shock pressures.) The effect of the scattering processes is to transfer energy from the longitudinal modes to transverse modes, and other degrees of freedom.

Because shock waves cause non-equilibrium excitation, some modes are 'superheated' for a finite time. This may greatly increase the rate at which atoms jump into sites of an alternative structure. This effect depends on the orientation of each crystal with respect to the incident shock. Similarly, the anisotropic strain generated by a shock wave may in general cause a phase transition to occur at a different (mean) isotropic pressure than in static compression.

Another estimate of the characteristic time required for a phase change to occur at the local level can be obtained by considering the distance which each atom must move to reach its new site, divided by the speed at which the atoms move at the temperature of interest (and taking into account superheating of modes caused by shock compression). Many paths may be possible between the initial structure and each alternative; any path or paths may be followed with varying probability. The set of possible paths may be estimated by Monte-Carlo studies of groups of N_{cm} atoms, where N_{cm} are common multiples of the numbers of atoms in the lattice cell for the initial structure and an alternate.

Nucleation and growth model

Polymorphic phase changes occur by the random creation of regions of the alternative phase, created by thermal (or mechanical) motion of the atoms. The phase change process can be investigated at the microstructural level using direct numerical simulation or statistical models, but it is desirable to have a macroscopic model suitable for use in continuum mechanics calculations.

Consider a material initially in a phase S_i, with a set $\{S_j\}$ phases available to it. Each phase has a unique EOS; for the purposes of polymorphism the important property if the free energy $f_i(\rho, T)$. In order to nucleate a region of a given phase j, a region of S_j must be created within S_i. This occurs by atomic-scale fluctuations, generally with an energy barrier E_{ij}. There is also an energy barrier E_{ji} for the reverse transformation; the difference between the barriers is the free energy, so the transformation to the phase of lower free energy is favoured (Fig. 2.3). The time-dependence of the phase transition is modelled as an Arrhenius rate with some associated attempt frequency f_{ij},

$$\mathcal{R}_{ij} = n_{ij} f_{ij} e^{-E_{ij}/k_B T^*}, \qquad (2.19)$$

where n_{ij} is the number density of atoms of i which are able to take part in this process. As with the model for the vacancy formation rate, an effective temperature T^* is used which takes into account non-equilibrium excitation of some vibrational modes.

In general, several processes may allow S_i to transform to S_j. For example, S_j may nucleate more easily at a dislocation, vacancy, grain boundary or impurity than in the bulk material. If multiple processes are possible,

$$\mathcal{R}_{ij} = \sum_k n_{kij} f_{kij} e^{-E_{kij}/k_B T^*}. \tag{2.20}$$

The attempt rates f are related closely to the phonon modes in each phase. These and the energy barriers vary with ρ and T^*. The density of atoms taking part, n, depends on the process (e.g. bulk or defect-nucleated).

Once a region of S_j has formed within the S_i, it grows outwards to consume the rest of the S_i. In general the growth process is non-isotropic. For finite-sized regions of S_j, there is generally an elastic strain energy as the lattices of S_i and S_j will generally need to distort in order to accommodate each other. The number of atoms taking part varies as the surface area of the growing region of S_j. If the region does not grow isotropically, then n_{kij} can vary with the volume fraction of S_j according to a shape function (in the simplest case, the surface area of a sphere). As with ignition and growth models in reactive flow [26, 27, 28], these geometrical terms eventually change from 'growth outwards' to 'growth inwards' when the regions of S_j coalesce to leave isolated regions of S_i. Rather than including the elastic strain energy explicitly, it is proposed to modify the energy barrier as the size of the region of S_j changes.

In a compound, the nucleation and growth rates might require atoms to hop over more than one barrier, e.g. to diffuse to a new position in a complicated structure. The Arrhenius formulation above can be generalised to treat multiple jumps.

Given a calculation of the nucleation rate $\mathcal{R}$, it is possible to estimate the spacing between each region of S_j. If the transformation rate is $\mathcal{R}$ per atom, then one nucleus of the new phase forms per

$$N = \left[\int \mathcal{R}\, dt\right]^{-1} \tag{2.21}$$

atoms. As the initial phase is transformed by growth from existing nuclei, a smaller volume is available for new nuclei to form. (The strain field generated in the partially transformed mixture might also influence the rate of producing new nuclei.) This result can be used to estimate the resulting grain size when the material transforms fully to S_j.

This model is sufficiently general to treat materials with multiple polymorphic phases, and energetics of transformation where metastable phases may form. At any instant, the material consists of fractions v_i of each phase. Each phase may in principle transform to any other phase via the set of rates $\mathcal{R}_{ij}$, which vary with ρ and T. (Accurate calculations of the $\mathcal{R}_{ij}$ may require additional microstructural information such as the number density of nuclei and the distribution of their sizes.) These rates can be integrated to obtain the time history of the transformation.

As described, the model requires a large amount of calibration against atomistic simulations to obtain energy barriers, attempt frequencies near defects etc. However, as a first estimate, a single Arrhenius rate can be used for each transformation based on a 'plausible' value for the energy barrier (and using the free energies from the EOS to predict the difference between forward and reverse rates), attempt frequencies from an average of the phonon frequencies, and simple geometrical terms for the growth from nuclei of the new phase. If the phase transition is diffusive rather than displacive, the Arrhenius rate should represent the rate-determining step.

Condensed/vapour two-phase region

In the polymorphism approach used here, the liquid/vapour region (Fig. 2.4) (and in principle the solid/vapour region, though the effect of the vapour is negligible for most materials) is treated in the same way as any other phase transformation.

Slightly different models were proposed for fluids of high and low density. Ideally these should give the same results where they meet at densities around the critical point. If not, a spurious additional 'phase transformation' would be predicted between the different models. It may be necessary to interpolate or to use an alternative way of predicting fluid properties (e.g. direct numerical simulation) in this region.

2.1.5 Non-equilibrium polymorphism in macroscopic simulations

If macroscopic simulations are to include the dynamics of phase changes explicitly, an EOS must be carried for each phase S_i and a rate law $\mathcal{R}_{ij}$ for the transformation between each pair of phases. Each macroscopic region of material is then described by the set of volume fractions f_i of the phases.

At the microstructural level, phase transitions proceed by the local reconfiguration of the atoms. As a small region of material changes phase, in general the state will not be in thermal or mechanical equilibrium with its surroundings, though the equilibration make take place quickly. In a dynamical situation, the external constraints may change during the course of the phase transition, i.e. when a mixture of phases is present. Since the different phases have different mechanical properties, their thermodynamic state may not change at the same rate. The challenge for a macroscopic calculation is, given the instantaneous f_i and thermodynamic state in each phase, calculate the evolution of the f_i and state in each phase over a finite interval of time. The time integration of the phase transformation and equilibration processes may be operator-split from the hydrodynamics, as discussed later in the context of hydrocodes. For simplicity, it is assumed here that this has been done, and the problem is to predict the evolution of the states and f_i at constant density and energy.

The phase change process converts material between types i. Considering phase changes independently of hydrodynamics and equilibration, material is transferred between components such that the density of the source is unchanged. (If phase changes could proceed in either direction then operator splitting could be applied to each phase change, or the net phase change calculated between each pair i, j and used in the net direction only.) This is a constant volume phase change, but the density ρ_j of the 'product' species will alter with the addition of reactant material of density ρ_i:

$$\Delta f_j = \Delta\lambda_{ij} f_i/\lambda_i, \quad \Delta f_i = -\Delta f_j, \tag{2.22}$$

where $\Delta\lambda_{ij}$ is the change in the mass fraction of component i which is added to component j (obtained by integrating the phase change rate $\mathcal{R}_{ij}$ over the time interval $\delta t'$); then

$$\Delta m_{ij} = \Delta f_j \rho_i, \quad \Delta E_{ij} = \Delta m_{ij}(e_i + q_{ij}), \tag{2.23}$$

are the densities of mass and internal energy transferred from i to j, where q_{ij} is any specific energy released in converting i to j, so

$$\rho_j \to \rho_j' = \frac{f_j\rho_j + \Delta m_{ij}}{f_j + \Delta f_j}, \quad e_j \to e_j' = \frac{f_j\rho_j e_j + \Delta E_{ij}}{\rho_j'(f_j + \Delta f_j)}. \tag{2.24}$$

The change in a free parameter x required to achieve equilibrium of a dependent parameter ϕ can be estimated as

$$\Delta x = -\frac{\phi - \bar{\phi}}{\partial\phi/\partial x}, \tag{2.25}$$

where $\bar{\phi}$ is the average value of ϕ which would be achieved in equilibrium. This calculation is like a Newton-Raphson approach to the root of an equation. For a conserved quantity, $\sum \Delta x_i = 0$. This may not be the case in practice because of rounding errors. If the residual is some value $\sum \Delta x_i = \epsilon$ say (which should be small) then the changes can be made conservative by taking

$$\Delta x_i' = \Delta x_i - f_i \epsilon. \tag{2.26}$$

If perfect equilibrium is desired, the parameter x can be changed until $\phi_i - \bar{\phi} \leq \eta$ for some tolerance η. For numerical stability, some fraction χ of the predicted $\Delta x_i'$ can be taken and the process repeated until equilibrium is reached. If there is a characteristic time τ associated with equilibration, then an exponential approach to equilibrium can be modelled by taking

$$\chi = 1 - e^{-\delta t'/\tau}. \tag{2.27}$$

For operator-split pressure equilibrium, ϕ is p_i and x is f_i. Now,

$$\frac{\partial p_i}{\partial f_i} = \frac{\partial p_i}{\partial \rho_i}\frac{\partial \rho_i}{\partial f_i}, \quad : \quad \rho_i = \frac{m_i}{V_i} = \frac{m_i}{V f_i} \quad \Rightarrow \quad \frac{\partial \rho_i}{\partial f_i} = -\frac{\rho_i}{f_i} \quad \Rightarrow \quad \Delta f_i = f_i \frac{p_i - \bar{p}}{\rho_i c_i^2} \tag{2.28}$$

where c_i is the bulk sound speed. If equilibrium takes place by isentropic expansion, the change in volume fraction is accompanied by a change in specific internal energy,

$$\Delta e_i \simeq -p_i \Delta v_i \simeq -\frac{\bar{p}\Delta f_i}{\bar{\rho}}. \tag{2.29}$$

For operator-split thermal equilibrium, ϕ is T_i and x is e_i. Then

$$\Delta e_i = -c_{vi}\left(T_i - \bar{T}\right), \tag{2.30}$$

where c_{vi} is the specific heat capacity of component i at constant volume.

2.1.6 Macroscopic model of non-equilibrium vibrations

As discussed, an effective temperature T^* is needed to relate time-dependent processes driven by vibrations during rapid deformations, because anisotropic stresses applied at high rates will excite different vibrational modes at different rates. The excitation of each mode then equilibrates over a finite length of time, essentially at constant energy.

A complete description of non-equilibrium vibrations would involve the rate of change of effective temperature T^* (or excitation level) for each vibrational mode as a function of the strain rate tensor

$$\dot{\epsilon}_{ij} = \frac{\partial[\vec{u}]_i}{\partial[\vec{r}]_j} \tag{2.31}$$

with respect to the current orientation of the lattice vectors [108], and the rate of change in the occupation of each vibrational mode given the occupation of all the other modes. The coupling between T^* for a mode of wavevector k and the strain rate tensor has the form

$$\dot{T}^*(\vec{k}) \propto |\dot{\epsilon}_{ij}.[\vec{k}]_j|^2, \tag{2.32}$$

where the constant of proportionality is chosen so that the integration over all modes gives the equilibrium temperature rise $\dot{T}$. (More rigorously, the coupling depends on the Fourier decomposition of $\dot{\epsilon}_{ij}$ in the $\vec{k}$ direction.) The coupling between different modes can be determined from the phonon-phonon interactions (anharmonic lattice contributions), and have the form

$$\dot{T}^* \propto T - T^*. \tag{2.33}$$

This level of detail is unnecessary and probably impractical for macroscopic simulations, but is important for understanding the processes involved and for predicting rates in a given situation. If a particular vibrational mode can be found which dominates in the process of interest, it can be followed in isolation. The parameters needed to describe the non-equilibrium process are then the state (ρ, e), the orientation of the appropriate vibrational mode $\vec{k}$ (following deformation of the material) and the effective temperature T^*.

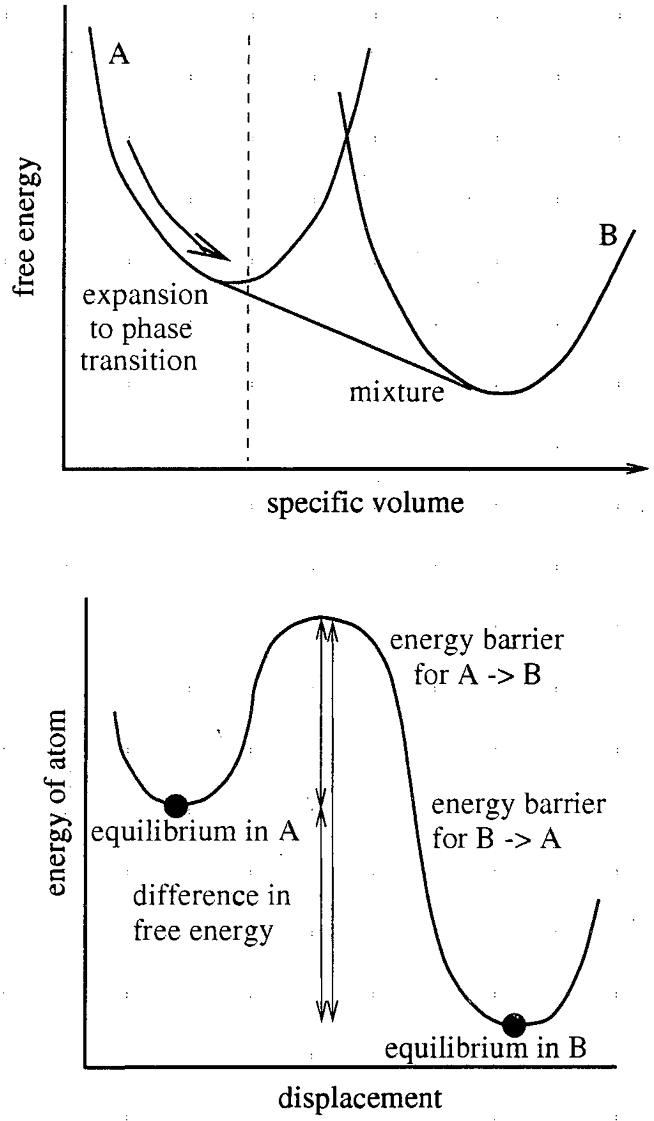
A
B
free energy
expansion to phase transition
mixture
specific volume
energy barrier for A -> B
energy barrier for B -> A
energy of atom
equilibrium in A
difference in free energy
equilibrium in B
displacement

Figure 2.3. Directionality of energy barrier to phase transition. The shape of the curve alters with p and T.

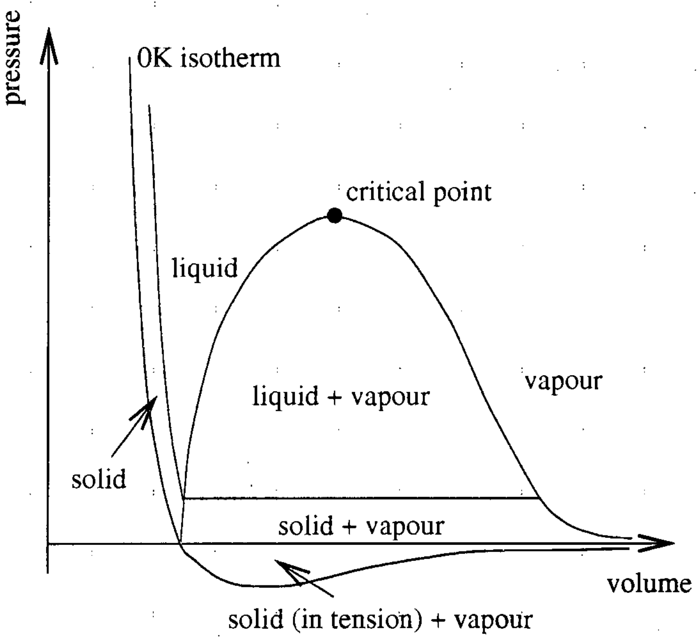
pressure
0K isotherm
critical point
liquid
liquid + vapour
vapour
solid
solid + vapour
volume
solid (in tension) + vapour

Figure 2.4. Schematic of the liquid – vapour region.

2.2 Electron ground states

If the configuration (relative position) of a set of atoms is held constant, the scale varied (changing the density), and the ground state of the electrons found with respect to stationary ions, the resulting variation of ground state energy with density ρ or specific volume v is called the 'frozen-ion cold curve', $e_c(v)$.

Although modern computers are powerful enough to allow reasonable calculations of systems containing several atoms with all the electrons treated explicitly, it is computationally much more efficient to model the core electrons – which are affected relatively little by each atom's environment – by simpler techniques. One way is to modify the potential in which the outer electrons move. The modified potential is called a pseudopotential.

Pseudopotentials may be applied in real or reciprocal (k) space. They may have different components for the different values of the quantum number l of the electron's angular momentum, making the potential non-local.

Pseudopotentials are typically expressed in an analytic form, and the parameters are adjusted so some features of the system reproduce those seen experimentally or by a calculation with the full set of electrons. If as is normal the pseudopotential is fitted to an isolated atom, it is then assumed that the core states do not change when the atoms are combined to produce condensed states – the frozen core approximation.

Since electrons are fermions, the wavefunctions must be antisymmetric with respect to particle exchange. This is an awkward requirement to assure in calculations. The usual way to proceed is to make use of the result [29, 30] that the effect of antisymmetrisation and repulsion between electrons (exchange and correlation) is exactly equivalent to a contribution in the Hamiltonian which is a functional of the electron density – 'density functional theory'. Unfortunately, no simple way has been found to obtain the appropriate functional; trial functionals are evaluated against explicitly antisymmetric and correlated test cases, such as a uniform electron gas. In most multi-atom calculations, the functional is taken to be a local function of the total electron density. This is the local density approximation (LDA). There are several variants of the LDA. The simplest is to assume that the effect of exchange and correlation is equal to that for a homogeneous electron field of the same density. Better agreement is obtained in some systems when the exchange-correlation energy is modelled as a function of the gradient of the local electron density, e.g. using the generalised gradient approximation

(GGA) [31, 32].

In the standard implementation of the pseudopotential method, the exchange-correlation contribution of the core states is absorbed in the total core energy, and the exchange-correlation energy of the outer electrons in the core region calculated by in effect making a linear expansion of the exchange-correlation operator

$$E_{xc}(\rho_c + \rho_o) \simeq E_{xc}(\rho_c) + E_{xc}(\rho_o). \tag{2.34}$$

This approximation is reasonable when the density ρ_o of the outer electrons in the core region is small.

2.2.1 Dirac band structure

Although most EOS-related calculations of electron ground states use Schrödinger's equation, the Dirac equation [38] is a more accurate representation, as it includes relativistic effects. The importance of relativistic effects on EOS properties is not intuitively obvious, so it would be useful to at least have the capability of including these effects in EOS predictions. Some relativistic effects have been included in other EOS predictions [33, 34, 35, 36, 37].

The writing of electron ground state programs suitable for EOS work is a long process. A Dirac-based program was developed during the course of this work, but not to the point where it was suitable for routine EOS calculations – the program requires further development before it is efficient enough. The theory and algorithms used in the Dirac program are presented here to demonstrate that it is possible to start from the Dirac equation – most commonly used for particle beams – and calculate atomic binding energies using methods that can be applied directly to crystalline materials.

Electronic wavefunctions

Following the well-tested standard model of quantum electrodynamics, we describe electrons in terms of a field $\psi(\vec{r})$ of Dirac bispinors [46]. A Dirac bispinor consists of two sets of two components:

$$\begin{pmatrix} \chi \\ \bar{\chi} \end{pmatrix} \quad : \quad \chi = \begin{pmatrix} \uparrow \\ \downarrow \end{pmatrix}, \tag{2.35}$$

describing the intensity of particle (χ) and antiparticle ($\bar{\chi}$) wavefunctions in spin $+\frac{1}{2}$ ($\uparrow$) and $-\frac{1}{2}$ ($\downarrow$) states. Each of these states is a complex function of position.

To be used correctly, the antiparticle wavefunctions must be set up with caution. The equations used here are based on the Feynman interpretation of the antiparticle properties of Dirac bispinors, namely that 'real' antiparticle states are negative energy eigenvalues of the particle states propagating backwards in time. If 'real' antiparticles were required, it would be advisable to check very carefully that they were represented properly when their effects – such as the sign of the electromagnetic field they generate – were calculated correctly. This should not matter when calculating equations of state.

4-space conventions

Relativistic effects and electromagnetism were treated in terms of quantities which transform as vectors in 4-dimensional space-time ('4-vectors') [44]. There are several possible choices for the definition of 4-vectors, so the definitions used here are described fully below.

The contravariant 4-position vector x^μ is defined as

$$x^\mu \equiv (ct, \vec{r}) \tag{2.36}$$

where c is the speed of light, t the time, and $\vec{r}$ the position vector in 3D space [38]. The (covariant) gradient operator in 4-space is [43]

$$\partial_\mu \equiv (\partial_0, \nabla) = \frac{\partial}{\partial x^\mu}. \tag{2.37}$$

∂_0 is therefore $\dfrac{\partial}{\partial(ct)}$.

The scalar product of two 4-vectors is

$$A.B = A^\mu B_\mu = A_\mu B^\mu. \tag{2.38}$$

Einstein's summation convention is used, so repeated indices are summed over so long as one is raised (contravariant) and the other lowered (covariant). Thus

$$A^\mu B_\mu \equiv \sum_{\mu=0}^{3} A^\mu B_\mu. \tag{2.39}$$

Contravariant and covariant vectors are related through the spacetime metric g,

$$\begin{aligned} A^\mu &= g^{\mu\nu} A_\nu & (2.40) \\ A_\mu &= g_{\mu\nu} A^\nu & (2.41) \\ g_{\mu\nu} &= g^{\mu\nu}. & (2.42) \end{aligned}$$

Thus,

$$A.B = g_{\mu\nu} A^\mu B^\mu = g^{\mu\nu} A_\mu B_\mu \tag{2.43}$$

In the absence of masses or energies which are large in a general relativistic sense,

$$g = \mathrm{diag}(1, -1, -1, -1). \tag{2.44}$$

Relativistic electromagnetism

The electromagnetic 4-potential A^μ is related to the electric (scalar, V) and magnetic (3-vector, $\vec{A}$) potentials by [44]

$$A^\mu = (V/c, \vec{A}). \tag{2.45}$$

A^μ can be found by solving Maxwell's equation [38, 44],

$$\Box^2 A^\mu = -\mu_0 j^\mu \tag{2.46}$$

where $\Box^2$ is the d'Alembertian operator $\partial^\mu \partial_\mu$ and μ_0 is the permeability of free space ($4\pi \times 10^{-7}$ H/m). j^μ is the 4-current density [44],

$$j^\mu \equiv q\rho(c, \vec{v}) \tag{2.47}$$

where q is the charge on a particle, ρ the particle density and $\vec{v}$ the particle velocity.

All solutions for the electromagnetic 4-potential satisfy Maxwell's equation. If it is desired to calculate the 4-potential from a known 4-current, Maxwell's equation is not particularly useful, because any number of solutions can be found by adding a constant to the electrostatic potential and a curl-free field to the vector potential. A more practical route is to deduce A^μ directly from the 4-current density. The standard expressions for Gauss' law and the Biot-Savart equation [49] can be combined compactly as

$$A^\mu(\vec{r}) = \frac{\mu_0}{4\pi}\int_V \frac{j^\mu(\vec{r'})}{|\vec{r'}-\vec{r}|}\,d\vec{r'}. \tag{2.48}$$

In effect we are using a Green's function to solve the time-independent Maxwell's equations.

The electromagnetic 4-potential can be used to introduce the effect on a single electron of the nuclei and other electrons. The nuclei can be described accurately as static charges, contributing to V through the Coulomb potential,

$$A^0_{\text{ions}}(\vec{r}) = \frac{1}{c}\sum_i \frac{Z_i e}{4\pi\epsilon_0|\vec{r}-\vec{r}_i|} \tag{2.49}$$

for a set of nuclei of charge $+Z_i e$ at positions $\vec{r}_i$. Other electrons can be described by their associated 4-current and hence a 4-potential through the Biot-Savart equation.

Evolution equation for a single electron

Electron wavefunctions satisfy the Dirac equation [38, 46]

$$(\not{p} - mc)\psi = 0 \tag{2.50}$$

where m is the particle rest-mass. $\not{p}$ is related to the 4-momentum operator p_μ and the local electromagnetic 4-potential A_μ by [43]

$$\not{p} = \gamma^\mu(p_\mu - qA_\mu) \tag{2.51}$$

where γ^μ is a Dirac gamma-matrix and $q = -e$ the charge on an electron. The 4-momentum operator is [38, 50]

$$p_\mu \equiv i\hbar\partial_\mu. \tag{2.52}$$

The gamma matrices are [46]

$$\gamma^\mu \equiv (\beta, \beta\alpha_i), \tag{2.53}$$

where

$$\alpha_i = \tilde{\alpha} \times \sigma_i, \tag{2.54}$$

in which multiplication is used in the sense that each element of $\tilde{\alpha}$ is multiplied by the corresponding *matrix* σ_i in α_i. An example is given below.

The standard form of the Dirac spin matrices is [46]

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{2.55}$$

There is no unique representation for β and $\tilde{\alpha}_i$. Standard choices [46] are the Dirac-Pauli representation

$$\beta = \mathrm{diag}(1, 1, -1, -1), \quad \tilde{\alpha} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{2.56}$$

and the Weyl representation

$$\beta = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \quad \tilde{\alpha} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{2.57}$$

As was stated above, the α_i are defined in terms of the product of $\tilde{\alpha}$ and the σ_i in the sense that each element of $\tilde{\alpha}$ is multiplied by the corresponding σ_i in α_i. For example, in the Weyl representation,

$$\alpha_2 = \begin{pmatrix} 0 & i & 0 & 0 \\ -i & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \end{pmatrix}. \tag{2.58}$$

Observables of systems described by a bispinor field are given by

$$\langle \psi | O | \psi \rangle \equiv \int_V \bar{\psi}(\vec{r}) O \psi(\vec{r}) \, d\vec{r} \tag{2.59}$$

where

$$\bar{\psi} \equiv \psi^\dagger \gamma^0 \tag{2.60}$$

and $\psi^\dagger$ is the transpose of the matrix of complex conjugates of the elements of ψ, i.e. (using $[A]_{ij}$ to denote components i, j of matrix A)

$$[\psi^\dagger]_{ij} \equiv ([\psi]_{ji})^*. \tag{2.61}$$

The Hamiltonian (total energy) of a particle field is given by the Schrödinger relation [50]

$$H = i\hbar \frac{\partial}{\partial t}. \tag{2.62}$$

Applying this to the Dirac equation,

$$H = c(\gamma^0)^{-1} \left(\sum_{a=1}^{3} -i\hbar\gamma^a \partial_a + q\gamma^\mu A_\mu + mc \right). \tag{2.63}$$

From the properties of the Dirac matrices [46], several factors cancel to give

$$H = -i\hbar c\vec{\alpha}.\nabla + q(V - c\vec{\alpha}.\vec{A}) + mc^2\beta. \tag{2.64}$$

This equation is equivalent to that used in previous work [48], but contains additional terms representing the magnetic field $\vec{A}$.

Electrons contribute to the electromagnetic field. Their contribution can be found by applying the Biot-Savart equation to the 4-current density j^μ. The 4-current density from a bispinor field is [46],

$$j^\mu = q\bar{\psi}\gamma^\mu\psi. \tag{2.65}$$

Bispinors for stationary states

Stationary states Ψ_i satisfy

$$H\Psi_i = H_i\Psi_i \tag{2.66}$$

where H_i is the energy of the ith stationary state. This equation applies with the same value H_i to each component of Ψ.

Two approaches may be followed in order to enforce this condition. If the basis set for Ψ freely spans each bispinor component then the solution could be constrained explicitly, e.g. by the use of a Lagrange multiplier,

$$H_i' \equiv H_i + \lambda \sum_j |[H\Psi_i]_j - H_i[\Psi_i]_j|^l, \tag{2.67}$$

where the index j runs over the bispinor components. Alternatively, the basis functions can be chosen to satisfy the constraint. In the Dirac-Pauli representation,

$$H\Psi = \begin{pmatrix} qV + mc^2 & -c\vec{\sigma}.(i\hbar\nabla + q\vec{A}) \\ -c\vec{\sigma}.(i\hbar\nabla + q\vec{A}) & qV - mc^2 \end{pmatrix} \begin{pmatrix} \Psi^+ \\ \Psi^- \end{pmatrix} = H_i \begin{pmatrix} \Psi^+ \\ \Psi^- \end{pmatrix}. \tag{2.68}$$

Positive energy (particle) solutions therefore satisfy

$$\begin{aligned} \Psi^- &= -\frac{c}{\eta}\left[\vec{\sigma}.(i\hbar\nabla + q\vec{A})\right]\Psi^+ && (2.69) \\ \nabla\Psi^- &= \frac{q}{\eta}(\nabla V)\Psi^- - \frac{c}{\eta}\vec{\sigma}.\left[i\hbar\nabla\otimes\nabla + q(\vec{A}\otimes\nabla + \nabla\otimes\vec{A})\right]\Psi^+, && (2.70) \end{aligned}$$

where

$$\eta \equiv H_i + mc^2 - qV. \tag{2.71}$$

When substituted into the Hamiltonian, the additional contraction with $\vec{\sigma}$ leads to the cancellation of the off-diagonal terms in $\nabla\otimes\nabla$, giving

$$\vec{\sigma}.\left[\vec{\sigma}.(\nabla\otimes\nabla)\right] \equiv \nabla^2. \tag{2.72}$$

The resulting kinetic energy term resembles closely the corresponding term in the Schrödinger equation. Since the binding energy of an electron is small compared with its rest-mass,

$$\eta \simeq 2mc^2, \tag{2.73}$$

so

$$-\frac{\hbar^2 c^2}{\eta}\nabla^2\Psi^+ \simeq -\frac{\hbar^2}{2m}\nabla^2\Psi^+. \tag{2.74}$$

It is interesting that, according to this derivation, the Hamiltonian includes contributions from the gradient of each component of the electromagnetic field.

Neglecting these and the rest-mass, and substituting in the equation for the minor components, the Hamiltonian for the major components becomes

$$H^{(+)} = -\frac{\hbar^2 c^2}{\eta}\nabla^2 + qV + \frac{c^2 q}{\eta}\vec{\sigma}.\vec{A}\left[i\hbar\vec{\sigma}.\nabla + q\vec{A}\right]. \tag{2.75}$$

Multiple electrons

It would be nice to follow the quantum field theoretic approach [38, 52] of treating Ψ as a single field representing all the electrons. In this case, Ψ must be a field of Grassmann numbers [52], which are members of an infinite set that anticommute among themselves. Unfortunately, adequate computer representations of Grassmann numbers have not yet been developed.

Instead, a separate wavefunction is calculated for each electron state. The wavefunction ψ_i of each state is expanded in terms of a basis set $\{\phi_j\}$:

$$\psi_i = \sum_j a_{ij}\phi_j, \tag{2.76}$$

for some set of real amplitudes a_{ij}.

An iterative procedure is used to find the wavefunctions. At each iteration n, some guess ${}_{[n]}\psi_i$ has been made of the wavefunctions, and of the occupation number f_i for each state. Given ${}_{[n]}\psi_i$, the Biot – Savart equation can be used to calculate the electromagnetic field generated by an electron in each state. Thus the electromagnetic field A_i^μ experienced by each state can be calculated given the ψ_i and f_i. (For each state i, this may be calculated by considering the wavefunction and occupation of all other states $j \neq i$, or a mean field may be used where the electromagnetic field generated by *all* the states is included. The mean field prescription is appropriate for some conventions of the local density approximation.)

For stability, an under-relaxation procedure is used when calculating the electromagnetic field experienced by each state at each iteration:

$${}_{[n]}A_i^\mu = (1-\lambda)_{[n-1]}A_i^\mu + \lambda A_i^\mu(\{{}_{[n]}\psi_i\}) \tag{2.77}$$

where $A_i^\mu(\{{}_{[n]}\psi_i\}$ is the electromagnetic potential determined from the $\{{}_{[n]}\psi_i\}$ and λ is the under-relaxation factor. The state occupations are under-relaxed in a similar manner.

If the ψ_i are orthonormal, then given ${}_{[n]}\psi_i$ and ${}_{[n]}A_i^\mu$, we can use the Hamiltonian equation

$$H\psi_i = E_i\psi_i, \tag{2.78}$$

multiply through by ψ_j and integrate over all space to obtain

$$\langle\psi_i|H|\psi_j\rangle - \delta_{ij}E_j = 0. \tag{2.79}$$

If the ψ_i are not orthonormal, then

$$\begin{aligned} \langle\psi_i|H|\psi_j\rangle &= E_j\langle\psi_i|\psi_j\rangle, && (2.80)\\ \Rightarrow \quad \langle\psi_i|H|\psi_k\rangle . \left[\langle\psi_k|\psi_j\rangle\right]^{-1} - \delta_{ij}E_j &= 0. && (2.81) \end{aligned}$$

The E_i and eigenvectors Q_{ij} are found from the solution of this eigenproblem. If the ψ_i are orthogonal then the eigenproblem is real and symmetric.

The modified set of eigenstates for use in the next iteration is then

$${}_{[n+1]}\psi_i = \sum_j {}_{[n+1]}a_{ij}\phi_j \quad : \quad {}_{[n+1]}a_{ij} = \sum_j Q_{ik[n]}a_{kj}. \tag{2.82}$$

The local density approximation can be used to find the contribution of exchange and correlation to the total energy (it does not affect the individual eigenstates). As a partial alternative, the exchange integral can be included in the calculation of the Hamiltonian matrix. The non-relativistic form is [42, 53]:

$$\delta H_{ij} = -\frac{1}{2}\frac{q^2}{4\pi\epsilon_0}\sum_{i\neq j}\int \bar{\psi}_i(\vec{r}_1)\bar{\psi}_j(\vec{r}_2)\frac{1}{|\vec{r}_1-\vec{r}_2|}\psi_i(\vec{r}_2)\psi_j(\vec{r}_1)\,d\vec{r}_1\,d\vec{r}_2. \tag{2.83}$$

If this term is included with basis functions of the determinantal type [53] then a Dirac – Hartree – Fock calculation is obtained.

The set of states ψ_i with energies E_i is calculated for a specific set of occupations f_i. This prescription works at finite temperatures, since the occupations can be calculated so as to correspond to excitation into states above the Fermi energy. The algorithm used to estimate the population of each band at finite temperatures is described in Section 2.4. Since the electromagnetic potential experienced by each state depends on the occupations of all other states (and of the same state, in the mean field approximation), the calculation of chemical potential should also be calculated self-consistently as part of the iterative scheme.

The chemical potential is only located to an accuracy sufficient to constrain the total number of electrons. At temperatures which are low compared with the spacing between the energy levels, the uncertainty in the chemical potential is comparable with the spacing between its adjacent energy levels.

If this is done then the electron states should be accurate even for temperatures which are high in comparison with the ionisation energies. The band structure is predicted in any case; the extra information obtained is the temperature-dependence of the band structure, or a rigorous calculation of the electron-thermal energy. However, the local density functionals may not be valid for $T > 0$.

A different approach might be possible for calculating the electron energy at finite temperatures. A temperature could be defined by modifying the electromagnetic background A^{μ} to reflect some appropriate spectrum of photons, e.g. a black body spectrum at the desired temperature. The electron ground state with respect to this potential may model the excitation of electrons from the ground state at zero temperature. The electromagnetic field would have to be made complex, reflecting the time-variation of the photon field.

If the eigenstates are expected to be significantly different from the initial guess or the basis set, then the Hamiltonian matrix should be calculated for a significantly larger set of states than are needed to contain all the electrons. The states obtained comprise the electron band structure.

At present, the code calculates as many eigenstates as it can. It would be more efficient when calculating ground states (though not band structures) to calculate only sufficient eigenstates to accommodate all the electrons. This development may be worth pursuing in the future.

Density functional forms

Functional forms included in the program are [53]:

Slater :

$$e_x(\vec{r}) = a n^b(\vec{r}) \tag{2.84}$$

for parameters a and b.

'Standard' Kohn-Sham parameters for exchange are, in atomic units, [41]

$$a = -\frac{3}{4\pi}(3\pi^2)^{1/3} \tag{2.85}$$

$$b = 4/3 \tag{2.86}$$

or [53]

$$a = -\frac{3}{\pi}(3\pi^2)^{1/3}. \tag{2.87}$$

The functional implemented in the software is

$$e_{xc} = -6\alpha \left(\frac{3n(\vec{r})}{8\pi}\right)^{\beta} n(\vec{r}). \tag{2.88}$$

Hedin-Lundqvist :

$$e_{xc}(\vec{r}) = e_x(\vec{r}) + e_c(\vec{r}), \tag{2.89}$$

where e_x is the Slater functional, and

$$e_c = -cg(x)n(\vec{r}) \tag{2.90}$$

where

$$x = r_s/a, \tag{2.91}$$

$$c = (2b/\pi\eta a)^{1/3}, \tag{2.92}$$

$$\eta = (4/9\pi)^{1/3}, \tag{2.93}$$

$$g(x) = (1+x^3)\ln(1+1/x) + x/2 - x^2 - 1/3, \tag{2.94}$$

$$\frac{4}{3}\pi n r_s^3 = 1. \tag{2.95}$$

Discretisation

The Dirac equation can be solved using algebraic forms for the field functions ψ and A^μ. The algebraic forms can be differentiated, and in principle integrated to give the Hamiltonian in closed form. The algebraic approach becomes less convenient as the complexity of the system increases.

The approach followed here is to use a discretised approximation to the field equations. If a field function f varies smoothly, it can be approximated to arbitrary accuracy by values averaged over sufficiently small regions δV_i of space around a set of points $\{\vec{r}_i\}$. The integral of f over all space becomes a summation:

$$\int_V f(\vec{r})\, d\vec{r} \simeq \sum_i f(\vec{r}_i)\delta V_i. \tag{2.96}$$

Basis functions

The set of eigenfunctions $\{\psi_i\}$ of the Hamiltonian can sometimes be found in algebraic form for sufficiently simple systems. Examples are the harmonic oscillator and the hydrogen atom. If the eigenfunctions are known for a system with a single particle, then the ground state is identical to the eigenfunction with lowest energy.

In general, it is not straightforward to find algebraic expressions for the eigenfunctions. However, any function ψ can be rewritten exactly in terms of a set of functions $\{\phi_i\}$ which completely spans the function (Hilbert) space containing ψ:

$$\psi = \sum_i c_i\phi_i \tag{2.97}$$

for some set of real amplitudes $\{c_i\}$. For example, if ψ is defined as a $\mathcal{C}$-valued function over all of $\mathcal{R}^3$, then it can be written in terms of the sum of the set of plane waves:

$$\psi(\vec{r}) = \sum_{\vec{k}_i} c_i e^{i(\vec{k}_i.\vec{r}-\omega_i t)} \tag{2.98}$$

where $\{\vec{k}_i\}$ is the set of all wavevectors (points in $\mathcal{R}^3$) and ω_i is the frequency of the mode $\vec{k}_i$, given by its energy, $E_i = \hbar\omega_i$.

The set of basis functions $\{\phi_i\}$ required to span the space of smooth functions over any finite region is infinite – a practical difficulty. However, the objective may be to predict the ground state to a finite accuracy; for instance to compare against experimental data with a finite accuracy. Any function ψ can be represented *to some accuracy* by expansion in terms of a *finite* set of basis states, N say:

$$\psi \simeq \sum_{i=1}^{N} c_i\phi_i. \tag{2.99}$$

If the $\{\phi_i\}$ are similar to the eigenstates of the Hamiltonian then only a few terms in the expansion may be needed.

In the presence of nuclei, many of the electron states become localised. Inner electrons are closely bound to the region about the nucleus. Depending on the shell structure, outer electrons may be delocalised as in the conduction bands of a metal, or localised as in the bonding orbitals of covalent atoms.

The archetypal series of localised basis states is the spherical harmonics. These provide a description of the eigenstates of a $(Z-1)$-ionised atom of atomic number Z – i.e. a nucleus with a single electron in orbit – which is exact to the accuracy of non-relativistic, spin-0 quantum mechanics. Localised states can often be described accurately in terms of a few spherical harmonics even when multiple electrons are present.

The archetypal series of delocalised basis states is the plane waves. These can be used to describe electron beams very accurately. Plane waves have advantages in their computational convenience, as they are fast to compute and it is easy to perform Fourier transforms on them.

Atomic conduction bands contain a significant plane wave element, but can also be quite localised. Methods of calculating the electron band structure in condensed matter have been developed to handle this effect by combining spherical harmonics and plane waves. Such work has resulted in the orthogonalised plane wave and pseudopotential techniques.

The choice of basis function has important consequences for the ability of a model to give accurate predictions of the ground state of a system. One way of implementing such calculations is to evaluate the basis functions in algebraic form at appropriate points in real or reciprocal space. The problem with this approach is that it can lead to a scheme which exploits the features of a particular basis set to full advantage, and which is therefore restricted to problems which are well-described by that basis set.

The approach followed here is to allow the basis set to be described in a discrete fashion in terms of its values at a set of points $\{\vec{r}_i\}$. To solve the Dirac equation, the gradient of each basis function must also be supplied at each of the points.

Since the wavefunction describes Dirac bispinors, each basis function describes the variation of a Dirac bispinor and its gradient with position. Thus at each point, each basis function must be defined as a 4-component array of complex numbers, and the gradient must be defined as a 3×4 matrix of complex numbers (i.e. 3 rows, 4 columns).

A calculation can be made without spin polarisation by making the spin-up and spin-down components equal in all basis functions.

Preprocessing programs were written to generate basis functions including plane waves, the first few spherical harmonics, Gaussians, or tabulated radial functions. In the case of the localised functions, multiple centres could be defined.

Symmetries and effects of the periodic lattice

The program calculates electron wavefunctions over a finite region of space.

Any functions defined over the region should exhibit the correct symmetry properties. In the current software implementation, the functions must have the appropriate symmetry as defined. This applies to the basis functions and the background electromagnetic potential. The symmetry is reflected in the volume weightings δV_i associated with ψ at each point.

Symmetry is also taken into account when calculating the electromagnetic field at each point caused by the charge current density at all other points. Applying symmetry to the current density is more than just a weighting, because symmetry operations can mix the components of the current density. For example, a rotation symmetry of 90° anticlockwise about the z-axis maps the x-component to the y-component, and the y-component to $-x$. For this reason, a 4×4 matrix must be defined between each pair of points to relate the current density at one (along with all its images from symmetry) to the electromagnetic potential at the other.

Stand-alone preprocessing programs were written to apply symmetry to fields and to calculate electromagnetic contributions.

An interesting consequence of this approach is that it should be possible to use symmetry operators which map to regions moving at different velocities. If anyone has an application for a spinning, shearing disc, then this scheme should cope.

The discrete, symmetrised expression of the Biot-Savart equation is

$$A^{\mu}(\vec{r}_i) = \frac{\mu_0}{4\pi} \sum_j \Omega_{ij} j^{\mu}(\vec{r}_j) \tag{2.100}$$

where

$$\Omega_{ij} = \sum_{\text{images of } j} \int_{\delta V_j} \frac{d\vec{r}}{|\vec{r} - \vec{r}_i|}, \tag{2.101}$$

where the integral is over the region considered to have a wavefunction sufficiently constant to be represented by a single point. The summation over 'images' refers to the images of a region obtained by applying the symmetry operations. The diagonal components of the matrices, Ω_{ii}, are not zero in general.

To summarise the requirements for using symmetry in the program, the following must be specified:

- Basis functions with the correct symmetry properties. They must join smoothly with their images under symmetry.
- The background electromagnetic potential with the correct symmetry.
- The volume associated with each point by symmetry, δV_i, so enable the integrals of ψ and ρ to be calculated correctly.
- The translation matrices which relate the charge current at any point (and all its images) to the electromagnetic potential at every other point.

The scheme for using symmetry described here handles a periodic lattice as well as symmetry operations inside each lattice cell. The 'joining smoothly' requirement is the same as applying periodic boundary conditions. The electromagnetic contributions from the translation operators which map a single lattice cell to the entire lattice add the energy caused by electromagnetic interactions between lattice cells.

Ewald sum

When the electromagnetic matrices and background field are calculated before use by this scheme, the summations must be expressed such that the $1/r$ terms from the Coulomb potential associated with the ions and the lowest moment of the electron charge distribution in each lattice cell are allowed to cancel exactly. This is the continuum version of the 'Ewald sum' performed for the point charges in an ionic crystal [40]. If this is not done, the contributions to the electrostatic energy of an infinite crystal from both the positive and negative charges will be infinite, because the sum over monopoles is divergent.

Consider the Gauss equation for the electrostatic potential:

$$\phi(\vec{r}) = \frac{q}{4\pi\epsilon_0}\left[\sum_{\text{ions}} \frac{Z_i}{|\vec{r}-\vec{r}_i|} - \int \frac{\rho(\vec{r}')}{|\vec{r}-\vec{r}'|}\, d\vec{r}'\right] \tag{2.102}$$

where q is the quantum of charge, Z_i is the net charge number of ion i and $\rho(\vec{r})$ is the electron probability density at point $\vec{r}$. Choosing a reference point $\vec{r}_c$, this equation can be expressed as

$$\phi(\vec{r}) = \phi_0(\vec{r}) + \phi_1(\vec{r}) \tag{2.103}$$

where

$$\phi_0(\vec{r}) \equiv \frac{q}{4\pi\epsilon_0} \frac{1}{|\vec{r} - \vec{r}_c|} \left[\sum_{\text{ions}} Z_i - \int \rho(\vec{r}')\, d\vec{r}' \right] \tag{2.104}$$

and

$$\begin{aligned} \phi_1(\vec{r}) \equiv \ & \frac{q}{4\pi\epsilon_0} \left\{ \sum_{\text{ions}} Z_i \left[\frac{1}{|\vec{r} - \vec{r}_i|} - \frac{1}{|\vec{r} - \vec{r}_c|} \right] \right. \\ & \left. - \int \rho(\vec{r}') \left[\frac{1}{|\vec{r} - \vec{r}'|} - \frac{1}{|\vec{r} - \vec{r}_c|} \right] d\vec{r}' \right\} \end{aligned} \tag{2.105}$$

If the net charge on the lattice is zero, then $\phi_0 = 0$ everywhere. ϕ_1 consists of dipole contributions – the moments of the charge distribution about $\vec{r}_c$ – and is therefore better-behaved than the summations in ϕ.

The $\vec{r}_c$ can be either a single point in space, or a separate point for each lattice cell (if each cell is electrically neutral). The latter has the better convergence properties, and should be used in practice.

The first term in ϕ_1 defines the electrostatic background. The second term defines the potential from the electron distribution. This equation can be generalised slightly to give the 4-potential.

Background fields

The external electromagnetic potential is supplied as a constant background field, usually used to bring in the effect of atomic nuclei. Additional background fields may be supplied for convenience: a 4-current density and constant wavefunctions, either of which may be used to represent electron fields which are to take place in the calculation only by their electromagnetic field and contribution to the local density functional.

Units

The Dirac program is intended to contain no assumptions about units. The values to certain physical parameters must be specified when the program is run. The parameters include the speed of light c and the permeability of free space μ_0. All quantities must be specified in a consistent set of units.

Particle physics is often performed, and the Dirac equation used, in 'natural units,' where $c = \hbar = 1$, and quantities such as mass and momentum are measured in convenient energy units, such as GeV. It should be possible to run the program using natural, SI or any other set of units. The program has so far been tested only in natural units. It should be safe to perform calculations in other units, but it is possible that some of the necessary constants have been omitted in some equations.

Program structure

The Dirac band structure program was written in C++, using the WXC++ class library. The program operates as follows:

Input:

- Read physical constants, sigma and gamma matrices, etc.
- Read LDA functional.
- Read problem description: number of field quanta (i.e. electrons), number of spatial points, volumes, background 4-potential, background 4-current density, electromagnetic connection matrices.
- Read basis functions, then normalise or orthogonalise if requested.
- Read initial guess at electron states (i.e. amplitudes with respect to basis functions, and occupation numbers), and choose subset to include in the optimisation.
- Read reference energies for calculating the minor components of the bispinor in each state. (These are replaced by the state energy in subsequent iterations of the optimisation scheme.)
- Read parameters to control optimisation: maximum change in total energy between iterations, numerical bandwidths, maximum numbers of iterations, relaxation factors for the 4-potential and occupation numbers.
- Read parameters to control the population of the states at finite temperatures (bisection accuracy and maximum iterations).

Prepare initial fields:

1. Calculate major spinor field for each state using basis functions and initial guess at amplitudes.
2. Calculate minor spinor field using guessed reference energies and the background 4-potential.
3. Normalise wavefunctions if requested.
4. Add effect of background 4-current to background potential.

Iterative optimisation: repeat for each iteration:

1. Calculate the 4-potential generated by each state.
2. Apply the under-relaxation algorithm to each 4-potential.
3. Form the Hamiltonian matrix, including the 4-potential *seen by* each state, and adding the Fock term if requested.
4. Calculate eigenvalues and eigenvectors, sorting in order of increasing state energy.
5. Predict occupation numbers at the requested temperature.
6. Apply the under-relaxation algorithm to the occupation numbers.
7. Calculate major spinor field for each state using basis functions and the amplitudes predicted from the eigenvectors of the Hamiltonian matrix.
8. Calculate minor spinor field using state energies and the 4-potential seen by each state.
9. Normalise wavefunctions if requested.
10. Calculate new 4-current density.
11. Calculate LDA energy and hence total energy.
12. Stop if energy has changed little from the previous iteration.

An electronic manual was written to document the input data more precisely.

Performance

As designed, the Dirac program calculates the wavefunction states and occupations for a set of electrons in a background potential. Arbitrary basis functions may be specified. The program makes no logical distinction between electrons around an isolated atom and electrons in a crystalline solid. [1] Thus it should be possible to test the physical equations and algorithms in the program by performing calculations of the binding energy of isolated atoms.

Calculations were performed of H, He^+ and He, in natural units with an energy scale of 1 eV. The wavefunctions were evaluated at 50 points in a radial line from the nucleus, the spacing between the points increasing geometrically.

It was found to be crucial to use basis functions and wavefunctions that were normalised to a high degree of accuracy. Any deviation ϵ from a norm of unity resulted in an error of ϵ times the rest energy of the electron – which could be much larger than the binding energy. The program was modified to allow basis functions and wavefunctions to be normalised during execution, and also to allow basis functions to be orthogonalised using a modified Gram – Schmidt scheme (Appendix H). These measures cured the problem. (There are complexities associated with simply removing the rest-mass from the Dirac Hamiltonian.)

The He calculations included a spin-up and a spin-down electron, each in the He 1s state. Calculations were performed in which each electron experienced the electromagnetic field of the nucleus plus the other electron ('individual fields'), and also in which each electron experienced the field of the nucleus plus both electrons ('mean field') – i.e. including itself.

Table 2.1 summarises binding energies obtained in the calculations. The single electron calculations were as accurate as might be hoped, given the relatively coarse spatial mesh. Multi-electron calculations using the Slater functional in the local density approximation gave quite accurate results for He with $\alpha \sim 2/3$ (2.4×10^{-3} in natural units) and the individual field model. This is a standard value for the local density approximation, and indicates that the multi-electron model performs correctly.

Calculations were performed to demonstrate that the energy did not change significantly as additional wavefunctions were added. As the number of allowed states increased, it seemed necessary to decrease the under-relaxation factors.

Calculations were performed at temperatures comparable with the difference in energy levels. The chemical potential decreased as temperature increased, and the occupations spread out among the available states.

No calculations were made to test the Fock term.

---

[1] At present, however, it is considerably more inconvenient to set up calculations with a repeating lattice cell, or with large numbers of basis functions. This should be a relatively straightforward task of making the process of setting up more automatic.

Table 2.1. Binding energies.

	experiment [54] (eV)	calculation (eV)	notes
H	13.7	13.61	
He^+	54.51	54.44	
He	79.21	40.95	individual fields, no LDA
		26.97	mean field, no LDA
		99.30	individual fields, Slater LDA ($\alpha = 1$)
		31.37	mean field, Slater LDA ($\alpha = 1$)
		79.33	individual fields, Slater LDA ($\alpha \sim 2/3$)

The results obtained suggest that the Dirac-based method developed should be capable of predicting ground state energies for atoms of higher atomic number, and of treating condensed matter. However, it is likely that improvements will have to be made to the numerical scheme to enable it to handle many electrons in an efficient and robust manner.

2.2.2 CASTEP

CASTEP [55, 56, 57, 58] is a computer program which finds the ground state of electrons with respect to a periodic array of potentials. The ground state is found from Schrödinger's equation, so CASTEP represents the outer electrons in a non-relativistic manner. However, the pseudopotentials may be generated so as to include relativistic effects in the inner electrons.

The wavefunction of the outer electrons is expanded as a series of plane waves. LDA and GGA are used to include the effects of exchange and correlation in a computationally convenient manner. CASTEP optimises the amplitudes of the basis set to find the ground state of the electrons.

Since the number of basis states is finite, there is an energy cutoff in the expansion. Analytic schemes can be used to extrapolate to an infinite cutoff, and to interpolate between jumps caused by the changing integral number of plane waves used in the basis set as the lattice cell changes size. This additional contribution is known as the Pulay stress [59, 60, 20, 61].

The electron field is represented on a discrete grid of locations in k-space. A high density of k-points is necessary to resolve the Fermi surface accurately in metals, although an entropy correcting scheme has been devised to reduce the density required [62].

A Gaussian smearing is applied to the electron energy levels during the energy minimisation to allow more interaction between states. The Gaussian width is decreased between successive iterations of the minimisation process. The effect of finite temperature on the outer electron states can be modelled by retaining some Gaussian smearing in the energy levels. The width of the Gaussian increases with temperature.

Symmetry operations are used to reduce the amount of computation required, permitting a calculation to be made using only those points in the wavefunction grid which are not related to each other by symmetry.

CASTEP can be used to perform fundamental 'molecular dynamics' – integrating forces to find the motion of the atoms – and the lattice cell can be allowed to relax to equilibrium with an applied pressure [57].

As well as producing the outer electron field at the chosen k-sites, the program can calculate several other useful quantities, such as the electronic band structure, the Fermi energy, the components of stress on the lattice cell and the force on each atom.

The pseudopotentials can be specified in real or reciprocal (k) space, and may have non-local (angular momentum-dependent) contributions. Exchange and correlation are modelled using the local density approximation, optionally with gradient corrections to the basic density-only approximation [31, 32, 58].

Ab initio pseudopotentials can be found by considering the energy levels or scattering cross-sections for the outer electrons of an isolated atom. These can be obtained from a rigorous all-electron structure calculation using a spherical harmonic basis set. The parameters in an empirical pseudopotential can then be optimised to reproduce the all-electron results. Pseudopotentials of sufficient complexity (e.g. involving an expansion in l-dependent or non-local terms) have been found to be transferable to condensed states.

The principal concern in using this scheme is the assumption that the pseudopotential does not change with compression. One measure of the validity of this 'constant pseudopotential' approximation is that the pseudopotentials should not overlap. This consideration places a constraint on the maximum compression allowable for each polymorphic phase.

2.2.3 INFERNO

For comparison, some calculations were also made with the 'INFERNO' code [48], which has been used previously for calculating EOS at high temperatures and pressures [19]. INFERNO solves the Dirac equation for an atom in jellium [41]. Exchange and correlation effects in the electron fields are included through the local density approximation (LDA).

The atom-in-jellium model was developed as a generalisation of Thomas-Fermi-Dirac (TFD) theory to represent condensed matter at densities between those where TFD is sufficient and those where band structure calculations are normally performed.

A self-consistent-field (SCF) solution is found for the coupled one-electron Dirac equations in a spherical region around the nucleus. The field outside this spherical region is modelled as 'jellium' – a uniform electron gas. For computational convenience, the energy of the jellium is approximated by its volume average – the muffin-tin approximation. The charge density is represented on a finite radial mesh.

The atom-in-jellium model has been extended to predict approximate lattice properties, by displacing the jellium and calculating the restoring force [66]. This problem can be solved by perturbation theory.

Finite temperatures are modelled using the average atom or mean field approximation. Numerical difficulties can be encountered at low temperatures (below 0.1 eV).

The atom-in-jellium calculations performed by INFERNO require much less computational effort than solving the corresponding band-structure problem. However, INFERNO cannot distinguish between different polymorphic phases, so it cannot predict phase diagrams. Furthermore, the model of atomic vibrations predicts only a single frequency, which is an extreme simplification of the phonon states found in condensed matter.

Version 80 of INFERNO was used, with slight modifications by the author to deal with some compiler complaints.

Atom-in-jellium models

The properties of the atom (such as energies) include significant contributions from the jellium. When calculating these properties, it is necessary to separate the contribution of the continuum electron states in the jellium. Three techniques have been developed to do this [48]:

Model A : Attribute to the atom the part of each energy integral originating inside the atomic sphere.

Model B : Calculate thermodynamic quantities from the difference in the chemical potential when an atom is inserted in the uniform electron gas, making it expand by the volume of the atomic sphere.

Model T : Calculate thermodynamic quantities from the difference in the total system volume when an atom is inserted in the uniform electron gas at constant chemical potential.

Since these techniques are based on fairly arbitrary assumptions, it is difficult to choose between them on theoretical grounds. The alternative models were compared, as described below.

2.3 Thermal motion of the atoms

2.3.1 Momentum-space methods: phonons

In the solid phase, atomic nuclei oscillate about an equilibrium position which is unchanged for a long time in comparison with the period of the oscillations. The relatively undistorted lattice found in crystalline solids allows displacements of the nuclei to propagate for large distances (compared with the interatomic spacing). The thermodynamic properties can then be described meaningfully in terms of the waves of displacement, rather than the displacement of an individual atom. The quanta of these waves are the phonon modes. [63, 40]

The phonon approach is preferred because it produces accurate results at low temperatures where the zero-point vibrational energy is important, and it allows comparison with experimental measurements of the phonon dispersion relation and density of states.

If nuclei interact through an effective harmonic potential between nearest neighbours, then the underlying group structure is Abelian, and phonons do not interact. Such phonons are termed quasiharmonic because although their interactions are ignored, the thermal expansivity of the lattice may be non-zero, because of the variation of the density of phonon states with compression. Of course, the 'harmonic' potential generally varies with density.

Quasiharmonic phonons

Consider a set of atoms $\{i\}$, mass m_i at positions $\vec{r}_i$. The total potential energy of the system Φ is a function of the $\vec{r}_i$. The position of each atom can be decomposed into an equilibrium value $\vec{r}_{0i}$, constant for each polymorphic phase, and the deviation $\vec{u}_i$ from equilibrium.

The equation of motion for each atom is

$$m_i \ddot{\vec{u}}_i = -\frac{\partial \Phi}{\partial \vec{u}_i}. \tag{2.106}$$

Φ can be expressed as a Taylor series in terms of the atomic displacements $\vec{u}_i$:

$$\Phi = \Phi_0 + \sum_i \frac{\partial \Phi}{\partial \vec{u}_i} \vec{u}_i + \sum_{i,j} \frac{\partial^2 \Phi}{\partial \vec{u}_i \partial \vec{u}_j} \vec{u}_i \vec{u}_j + \ldots \tag{2.107}$$

where Φ_0 is the frozen-ion cold curve energy. In equilibrium, $\frac{\partial\Phi(\vec{u}_i,\vec{u}_j)}{\partial\vec{u}_j} = 0$. Making the quasiharmonic approximation, terms cubic and higher in the $\vec{u}_i$ are neglected, so for small displacements from equilibrium,

$$\frac{\partial\Phi}{\partial\vec{u}_i} \simeq \sum_j \frac{\partial^2\Phi}{\partial\vec{u}_i\partial\vec{u}_j}.\vec{u}_j \equiv \sum_{j\neq i}\mathbf{D}(\vec{r}_i - \vec{r}_j)\vec{u}_j. \tag{2.108}$$

If the atomic structure can be represented by N atoms, then $\mathbf{D}$ has $9N^2$ components, covering all possible pairs of atoms and coordinate directions. $\mathbf{D}$ can be thought of of an $N \times N$ matrix of 3×3 matrices. In this work, $\mathbf{D}$ is referred to as the 'stiffness matrix' of the lattice. This distinguishes it from the dynamical matrix for a given phonon mode, defined below. Components of $\mathbf{D}$ are referred to using Roman letters for the index of each atom (i and j) and Greek indices for the coordinate directions (α and β).

Normal modes of the atomic motion are found by looking for wave-like solutions

$$\vec{u}_i(t) \propto \vec{\epsilon}e^{i(\vec{k}.\vec{r}_i-\omega t)} \tag{2.109}$$

where $\vec{\epsilon}$ is the polarisation vector, $\vec{k}$ the wavevector and ω the angular frequency of the mode with $\vec{\epsilon}$ and $\vec{k}$. Normal modes are thus the eigensolutions of

$$\omega^2\vec{u}_i = \sum_j \frac{\partial^2\Phi}{\partial\vec{u}_i\partial\vec{u}_j}.e^{i\vec{k}.(\vec{r}_i-\vec{r}_j)}\frac{\vec{u}_i}{\sqrt{m_i m_j}}. \tag{2.110}$$

The matrix on the right-hand side is the dynamical matrix, $\tilde{\mathbf{D}}$:

$$[\tilde{\mathbf{D}}(\vec{\mathbf{k}})]_{ij\alpha\beta} \equiv \sum_j \frac{\partial^2\Phi}{\partial[\vec{u}_i]_\alpha\partial[\vec{u}_j]_\beta}.e^{i\vec{k}.(\vec{r}_i-\vec{r}_j)}\frac{1}{\sqrt{m_i m_j}}, \tag{2.111}$$

where square brackets are used to denote an element of a matrix.

Phonons can be regarded simply as normal modes populated by Bose-Einstein statistics [63, 40, 2].

The eigenproblem can be solved using wholly real operations by constructing the symmetric matrix $\tilde{\mathbf{D}}'$ of rank double that of the complex original $\tilde{\mathbf{D}}$ [51]:

$$\tilde{\mathbf{D}}' \equiv \begin{pmatrix} \Re\tilde{\mathbf{D}} & -\Im\tilde{\mathbf{D}} \\ \Im\tilde{\mathbf{D}} & \Re\tilde{\mathbf{D}} \end{pmatrix}. \tag{2.112}$$

Symmetries of a general lattice

If Φ is a smooth potential, its second derivatives are independent of the order of differentiation. Thus

$$\frac{\partial^2\Phi}{\partial[\vec{u}_i]_\alpha\partial[\vec{u}_j]_\beta} = \frac{\partial^2\Phi}{\partial[\vec{u}_j]_\beta\partial[\vec{u}_i]_\alpha}, \tag{2.113}$$

or

$$[\mathbf{D}]_{i\alpha,j\beta} = [\mathbf{D}]_{j\beta,i\alpha}. \tag{2.114}$$

Phonons represent internal motions of a lattice about its centre of mass. Parallel motion of all the atoms – bulk motion of the crystal – is included as three zero-frequency modes per atom at $\vec{k} = 0$. If the stiffness matrix is not calculated perfectly, it may include contributions corresponding to bulk motion. Unfortunately, the stiffness matrix tends to be quite sensitive to inaccuracies and inconsistencies in calculating its elements [64]. More reliable results may be obtained by applying this constraint, which is simply that phonons exert no net force on the lattice, directly to the stiffness matrix:

$$\sum_j [\mathbf{D}]_{i\alpha,j\beta} = 0 \quad \forall i, \alpha, \beta. \tag{2.115}$$

If the point group has inversion symmetry, $\mathbf{D}(\vec{r}_i - \vec{r}_j) = \mathbf{D}(\vec{r}_j - \vec{r}_i)$. If U is a smooth function of the $\vec{u}_i$, the order of differentiation does not matter. Thus

$$[\mathbf{D}]_{\alpha\beta ij} = [\mathbf{D}]_{\alpha\beta ji} = [\mathbf{D}]_{\beta\alpha ij} = [\mathbf{D}]_{\beta\alpha ji} \tag{2.116}$$

[40], allowing the symmetry condition to be applied simply by recalculating

$$[\mathbf{D}]_{i\alpha,i\beta} = \sum_{j\neq i} [\mathbf{D}]_{i\alpha,j\beta} \quad \forall \alpha, \beta. \tag{2.117}$$

If the point group does not have inversion symmetry, the symmetrisation of $\mathbf{D}$ cannot proceed in such a straightforward way. Iterative schemes have been developed so that the commutivity of partial differentiation can be enforced and bulk motions removed from $\mathbf{D}$ [64]. In the scheme chosen, each iteration consists of an application of the following assignments in turn:

$$
\begin{aligned}
[\mathbf{D}]_{i\alpha,j\beta} &\leftarrow \frac{1}{2}\left([\mathbf{D}]_{i\alpha,j\beta}+[\mathbf{D}]_{j\beta,i\alpha}\right) && (2.118)\\
[\mathbf{D}]_{i\alpha,i\beta} &\leftarrow -\sum_{i\neq j}[\mathbf{D}]_{i\alpha,j\beta} && (2.119)\\
[\mathbf{D}]_{i\alpha,j\beta} &\leftarrow \frac{1}{2}\left([\mathbf{D}]_{i\alpha,j\beta}+[\mathbf{D}]_{j\beta,i\alpha}\right) && (2.120)\\
\sigma_1 &\leftarrow \sum_{i\neq j}[\mathbf{D}]_{j\beta,i\alpha} && (2.121)\\
\sigma_2 &\leftarrow \sum_{i\neq j}[\mathbf{D}]_{j\alpha,i\beta} && (2.122)\\
\sigma_d &\leftarrow \sigma_1-\sigma_2 && (2.123)\\
\text{if } |\sigma_d|>\epsilon \text{ then for } j\neq i: & && (2.124)\\
\sigma &\leftarrow \frac{\sigma_d}{2(N-1)} && (2.125)\\
[\mathbf{D}]_{j\beta,i\alpha} &\leftarrow [\mathbf{D}]_{j\beta,i\alpha}-\sigma && (2.126)\\
[\mathbf{D}]_{j\alpha,i\beta} &\leftarrow [\mathbf{D}]_{j\alpha,i\beta}+\sigma && (2.127)
\end{aligned}
$$

where ϵ is a numerical cutoff, and the '$\leftarrow$' symbol is used in the computational 'let $x = y$' sense. This algorithm works by assessing the amount by which the relation

$$\sum_{i\neq j}[\mathbf{D}]_{i\alpha,j\beta}=\sum_{i\neq j}[\mathbf{D}]_{i\beta,j\alpha} \qquad (2.128)$$

– a consequence of the other symmetry conditions – is violated because of errors in calculating $\mathbf{D}$, and adjusting the non-zero elements by the same amount so as to satisfy the relation [64]. In practice, a single application of this algorithm was sufficient for the symmetry equations to be satisfied to about 1 part in 10^{-7}, in amu-Å-eV units.

Calculation of elements in the dynamical matrix

In principle, elements in the dynamical matrix could be calculated explicitly by perturbing pairs of atoms from their equilibrium position and calculating the ground state energy Φ in each new configuration:

$$\mathbf{D}(\vec{r}_i-\vec{r}_j)\equiv\frac{\partial^2\Phi}{\partial\vec{u}_i\partial\vec{u}_j}\simeq\frac{[\Phi(\vec{u}_i,\vec{u}_j)-\Phi(\vec{u}_i,0)]-[\Phi(0,\vec{u}_j)-\Phi(0,0)]}{|\vec{u}_i||\vec{u}_j|}, \qquad (2.129)$$

for sufficiently small displacements $\vec{u}_i$ and $\vec{u}_j$. Unless the ground state can be obtained quite quickly, this is likely to be an extremely demanding procedure.

If the interaction between the atoms can be represented by an interatomic potential (IAP), then the forces and stiffness can be obtained by algebraic differentiation. This procedure is very efficient for most common forms of IAP, but is limited by the validity of using an IAP.

Matrix elements can also be obtained without the need to make pairs of displacements, by calculating the force on each atom if a single atom is perturbed from its equilibrium position. The forces can be calculated from the charge distribution of the electrons, and hence *ab initio* via the electron ground state. This method was preferred as it avoids the need to assume explicit forms for an interatomic potential. It explicitly includes the effect of displacing atoms on the electron band structure at $T = 0$. In principle, the effect of the mean atomic displacement on the electron-thermal energy could be taken into account via the change in band structure.

Consider a set of atoms in a calculation of the electron ground state. If a calculation is performed with atom i perturbed by some finite displacement $\vec{u}_i$ from its equilibrium position, it is possible to calculate the residual force on each atom (e.g. from the electron ground state calculation). For a sufficiently small value of $\vec{u}_i$,

$$\mathbf{D}(\vec{r}_i - \vec{r}_j) \equiv \frac{\partial^2 \Phi}{\partial \vec{u}_i \partial \vec{u}_j} \simeq \frac{\partial \Phi(\vec{u}_i)}{\partial \vec{u}_j} \frac{1}{|\vec{u}_i|}. \tag{2.130}$$

$\frac{\partial \Phi(\vec{u}_i)}{\partial \vec{u}_j}$ is simply the force $\vec{f}_j$ on atom j when atom i is displaced.

Because partial differentiation operators commute for a smooth function, a row and column of the eigenproblem can be determined from an electron ground state calculation with a displacement $\vec{u}_i$ along one of the coordinate directions, by dividing the force on each atom by the displacement. Other elements can be generated using symmetry, as described below. In simple terms, a force divided by the displacement which caused it is like the spring constant or stiffness of a linear spring. Thus one way of describing $\frac{\partial^2 \Phi}{\partial \vec{u}_i \partial \vec{u}_j}$ is to call it the 'stiffness matrix' of the lattice.

Spatial derivatives of the force on each atom can be obtained directly by including the necessary algebraic derivatives internally in the electron ground state code. The restoring forces are then calculated for infinitesimal displacements of pairs of atoms from their equilibrium positions. This approach has been demonstrated for some ground state codes [66], but has not yet been implemented in CASTEP. It is an example of adjoint differentiation of a forward model [67]; this typically introduces an order of magnitude increase in complexity in a computer program.

In the present work, electron states were calculated for a lattice cell with periodic boundary conditions. This meant that the forces were generated not only by the perturbation of the atom in the lattice cell but also from the images of the perturbed atom in the periodic boundary scheme. The effect of the images was reduced by performing the calculations with a supercell consisting of several of the lattice cells used to determine the cold curve. Typically, 2 × 2 × 2 lattice cells were used. In principle, the forces caused by the perturbation of a single atom could be found by repeating these calculations with increasing numbers of supercells, $N \times N \times N$ say, and estimating the $\vec{f}_j$ as $N \to \infty$. However, each calculation with a perturbed atom required far more computational time than the corresponding calculation for a perfect lattice, because the perturbation breaks the symmetry of the lattice. Another way of removing the effect of the perturbed images is to assume that the force is a function of distance from the perturbing atoms, $\vec{f}(\vec{r} - \vec{r}_i)$, such as radial dependence $|\vec{f}|(|\vec{r} - \vec{r}_i|)$. The force obtained from the electron states is taken to be a superposition of contributions from the perturbed atom and its images k:

$$\vec{f}_j = \sum_k \vec{f}(\vec{r}_j - \vec{r}_{i,k}). \qquad (2.131)$$

Given a functional form for $\vec{f}(\vec{r})$, the effect of the images can be removed. The form can be estimated by comparing forces for different sizes of supercell, and/or looking at systematic variations in for forces between different atoms in a single calculation. A simpler approach [65] is to set the $\vec{f}_j$ to zero if atom j is further from the perturbed atom than from any of its images, and to divide the force by the number of nearest images at an equal distance, to correct their effects on atoms lying midway between perturbed atoms (Fig. 2.5). A computer program was written to automate this process.

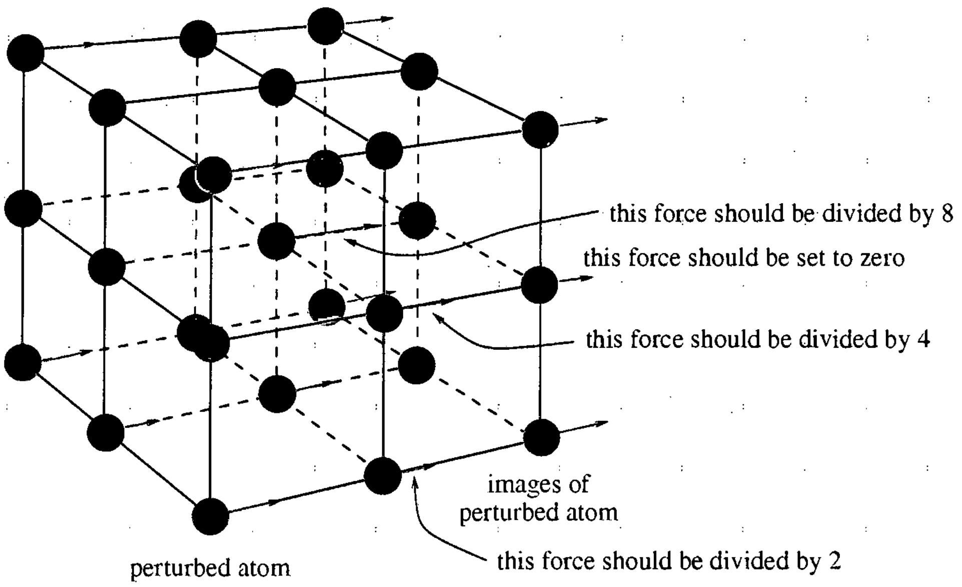
this force should be divided by 8
this force should be set to zero
this force should be divided by 4
images of
perturbed atom
perturbed atom
this force should be divided by 2

Figure 2.5. Simple scheme for correcting for the effect of perturbations in the position of image atoms.
(Illustration for 2 × 2 × 2 array of simple cubic cells.)

Symmetry of a specific lattice

Each of the perturbed-atom ground state calculations is in principle as straightforward as the corresponding calculation of the perfect (unperturbed) structure on the cold curve. However, the perturbation reduces the symmetry of the structure. For a given number of points in reciprocal space, more are unique. These calculations therefore take significantly more computer time. Fortunately, the symmetry of the atomic structure allows many of the elements of the eigenproblem matrix to be filled in from a single calculation. For example, in a structure with cubic symmetry, a displacement along the $(1,0,0)$ direction is equivalent to displacements along the $(0,1,0)$ and $(0,0,1)$ directions. The only difference is the orientation of the forces (a simple matrix multiplication) and the indices j of the atoms to which each force relates.

Each symmetry operation on the lattice can be defined as a rotation (matrix) and a translation (vector). This primitive description in turn defines a permutation among the indices j of the atoms in a structure with periodic boundary conditions. In other words, under the symmetry operation, each atom j maps to some atom j' in the repeating lattice cell. When using symmetry operations to propagate the results from a number of calculations of atomic forces to other elements of the eigenproblem matrix, it is necessary to consider explicitly only the rotation matrices and the permutation of atoms.

Density of phonon states

To predict the density of phonon states $g(\omega)$ for a structure with a given set of lattice parameters, the phonon eigenproblem was solved for each of a set of wavevectors $\vec{k}$. These were chosen randomly with a uniform distribution over the Brillouin zone, by taking the components to be three independent random numbers with uniform distribution between 0 and 1. Because these points in reciprocal space were random, their number is not reduced by symmetry operations, i.e. all were unique. This increases the statistical accuracy of the density of states. By contrast, if $\vec{k}$ points are chosen to be distributed evenly in a rectangular array, the symmetry of the lattice means that many of the points may be related to each other by symmetry operations, and hence are not unique.

The rate of convergence of the density of states for different distributions of $\vec{k}$ points was investigated by performing calculations with $10 \times 10 \times 10$ and $30 \times 30 \times 30$ points distributed evenly, and with 1000 and 10000 points distributed randomly. The density of states with $10 \times 10 \times 10$ evenly-spaced points was very spiky; much more so than with 1000 randomly-distributed points. Even at the level of $30 \times 30 \times 30$ evenly-spaced points, the density of states from 10000 randomly-distributed points was significantly smoother. (Fig. 2.6.)

Although large numbers of $\vec{k}$ points were needed to obtain a smooth density of states, the number needed for a converged equation of state was much smaller. This number was estimated by calculating the atom-thermal contribution through the density of states with different numbers of $\vec{k}$ points, and predicting the shock Hugoniot for each equation of state. It was found that ~ 1000 $\vec{k}$ points, collected into ~ 100 bins, was adequate for a converged shock Hugoniot.

Populating the phonon modes

The variation of lattice-thermal energy with temperature was found by populating the phonon modes according to Boltzmann statistics [40]. At a given density, the

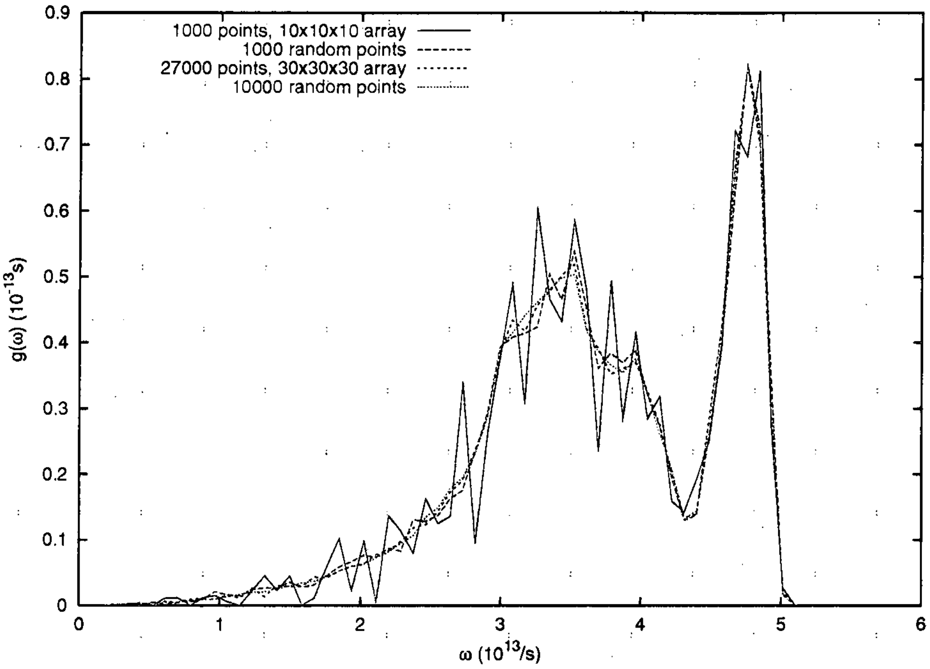

Figure 2.6. Statistical noise in density of phonon states for different numbers and distributions of $\vec{k}$ points.
(Illustration for aluminium with an interatomic potential of the Finnis – Sinclair form.)

lattice thermal energy is

$$E_l(T) = \frac{1}{2}\sum_i g(\omega_i)\hbar\omega_i + \sum_i g(\omega_i)\frac{\hbar\omega_i}{e^{\hbar\omega_i/kT} - 1}, \qquad (2.132)$$

from which $e_l(v, T)$ can be found by normalising to 3 modes per atom.

Negative eigenvalues of the dynamical matrix were sometimes found, giving imaginary phonon frequencies. Imaginary frequencies indicate an instability in the lattice: the force pushes the atom away from its equilibrium position, which is a maximum rather than a minimum of energy. This situation may be caused by the occurrence of a major phase transition (Fig. 2.7), or a minor perturbation of the structure (Fig. 2.8) – infinitesimal displacements will not distinguish between these possibilities, though finite displacements may be used to investigate the structure of the potential field experienced by each atom. In any case, the local potential is no longer quadratic, and the quasiharmonic model may not be accurate.

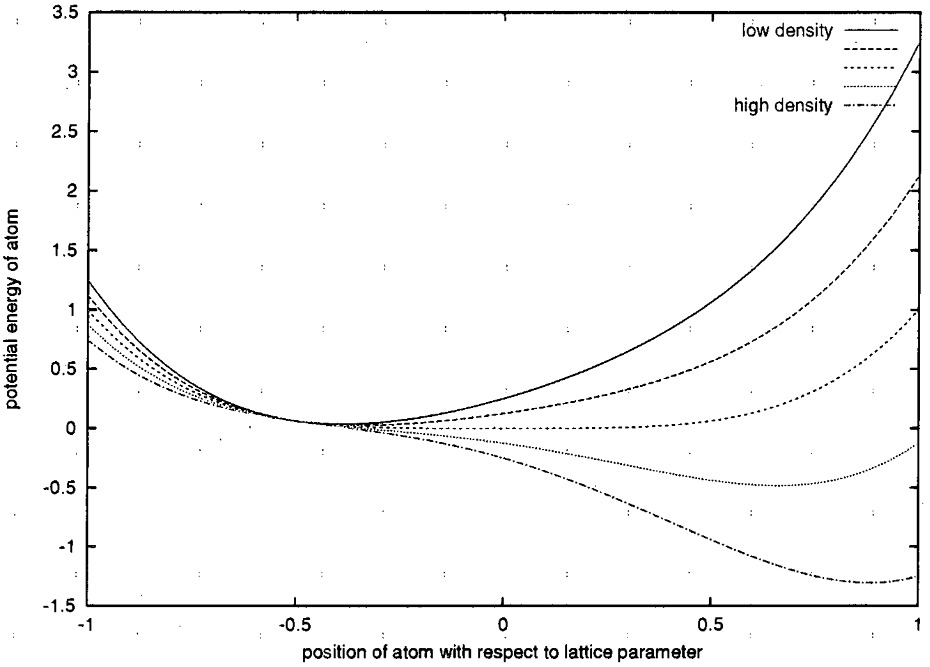

Figure 2.7. Potential function allowing a major phase change.

In the present work, we wish to generate an EOS for each polymorphic phase over a wide range of mass density and temperature. This is partly to allow phase boundaries to be defined more accurately, and partly so that metastable phases can be included (with a rate law for transitions to the stable state, as discussed earlier). We require a prescription for treating phonon states with imaginary frequencies, i.e. to allow the internal energy to be predicted when the structure is unstable. The following methods were considered:

1. Ignore states with imaginary frequencies, either leaving fewer than 3 modes per atom or re-scaling the density of states to give 3 modes per atom. These approaches are justified when a minor perturbation occurs to the structure.
2. Treat the states as freely translational, contributing a heat capacity of $k_B/2$ at all temperatures.

Equilibrium EOS and phase boundaries were found to be insensitive to the method chosen.

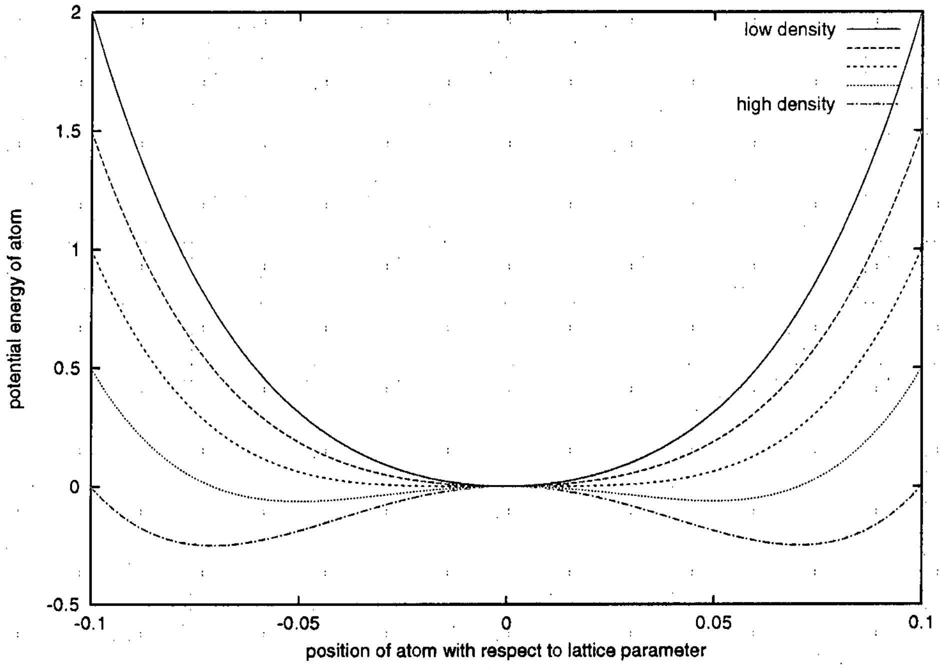

Figure 2.8. Potential function allowing a minor perturbation to the structure.

Anharmonicity; effect of different displacements

If the interatomic potentials were harmonic, then the restoring force on each atom would be directly proportional to its displacement from equilibrium. In general, there is no reason why interatomic potentials should be harmonic: the restoring force could increase more or less rapidly than linearly with displacement.

When calculating the stiffness matrix, atoms are displaced by a finite amount. If calculations are made for a series of displacements, a curve can be drawn allowing the stiffness to be inferred for infinitesimal displacements.

In reality, the zero-point energy and any thermal vibration means that the mean square displacement of the atoms is finite, so calculations for a finite temperature could use some 'effective stiffness matrix' for displacements of the expected magnitude. This is one way of incorporating anharmonic effects, and is an instance where using linear response theory to predict the stiffness matrix from infinitesimal displacements does not help directly.

'Typical' displacements

Given the stiffness, it is possible to estimate the mean displacement of an atom from equilibrium as a function of temperature. For the purposes of this estimate, all the stiffnesses calculated at any value of the lattice parameter are equivalent, as they are quite similar.

For a harmonic potential, the force f is proportional to the displacement d:

$$f = -\alpha d \tag{2.133}$$

where α is the stiffness. The potential energy e is given by

$$e = e_0 + \frac{1}{2}\alpha d^2 \tag{2.134}$$

where e_0 is the potential energy at the equilibrium position. The classical heat capacity of a single potential mode such as this is $k_B/2$, so the mean potential energy at a temperature T is $k_B T/2$. A mean displacement as a function of temperature is therefore

$$d \sim \sqrt{\frac{k_B T}{\alpha}}. \tag{2.135}$$

The magnitude of displacements may also be of use in relating the breadth of a diffracted X-ray beam to the material temperature.

Interpolating densities of states

It is computationally expensive to find atomic restoring forces from an electron ground state calculation. It would be preferable to avoid performing these calculations for every value of lattice parameter used in the cold curve. With this in mind, schemes were considered to enable densities of states to be interpolated between a smaller number of rigorous force calculations.

In principle, the interpolation can be performed at any stage, from the force components through the stiffness matrix to the phonon frequencies and the density of states itself. Processing the forces to give a complete stiffness matrix, and then solving the eigenproblem for each of a large number of wavevectors requires a significant amount of computer time (though far less than the electron ground state calculations). There are some practical problems in trying to interpolate between densities of states, since they tend to vary by a scaling process (which makes it awkward to manipulate bins between different lattice spacings) and by blocks of modes transferring between regions of the curve, while preserving the total number of modes. Accordingly, it was decided to approach the problem by interpolating the values of the frequencies themselves between calculations at adjacent values of the lattice parameter. These interpolated phonon modes were then collected into bins to give $g(\omega)$ and hence $e_l(T)$ in exactly the same way as the rigorously-calculated modes.

This technique suffers from potential pitfalls. Numerical schemes used to calculate the eigenvalues which correspond to individual modes do not generally find the eigenvalues in order, but locate each one iteratively. Given a guess at an eigenvalue, the nearest value to the guess is usually found first. The set of eigenvalues may be found in a random order, so the correspondence between the modes at adjacent values of the lattice parameter may be lost ('mode incoherence').

The interpolation of frequencies may be carried out in a number of different ways. Even choosing linear interpolation, this could be made linear in lattice parameter, its reciprocal, other powers (corresponding to density, specific volume, etc) or other functions. Various choices of interpolating function were tried and evaluated, as discussed later.

Much of the detail of these messy numerical considerations is lost in the integrations required to generate the EOS.

2.3.2 Atomic motion

An alternative way of estimating atom-thermal energies is by direct simulation of a large ensemble of atoms. If the motion of atoms from their equilibrium positions is treated explicitly then co-operative non-linear effects should be treated properly. On the other hand, the real-space treatment of atoms is less amenable to the inclusion of quantum mechanical effects such as the zero point energy. (This can be estimated in a semi-classical way.)

Direct simulations fall into two main classes: molecular dynamics and Monte-Carlo. Their use dates back to the 1950s or before; the algorithms are summarised in Appendix D.

In both cases, an ensemble of atoms is defined in the desired structure. The ensemble should be large in order to allow long-range co-operative effects to be treated. The accuracy of simulations of a given size is readily tested by performing occasional simulations of a different size – ideally larger – to determine to what extent the result has converged with system size. Choosing a temperature, the motion of the atoms is simulated until the energy has settled to a steady mean value. (The instantaneous value will generally fluctuate around a steady average.) Once the system has equilibrated, observables can be calculated by time averaging. [68]

In the present work, the main observable used was the internal energy, since the prescription for calculating EOS produced a consistent pressure given the energy. In some cases, the potential energy was observed to decrease significantly from the value for a perfect lattice at $T = 0$. This indicates a phase transition. No *a priori* way was found to predict the orientation of a new phase relative to the initial phase – in any case, this is not possible in general. It is not always easy to identify a phase by inspection (looking at a picture of the atoms in 3D), particularly at elevated temperatures where the atoms may be displaced significantly from their equilibrium positions. Instead, attempts were made to identify new phases by performing 'numerical powder crystallography' on the ensemble. To this end, a method was devised for predicting powder diffraction patterns from the positions of an ensemble of atoms, as described in Appendix E.

The ensembles used were too large to be amenable to quantum mechanical calculations of the electron ground state, so interatomic potentials were used to govern the motion of the atoms. The potentials were fitted to ground state calculations of reference states – usually the cold curve for one or more polymorphic phases – and were thus fairly consistent with the rigorous treatment of the electrons.

2.4 Thermal excitation of the electrons

The electron-thermal contribution $e_e(v,T)$ can be calculated from the electron band structure of a substance. The band structure itself can be obtained by a variety of techniques, in particular by finding the eigenvalues of the Hamiltonian of the electrons in the presence of a configuration of atoms.

A simpler can also be obtained using just the Fermi energy and the density of levels around that energy, or the effective number of conduction electrons.

2.4.1 Electron band structure

Given a particle field $\Psi(\vec{r})$ in the presence of a background potential $V(\vec{r})$, quantum mechanics allows the set of stationary states $\{\psi_i(\vec{r})\}$ to be calculated. The stationary states are those with a constant energy E_i.

The most simple procedure is to calculate the set of energy levels $\{E_i\}$ which could be occupied by a single particle state. In many applications, in particular atomic physics as required in EOS work, multiple particles are present. Almost all relevant atomic states possess several (or many) electrons. If the particles did not interact other than by the Pauli exclusion principle, then the single particle states $\{\psi_i\}$ could simply be populated by as many particles as are actually present in the system, N say. The ground state would be found by occupying the available states with the individual particles in order of increasing energy E_i. The ground state energy would be $\sum_{i=0}^{N-1} E_i$ (folding up and down spin states into a single series of energies).

However, electrons do interact with each other through the electromagnetic field and the constraint of antisymmetry which must be imposed on the wavefunctions of individual electrons. As a result, the energies $\{E_i : 0 \leq i \leq N\}$ of the individual electrons in the ground state are different to the single particle values. Moreover, the energy levels available in excited states $\{E_i'\}$ depend on the excitation state. For example, if inner electrons in an atom are excited to higher energies, then all the energy levels may change slightly. The outer levels will decrease because the shielding effect of some of the inner electrons – those which have been excited – is removed. This effect is most important at high temperatures, when a significant proportion of the electrons are in excited states.

The approach proposed here assumes that a unique (i.e. temperature independent) set of energy levels can be found. At lower temperatures, the energy levels predicted using ground state occupations should be quite accurate. At high temperatures, it may be possible in some electron programs to recalculate the energy levels at an elevated temperature (i.e. with some population of excited states) and use these modified states for calculating the energy at similar temperatures.

Note that 'low' temperatures are those where the electron-thermal energy is unimportant. These may still be several tens of thousands of K.

The picture of particles occupying a set of states is equally valid for an isolated atom, atoms in condensed matter or indeed electrons 'in a box' – i.e. in the absence of any nuclei. However, density functional theory as originally developed [30] applied to the ground state only; extensions to finite temperatures have been developed quite recently. The standard LDA prescription is not strictly applicable to finite temperatures.

In CASTEP, the electron wavefunctions are represented at a finite set of positions in reciprocal space, $\{\vec{k}_i\}$. The 'raw' number of these wavevectors is reduced by symmetry operations, so as a result each has a weight w_i. At each wavevector, electrons can occupy discrete energy levels. CASTEP is designed to calculate the energy of each band at each wavevector, $\{E_{ij}\}$. Each band can contain up to 2 electrons. This band structure refers to the outer electrons only, i.e. those which are not incorporated in the pseudopotential.

2.4.2 Electron-thermal energy

Let us assume that a set of occupation-independent electron energy levels $\{E_i\}$ can be found for a substance with a given arrangement of nuclei (polymorphic phase) and density.

If these energy levels are sufficiently accurate and complete, they may in fact be used to calculate the energy on the cold curve e_c. This is simply the ground state energy of the electrons, where the lowest states fully occupied up to a maximum at the Fermi energy E_F, above which no states are occupied. In this case, the cold curve and electron-thermal energy would be combined:

$$e(v,T) = e_e(v,T) + e_l(v,T). \tag{2.136}$$

This is not necessary for EOS at 'low' temperatures, where the only energy states which matter are those close to E_F.

If the number of energy levels calculated is very large – as in 'real' systems – then it may be represented more efficiently as a distribution function $g(E)$ where the number of levels with energy between E and $E + \Delta E$ is $g(E)\Delta E$. The distribution may be obtained from a sufficiently large set of energy levels, by obtaining the frequency with which they fall in a series of energy bins as finely-spaced as is necessary to resolve the density of levels for any particular application. Summations over discrete energy states become integrals over E, weighted by $g(E)$.

The probability of occupation of a fermion state of energy E at a temperature T is [2]

$$f(E) = \frac{1}{e^{(E-\mu)/kT} + 1}, \tag{2.137}$$

where the chemical potential μ constrains the total number of electrons: when $T = 0$, f is 1 for states where $E < \mu$ and 0 for states where $E > \mu$. Thus at $T = 0$, μ is equal to the Fermi energy E_F.

For a system described by a density of energy levels $g(E)$, the expectation value of the internal energy is

$$\langle E \rangle = \int_{-\infty}^{\infty} E f(E) g(E)\, dE. \tag{2.138}$$

This relation allows the internal energy to be calculated as a function of temperature for the EOS.

Given the density of levels $g(E)$, the chemical potential μ can be found at any temperature T by the constraint on the total number of electrons N:

$$\int_{-\infty}^{\infty} f(E) g(E)\, dE = N. \tag{2.139}$$

The relation is not readily invertible to find the variation of μ with T, so μ must in general be found by a process such as iteration. At $T = 0$, μ is equal to E_F. As $T \to \infty$, $\mu \to -\infty$. For temperatures of a few eV or below, μ does not decrease very much below E_F, and it is possible to choose a constant lower bound below which μ does not fall over the whole EOS. The precise value of this bound does not matter because the population functions vary smoothly with μ. Since the values of μ must lie between E_F and the lower bound, the relation could be solved robustly by bisection. At any stage in the iterative solution, if the current value of μ is too great then the resulting number of electrons will be greater than the number desired. This can be seen trivially for $T = 0$, where if $\mu < E_F$, $\int_{-\infty}^{\infty} f(E)g(E)\,dE < N$.

At 'low' temperatures, μ remains close to E_F and its variation can probably be ignored. By contrast, $g(E)$ can vary very rapidly with E. This is the reason why the electronic specific heat capacity of many metals is significantly different to the value predicted by free electron theory [40]. The density of energy levels can also change significantly between different polymorphic phases.

Given an electron structure scheme which operates at finite temperatures, a set of energy levels $\{E_i'\}$ modified by excitations can be determined. This is equivalent to a modified density of energy levels $g(E, T)$. A relation of this form could be incorporated in the equations above to give an electron-thermal energy which should be valid to arbitrarily high temperatures.

2.4.3 Low temperature behaviour

If the $k_B T$ is much less than the Fermi energy E_F and the density of levels varies slowly with E in the region of E_F, then the electronic heat capacity c_{ve} at constant volume is approximated well by [40, 15]

$$c_e = \frac{\pi^2}{3} g(E_F) k_B^2 T. \tag{2.140}$$

The resulting heat capacity may be specific, volumetric or per atom, according to the definition of $g(E)$.

Thus, given E_F and $g(E)$, the electronic heat capacity can be calculated quite easily at temperatures which are low in comparison to the energy range over which $g(E)$ remains effectively constant.

2.4.4 Quantum free electron model

If $g(E_F)$ is not readily available, then it can be estimated for metals by using the Sommerfeld (quantum free electron) model [40, 69]. The Fermi surface overlaps the conduction bands, so the density of levels can be estimated by treating the conduction electrons as free, non-interacting particles in a box. In this case [69] the density of levels per unit volume is

$$g(E) = \frac{(2m_e)^{3/2}}{2\pi^2\hbar^3}\sqrt{E} \tag{2.141}$$

where m_e is the mass of an electron and $\hbar$ is Planck's constant h divided by 2π. By populating this distribution with electrons up to the Fermi energy, one finds that

$$g(E_F) = \frac{3}{2}\frac{n}{E_F} \tag{2.142}$$

where n is the density of conduction electrons. The electronic specific heat capacity is then

$$c_{ve} = \frac{\pi^2}{2}\frac{k_BT}{E_F}\frac{nk_B}{\rho}, \tag{2.143}$$

where ρ is the mass density. The form given here is convenient as it explicitly includes the Fermi energy E_F. Although the derivation is based on free electron theory, more rigorous calculations of E_F can be inserted instead to give an alternative estimate of the electron-thermal energy:

$$e_e(\rho, T) = \frac{\pi^2}{4}\frac{k_BT}{E_F(\rho)}\frac{n(\rho)k_BT}{\rho}. \tag{2.144}$$

(Arranging the equation in this way emphasises the factor k_BT/E_F which reduces the number of electrons taking part, and the factor nk_BT which is a measure of the thermal energy of n excited electrons.)

The number of conduction electrons per atom can be estimated in a variety of ways, such as taking the preferred valence Z_v for the metal. Then,

$$n = \frac{Z_v\rho}{m_a} \tag{2.145}$$

where m_a is the mass of an atom. However, the effective number of conduction electrons is also liable to increase with compression, as more electron shells overlap between adjacent atoms.

The 'pure' free electron model results in an inaccurate electronic heat capacity for many substances. Postulating that $g(E_F)$ is in error, it is possible to use the ratio between experimental and free electron heat capacities [40] to adjust the estimates. The ratio of the heat capacities can be applied directly to $e_e(T)$. It is however a dubious assumption that the same ratio is valid over a wide range of densities.

If E_F is not known, then the free electron value can be used [40, 15]:

$$E_F = \frac{\hbar^2 k_F^2}{2m_e}, \tag{2.146}$$

where $\hbar = h/2\pi$ and k_F is the Fermi wavevector,

$$k_F = (3\pi^2 n)^{1/3}. \tag{2.147}$$

Following this sequence of approximations,

$$E_F = \frac{\hbar^2 (3\pi^2 Z_v \rho / m_a)^{2/3}}{2m_e}, \tag{2.148}$$

so that $E_F \propto \rho^{2/3}$ with a constant of proportionality

$$\alpha \equiv \frac{\hbar^2 (3\pi^2 Z_v / m_a)^{2/3}}{2m_e}. \tag{2.149}$$

2.5 Correcting inadequacies in *ab initio* calculations

It was found that purely *ab initio* predictions of the EOS over-predicted the density at STP by a few percent. This is thought to be a consequence of the use of the local density approximation (LDA) to the electronic exchange and correlation, i.e. the effect of this simplification of the electron treatment. The predictions were not significantly better when the generalised gradient approximation (GGA) was used instead.

The density in static compression and shock wave experiments can be obtained to a considerably greater accuracy than this, and it is desirable that EOS used for engineering calculations be as accurate as is warranted by the uncertainty in the experimental data. For this reason, methods of correcting the *ab initio* EOS were investigated.

In principle, any amount of experimental data could be used in adjusting an EOS. Ideally, only minimal adjustments should be made to a model which is supposed to contain a high degree of physical fidelity, otherwise the adjustments are may make the model less self-consistent. In the present work, the only adjustments considered were a global pressure offset over the whole EOS surface: one or two parameters for which the starting values were defined by the *ab initio* model.

There are many possible ways to adjust an EOS so that it passes through a single reference state $p_r(\rho_r, T_r)$, e.g. STP or $p = 0$, $T = 0$. The following simple methods can be used to remove an inaccuracy in p at ρ_r (equivalent to in inaccuracy in ρ at p_r) if p increases with ρ in the vicinity of the reference state:

Pressure offset: Add a constant to the pressure over the whole EOS:

$$p(\rho, T) \to p'(\rho, T) = p(\rho, T) + \Delta p. \tag{2.150}$$

Density scaling: Scale the density:

$$\rho \to \rho' = \alpha\rho. \tag{2.151}$$

Some other possibilities can be discarded on physical grounds. A density offset $\rho \to \rho + \Delta\rho$ is unphysical in the case of an *ab initio* model designed (reasonably) to give $p = 0$, $e = 0$ as $\rho \to 0$. A pressure scaling $p \to \alpha p$ is undesirable because one reference state might be $p = 0$, where the scaling would have no effect.

The pressure offset and the density scaling were investigated by applying them to an early version of the *ab initio* EOS for aluminium, adjusting the EOS to pass through the density measured at STP. The ambient isotherm and principal Hugoniot were then generated from the modified EOS. The pressure offset was found to give significantly better agreement over the rest of the EOS than did the density scaling.

It is in any case philosophically desirable to adjust the pressure rather than the density. In the EOS-constructing method developed here, the density and temperature are independent parameters for each value of which the energy and pressure are calculated using models of the physics. The models are not perfect – the inadequacy of the LDA is a prime example, but not the only one – so it is more logical to accept that adjustments should be made to p and e rather than ρ or T.

The complete EOS is built by first generating $e(\rho, T)$ and then using it to calculate $p(\rho, T)$. If the pressure offset is applied to $p(\rho, T)$ only then the pressures will no longer be consistent with the energies. A more rigorous approach is to apply the correction to the internal energy, and then calculate the pressure using the modified internal energy. Given a desired pressure correction Δp, the specific internal energy should be modified by adding a linear function of specific volume $v = 1/\rho$:

$$e(v, T) \to e'(v, T) = e(v, T) + (\Delta e_0 - v\Delta p), i.e. \tag{2.152}$$

$$e(\rho, T) \to e'(\rho, T) = e(\rho, T) + (\Delta e_0 - \Delta p/\rho). \tag{2.153}$$

The energy offset Δe_0 is essentially a free parameter, and can be chosen so that the energy remains constant at some reference state.

An even better approach might be to adjust the exchange-correlation model (e.g. the LDA or GGA parameters) until the reference state is matched. This would allow the correction to influence the stiffness of the lattice with respect to perturbations in atom position, and would also predict a modified electron band structure.

The importance of corrections of this order was estimated for aluminium by generating the lattice-thermal energy via interatomic potentials (IAPs) rather than using quantum mechanical calculations of the lattice stiffness. IAPs of the Morse form were fitted to the FCC cold curve as described later, and then used to predict the density of phonon states. Applying the energy tilt applied to the frozen-ion cold curve, a modified IAP was calculated to fit this cold curve, and then used to predict modified phonon frequencies. The pressure offset produced a small change in phonon frequencies, but the sensitivity of the ambient isotherm and principal Hugoniot to the difference in lattice-thermal energy was very small.

2.6 Summary of method for predicting equations of state

Equations of state were generated by using the steps listed below. The names of significant computer programs are also given; electronic manuals were written for most new programs.

For each phase:

For each density (i.e. set of lattice parameters):

Find equilibrium state: Calculate electron ground state (using `CASTEP`), varying internal degrees of freedom until the state of isotropic stress is found. Energy at this state lies on the frozen-ion cold curve. Units are also converted from eV-amu-Å to Mg-MJ-GPa-mm-μs (using utility program `function`).

Electron-thermal contribution:

1. Calculate band structure (`CASTEP` in band structure mode).
2. Convert to density of levels (programs `BAND2MODE` and `MODE2DOL`).
3. Calculate electron-thermal energy at each temperature along isochore, by applying Fermi-Dirac statistics and determining the chemical potential (program `DOL2ET`).
4. Convert units from eV-amu-Å to Mg-MJ-GPa-mm-μs (utility program `function`).

Lattice-thermal contribution: [2]

1. For each displacement required by symmetry, calculate the force on each atom when the appropriate atom is displaced (`CASTEP`).
2. Modify the forces to take account of the finite size of the supercell (program `PERTFORCE`).
3. Combine the forces, using symmetry to fill the stiffness matrix (program `FORCE2PHI`).
4. Apply translational symmetry to the stiffness matrix (program `PHISYMMCEN` for a centrosymmetric structure, or `PHISYMM` in general).

---

[2] As mentioned in Section 2.3, the calculation of the stiffness matrix and the application of lattice symmetries were based on previous work by Stewart Clark, Michele Warren and Graeme Ackland at the University of Edinburgh [64, 65]. The programs used here were however written from scratch.

        5. Calculate the vibrational frequencies for each candidate phonon wavevector (program `PHI2PHON`).
        6. Collect frequencies into statistical bins, to produce the density of states (utility program `bin`).
        7. Adjust the density of states to treat imaginary modes as freely translational, if desired.
        8. Calculate atom-thermal energy at each temperature along isochore, by applying Maxwell-Boltzmann statistics (program `PHONS2ET`).
        9. Convert units from eV-amu-Å to Mg-MJ-GPa-mm-μs (utility program `function`).

    Collect contributions into tables:

        1. Cold curve points: set of v, e values (WXC++ 'table1d' format).
        2. Electron-thermal energy: rectangular array

           ```
           number_of_volumes number_of_temperatures
           For each volume:  specific volume
           For each temperature:  temperature
           For each volume:  For each temperature:  specific internal
           energy
           ```

           (WXC++ 'table2d' format).
        3. Atom-thermal energy: as for electron-thermal energy.

    Specific internal energy: combine contributions into a single table (utility program `combine`).

    Free energy: integrate along each isochore to calculate $f(T)$ (program `ENTROPY`) and combine isochore data into a rectangular $f(v, T)$ table.

Equilibrium surface: Using the free energy $f(v, T)$ for each phase calculated, predict the phase diagram (program `PHASEDIAG`) and equilibrium free energy $f(v, T)$ (program `EQMEOS`).

Equilibrium equation of state: Given the equilibrium free energy table $f(v, T)$, apply thermodynamics to deduce the specific internal energy $e(v, T)$ and pressure $p(v, T)$, both in tabular form (program `VTF2EOS`).

***Ab fere initio* equation of state:**

    Discrepancy: extract the pressure at the known STP density (programs `tabinterp` for bilinear interpolation or `tabinterp1` for biquadratic). The discrepancy is the difference between this pressure and 1 atm (i.e. 10^{-4} GPa).

Correction: formulate the energy correction and apply to the energy table for each phase (repeat with the correcting term in program `combine`, as above).

Regenerate equation of state: repeat calculations of free energy, equilibrium properties and EOS, as above.

Additional calculations were performed to investigate the sensitivity to resolution in reciprocal space (electrons or phonons), plane wave energy cutoff, atom displacement, resolution of binning in density of states, etc. UNIX shell scripts were written to carry out common sequences of calculations (such as the generation of the cold curve, or the thermal contributions along each isochore) automatically.

In some instances, calculations were performed at a single density with a variety of values for the plane wave cutoff energy, in order to determine the Pulay stress. In practice, most calculations were performed with a sufficiently large number of plane waves that the effect of the Pulay stress was negligible.

The procedure was modified slightly when investigating the use of interatomic potentials for the atom-thermal energy, or simpler electron-thermal models.

The `K290` program was used to deduce the symmetry of each lattice and hence find the unique set of points in reciprocal space. The symmetry operations were deduced for each lattice type, and also for each unique displacement of an atom from equilibrium required when calculating *ab initio* phonons. The symmetry operations used in propagating restoring forces around the elements of the dynamical matrix were those deduced with all atoms at their equilibrium positions.

Typically, the density range used was from a little below STP density (e.g. 90% of STP density) to about double STP density. For each phase, the density range was converted to a range of lattice parameters, rounded to 0.1 Å. Calculations were performed on isochores at intervals of 0.1 Å. EOS tables then included 15 to 20 isochores. (*Ab initio* phonons were typically calculated at larger intervals and interpolated onto other isochores.

The *ab initio* restoring force calculations were based on displacements of 0.001, 0.01, 0.0 and -0.001 of the corresponding lattice parameter for the supercell. The values of ±0.001 represent an 'infinitesimal' displacement; 0.01 represents a displacement of the same order as would be expected at temperatures ~1000 K, as predicted along the shock Hugoniot. The sensitivity of the EOS to displacement was found to be very small, as discussed in the context of sample materials in Chapter 3.

The temperature range used was 0 to 10000 K. A geometrical series of 200 intervals was used, with an expansion ratio of 1.01. The smallest interval was thus 15.8328 K (i.e. the first two isotherms were at 0 and 15.8328 K) and the largest was 114.69 K (i.e. the last two isotherms were at 9885.31 and 10000 K). The upper limit of 10000 K was chosen to allow estimates to be made of the shock Hugoniot well into the liquid regime. To a first approximation, the EOS in the liquid regime at temperatures above the solid-liquid transition is similar to the EOS of the solid at the same density [1]. Long-range order is lost in the liquid, but short-range order is expected to be similar to that seen in the solid, so the cold curve calculated for the nearest solid phase below melt should often be reasonable in the liquid. Quasiharmonic phonons are not an appropriate model, but it is likely that all vibrational modes are active anyway at or below the melt temperature, so the Dulong – Petit value to which the phonon treatment asymptotes should be reasonably accurate. Assuming that the electron band structure is not perturbed significantly by the changes in atomic configuration caused by melting at high pressure, the band structure itself should still be valid at such temperatures (10000 K is less than 1 eV).

The electron ground state calculations typically used an array of $10 \times 10 \times 10$ points in reciprocal space. Calculations with $5 \times 5 \times 5$ points were used to check the convergence. This limited the resolution in the band structure and density of electron energy levels. Because of the relatively small number of points, these were typically applied directly in calculating the electron-thermal energy; the density of levels being used for graphs only. In contrast, the number of (randomly-distributed) points in reciprocal space used to calculate phonon frequencies was varied until the density of states had converged. (The density of binning was varied until the density of states looked smooth while preserving important features.) Typically, 10^4 to 10^5 points were used, collected in bins 0.05 to 0.5 frequency units wide (in eV-amu-Å units).

Both ranges and intervals were chosen so that shock Hugoniots were converged to a reasonable accuracy, small in comparison with the range of the Hugoniot and reasonably small compared with the scatter in the experimental data. The resolution was tested by generating EOS with half the resolution. The temperatures were certainly fine enough for accurate EOS predictions; the densities were just fine enough for reasonable predictions. (It requires much more computational effort to add extra isochores than to add extra isotherms.)

The resulting EOS were in effect optimised for compressions into the strong shock regime. The method developed should also allow reasonable predictions to be made of, for example, the thermal expansivity near STP; however, the tabulations used in the present work were too coarse for such predictions to be accurate.

Chapter 3

Application to a selection of materials

3.1 Aluminium

Aluminium was chosen as the first material to analyse because it is a simple metal, forming the FCC structure to at least 200 GPa. A considerable amount of experimental data is available. The pseudopotential had been used with good results at normal density, and calculations using other techniques have been published. Furthermore, aluminium is a metal of technological significance, so an accurate and thermodynamically complete EOS would be valuable for a range of applications.

3.1.1 INFERNO isotherms

In order to evaluate the atom-in-jellium models used in INFERNO (summarised in Section 2.2), calculations were made using each model in turn.

INFERNO's option settings were essentially as recommended [48] (Table 3.1). Calculations were made with atomic sphere radii between 1.5 and 4 Bohrs, at intervals of 0.01 Bohr, at a temperature of 0.01 eV. No numerical difficulties were encountered in the case of aluminium. The calculation of each state required only a few iterations; generally under ten.

The pressure was found by numerical differentiation of the energy – volume dependence. States with pressures outside the range 0 and 400 GPa were discarded for the purposes of this study. Calculations were also performed with the mesh extending to a smaller radius. This made no difference to the results. Some calculations were repeated at a temperature of 0.03 eV to indicate whether the electron – ion energy would be likely to change much between 0.01 eV and 0 K.

The calculated state data were used to infer the density at which $p = 0$, for $T \simeq 116\,\mathrm{K}$ (0.01 eV), and compared with the reference density of $2.77\,\mathrm{Mg/m^3}$ [110] (Table 3.2). The results from calculations at 116 K and 348 K did not differ before the fourth significant figure.

By far the best match was from Model B, which differed from the data by less than 1%.

The INFERNO calculations were compared with measurements of the ambient isotherm to 220 GPa [71]. The difference between calculations at 116 and 348 K was not significant – everywhere less than 1 GPa. (Fig. 3.1).

Model B again produced the best results. Its compression curve lay about 4% below the reference curve at $v/v_0 \sim 0.7$, and about 10% below at $v/v_0 \sim 0.45$.

Table 3.1. Important INFERNO option settings.

Option	Setting
relativity	on
radial mesh	exponential: $r_i = 60e^{(i-421)/32}$
mixing parameter	0.3
effective infinite radius of bound state	75
convergence criterion	10^{-6}
mode	1 (find chemical potential given atomic volume)
local density model	Hedin-Lundqvist
parameter in local density exchange model	2/3 (Kohn-Sham)
Coulombic tail for last electron in atomic sphere	off
Lagrange multiplier in charge neutrality constraint	type 2
chemical potential of atom relative to jellium	0
energy meshpoints for R-matrix scheme in continuum	5
largest angular momentum quantum number in explicit solution	5

The results from the other models lay tens of percent below the reference curve.

A brief study found that no simple scaling or shift would bring the calculated curves into good agreement with the data.

The 'Model B' variant of the atom-in-jellium model appeared to produce the most accurate cold compression data for aluminium between 0 and 400 GPa. Even so, the pressures predicted may be in error by of the order of 10%.

INFERNO could however calculate a state point about two orders of magnitude faster than a band structure code. Although it cannot calculate states at 0 K, it seemed to operate reliably at temperatures well below 0.1 eV.

The electron – ion isotherm predicted at a temperature of 0.01 eV was the same as that predicted at 0.03 eV to at least three significant figures, suggesting that the 0 K isotherm should be represented to at least this accuracy by the 0.01 eV isotherm.

Table 3.2. Densities at $p = 0$.

Source	density (Mg/m^3)
Reference data [110]	2.77
Model A	2.49
Model B	2.75
Model T	2.51

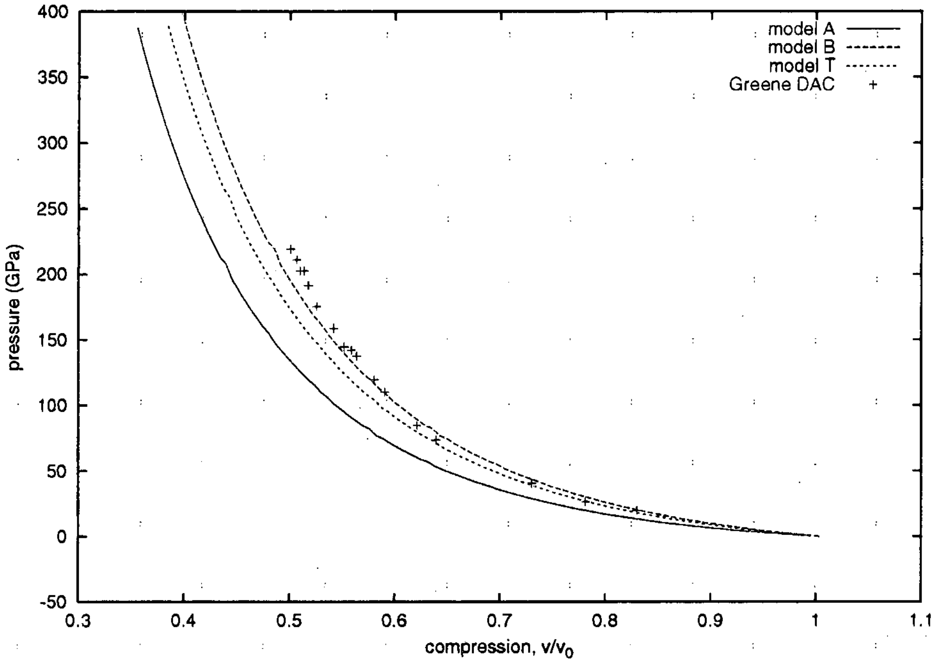

Figure 3.1. INFERNO cold curves for aluminium.

3.1.2 CASTEP cold curve

A reciprocal-space pseudopotential was used [56], incorporating all but the outer 3 electrons. It was derived by adjusting its parameters to reproduce the scattering properties for the outer electrons in an isolated atom calculated using all-electron methods [62], and is therefore wholly *ab initio*. The non-local part of this pseudopotential has contributions for $l = 0$ and $l = 1$.

CASTEP version 5.1 was used. This version incorporates the entropy correction to stabilise the convergence of metal calculations performed with Gaussian smearing. The entropy correction was found to make a negligible difference. Version 5.1 does not include non-linear core corrections. The program was modified to increase the number of k points allowed.

To find the ground-state of each configuration of interest, the program was allowed to perform 45 iterations towards the ground state from a random initial state. By then, the energy and stress changed very little between iterations.

Pulay stresses [59, 60, 20, 61] were taken into account. The total energy E_T of a 4-atom FCC lattice cell was calculated with an energy cutoff E_c of 250, 300 and 350 eV. $dE_T/d(\ln E_c)$, used by CASTEP in determining the Pulay stress, was estimated from the results. The correction to the ground-state energies was negligible.

The effect of the density of points used to represent the electron field in k-space was investigated by performing sets of calculations with 125 (5^3) and 1000 (10^3) points in the Brillouin zone. The number of k-points actually required in the calculations was considerably less than this, since the Brillouin zone could be reduced to a smaller unique wedge by symmetry. The calculations with 125 k-points produced pressures slightly below those with 1000 k-points, by less than 1 GPa. This suggests that the calculations with 1000 k-points were converged in this respect to better than 1 GPa. 1000 k-points were used as standard.

The absolute ground-state energy of an isolated atom was found by performing calculations on the simple cubic (sc) lattice with a lattice side a of 6, 8 and 10 Å. The ground-state energy was found to about ±0.02 eV by extrapolating to infinite cell size using inverse powers of a.

The effect of the gradient correction to the LDA [31, 32] was investigated by repeating calculations of all configurations using this modified model of the exchange-correlation energy. This is not strictly a proper treatment, because the pseudopotential used was determined with the LDA. It is possible that re-optimising the pseudopotential using the GGA would produce different results, thus the GGA calculations presented here should be regarded with caution.

Calculations were performed to predict the cold compression curves (energy e and pressure p as functions of density ρ or specific volume v) for aluminium in the FCC and BCC structures, using lattice cells containing 4 and 2 atoms respectively. Calculations were also performed of the ideal HCP structure, using lattice cells containing 2 or 4 atoms (HCP primitive cell or tetragonal cell respectively). The calculations predicted that FCC remains the stable phase at $T = 0$ over the range $1.5 \leq \rho \leq 3.0\,\mathrm{amu/Å}^3$ ($2.6 \leq \rho \leq 5.5\,\mathrm{Mg/m}^3$, giving pressures up to about 200 GPa). When the density exceeds $3\,\mathrm{amu/Å}^3$ ($5\,\mathrm{Mg/m}^3$) or so, the energy curves become too close to distinguish reliably with the calculations presented here. (Fig. 3.2.)

The difference between LDA and GGA calculations was an offset of 0.01 eV/amu, GGA having the higher energy.

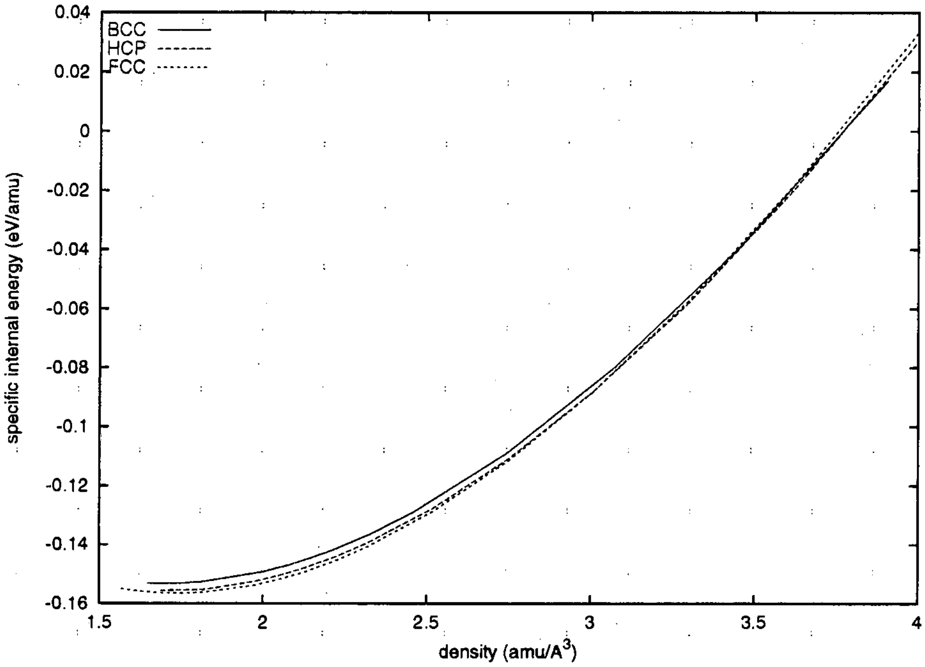

Figure 3.2. Predicted ground state energy for aluminium.

The ground state calculations also produced values of the Fermi energy (Fig. 3.3) and the electronic band structure. These were used in predicting the electron-thermal contribution to the EOS. In CASTEP, the Fermi energy is calculated with respect to the band energies; these contain an absolute offset which depends on the pseudopotential. The predicted values for Fermi energy could be represented accurately by a free electron form (see Section 2.4) incorporating an offset,

$$E_F = E_0 + \alpha\rho^{2/3}. \tag{3.1}$$

E_0 and α were obtained by fitting a straight line to $(\rho^{2/3}, E_F)$ pairs. E_0 was found to be $-7.998\,\mathrm{eV}$ and α $8.678\,\mathrm{eV.m^2/Mg^{2/3}}$, with a residue (sum of squared deviations in E_F) of $\sim$0.017.

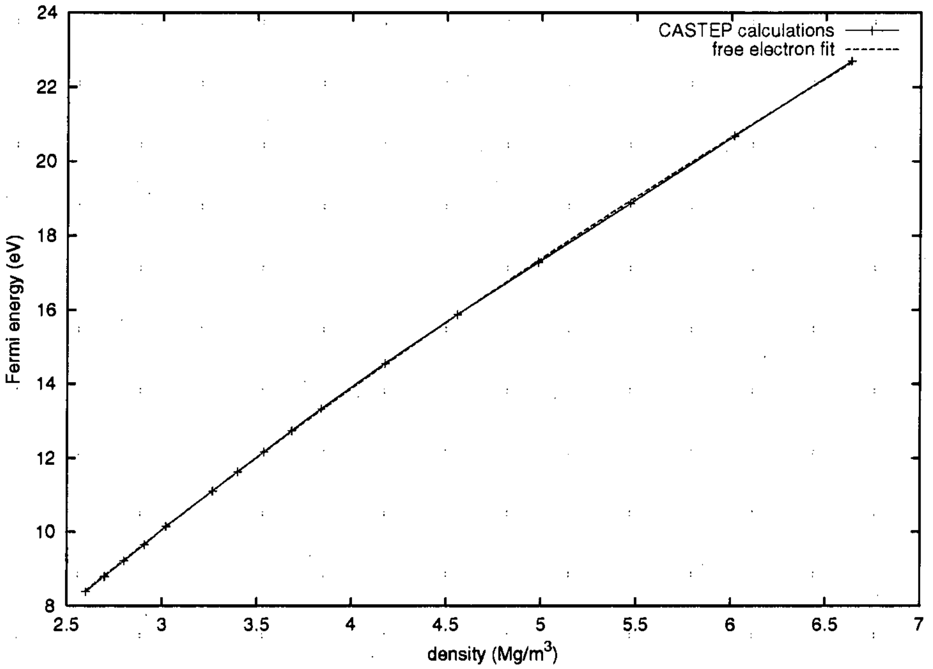

Figure 3.3. Predicted Fermi energy for aluminium, compared with variation for free electrons.

The value deduced for α would correspond to 5.2 conduction electrons per atom (assuming the effective mass is the rest mass of a free electron). The valence of aluminium is 3; the difference can be regarded as caused by the concentration of the electron wavefunctions into a smaller volume of space by the ions or as a smaller effective mass. The predicted variation of Fermi energy with density cannot be reproduced at all accurately with 3 conduction electrons. This discrepancy demonstrates the need to treat electrons accurately when predicting EOS – it is not adequate to assume that electronic effects can be modelled by treating the valence electrons as free, as has been done previously [16].

Grüneisen parameter

For comparison with empirical EOS, the Grüneisen parameter $\gamma(v)$ was deduced from the FCC cold curve using the Dugdale-MacDonald relation [16]. The temperature dependence of γ was not considered.

Analytic functions were fitted to sets of adjacent states calculated on the cold $p(v)$ curve. These functions were differentiated to obtain $\gamma(v)$. The results were compared with different functional fits to different numbers of adjacent points. When the number of points used in a functional fit was greater than the number of fitting parameters, the fit was least squares; otherwise, it passed precisely through each point. It was found that quadratics fitted to sets of 3, 4 and 5 points, and cubics fitted to sets of 4 and 5 points produced $\gamma(v)$ with similar behaviour (Fig. 3.4), whereas exponentials ($p = a\exp(bv)$) fitted to 2, 3, 4 and 5 points were quite different. Such exponentials cannot model the region near $p = 0$, but it seems that their behaviour is inadequate even at finite pressures. (Fig. 3.4)

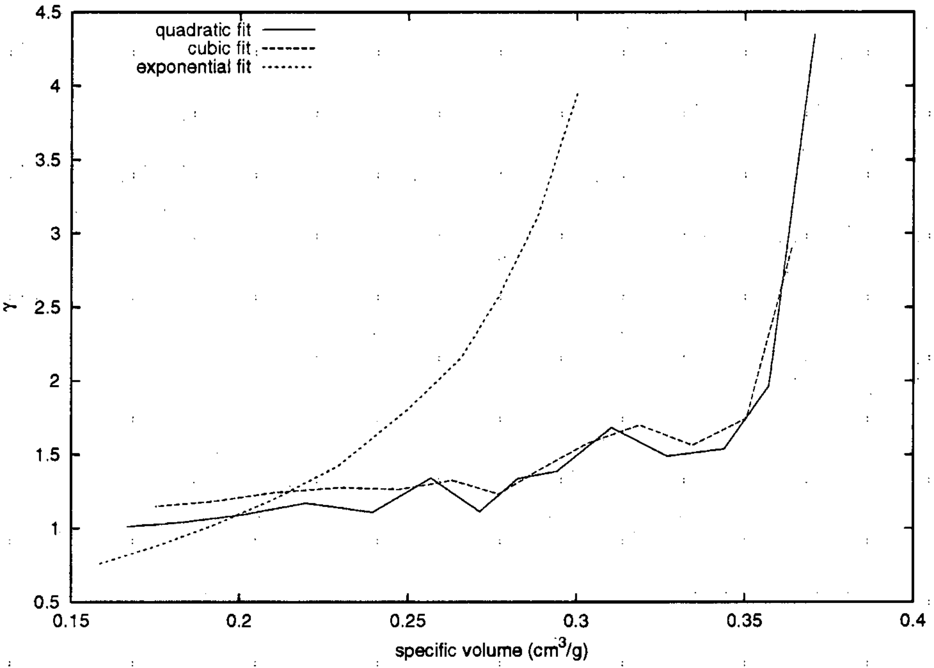

Figure 3.4. Grüneisen $\gamma(v)$ deduced for aluminium.

The behaviour predicted of $\gamma(v)$ indicated a linear region for $v < v_0$, then a steep increase. For comparison with empirical EOS optimised for compression experiments, the steep increase was ignored and a linear variation obtained from the (v, γ) points deduced from polynomial fitting:

$$\gamma(v) = 0.358 + 3.748v. \tag{3.2}$$

It is instructive to compare the deduced $\gamma(v)$ behaviour with results from other workers [19, 70]. The variation derived from polynomial fits was reasonably consistent with the accepted behaviour of γ. $\gamma(v)$ can also be deduced from other empirical EOS forms using the Dugdale – Macdonald relationship (see Appendix G). Here, published parameters for Al were used, obtained by fitting to experimental data, rather than by optimising the parameters to reproduce the CASTEP calculations. It can be seen that with the exception of the piecewise exponential fit to the *ab initio* calculations, all the other calculations of $\gamma(v)$ are in reasonable agreement over most of the compression range. There is significantly more variation around equilibrium density; this is also reflected in the wider range of values obtained by experiment [19]. Around $p = 0$, the uncertainties in deriving γ are significantly greater. The calculation of $\gamma(v)$ involves the second derivative of the pressure with respect to density. Differentiating tends to highlight differences between functional forms. For example, the switch between forms used in the EOS by Bushman et al is clearly evident as a sudden drop in γ. (Fig. 3.5)

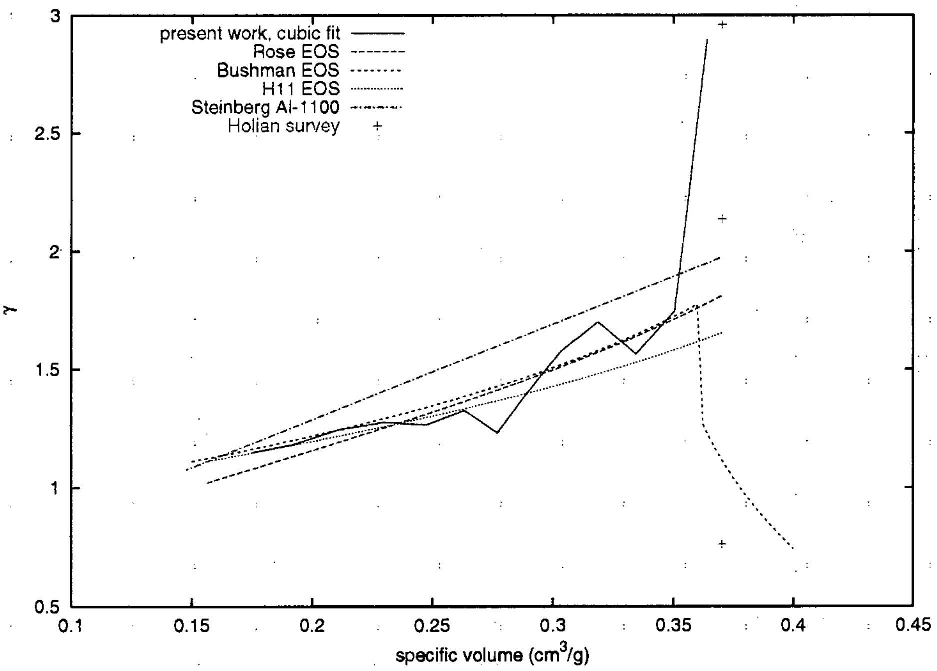

Figure 3.5. Grüneisen $\gamma(v)$ deduced from different models for aluminium.

3.1.3 Lattice-thermal contribution

The lattice-thermal energy $e_l(v,T)$ was predicted from the restoring forces predicted *ab initio* by displacing each atom by a finite amount, and also using interatomic potentials deduced from the cold curve.

In the FCC lattice, the high degree of symmetry allows all elements to be obtained by displacing the atom at $0,0,0$ in the x-direction. The FCC lattice has inversion symmetry, so the dynamical matrix could be symmetrised in a deterministic way (Section 2.3).

The density of phonon states at normal density was compared with experimental data deduced from neutron scattering [40] (Fig. 3.6).

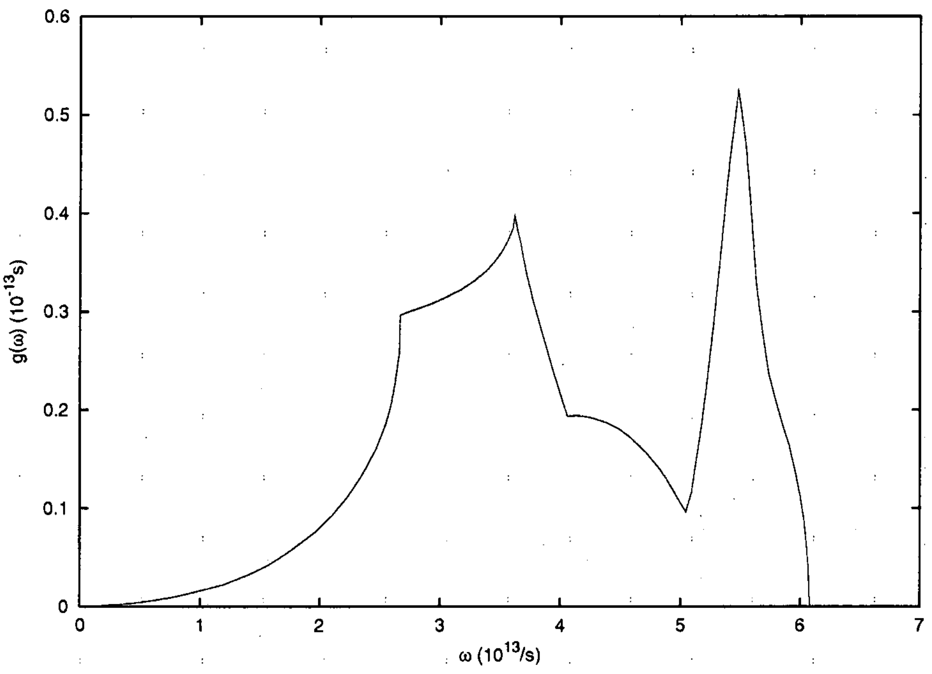

Figure 3.6. Density of phonon states in aluminium, from neutron scattering measurements.

Interatomic potentials

If the interaction between atoms can be represented by a potential function, then the motion of the atoms can be predicted with much less effort than if a full quantum-mechanical calculation is necessary. Simple interatomic potentials are not generally valid for metals, because the energy of the conduction electrons should be treated separately. However, simple potentials have been used by various workers in predicting the EOS of metals, with some success [16]. Accordingly, interatomic potentials were assessed for their accuracy in reproducing the difference in energy between polymorphic phases and in predicting the density of phonon states. The functional forms of the IAPs used are described in Appendix C.

Inverse power potentials were obtained by fitting only even powers of r. A deterministic method was developed for fitting the inverse power potentials – or any linear combination of constant functions of interatomic separation – which was capable of using energy, pressure and higher derivatives of volume, as described in Appendix C. It was possible to find potentials which reproduced calculated LDA or GGA energies. For a given number of terms, the fit to FCC energies was more accurate than the fit to the corresponding BCC energies. Lower powers of r were found necessary in order to give the maximum accuracy. (Table 3.3.)

Table 3.3. Inverse power parameters for aluminium.

fit to	A_i				
	$i = 4$	$i = 6$	$i = 8$	$i = 10$	$i = 12$
FCC LDA	-83.148	1193.0	-3325.9	3056.5	
	-77.527	956.36	-958.33	-6495.4	13903.
BCC LDA	-86.445	1296.7	-3891.6	3871.4	
	-95.411	1672.6	-7616.6	18657.	-21030.
FCC GGA	-70.498	929.25	-2035.9	1016.8	
	-59.063	447.88	2780.9	-18417.	28286.
BCC GGA	-72.803	1009.1	-2492.6	1708.8	
	-95.849	1975.1	-12067.	39713.	-54053.

Units are such as to give potential energy in eV for distances in Å.

Considering either LDA or GGA, potentials were matched to $e(v)$ calculations for a single lattice type. It was found that a potential optimised to match one lattice type did not in general match a potential optimised for the other lattice type. This effect was smaller with fewer powers of r. Also, a potential optimised to match one lattice type did not reproduce energies for the other lattice type – in fact, the predicted FCC-BCC energy difference was uniformly much smaller than suggested by the quantum mechanical calculations. (Fig. 3.7.)

For calculating the density of phonon states, the potential used was that

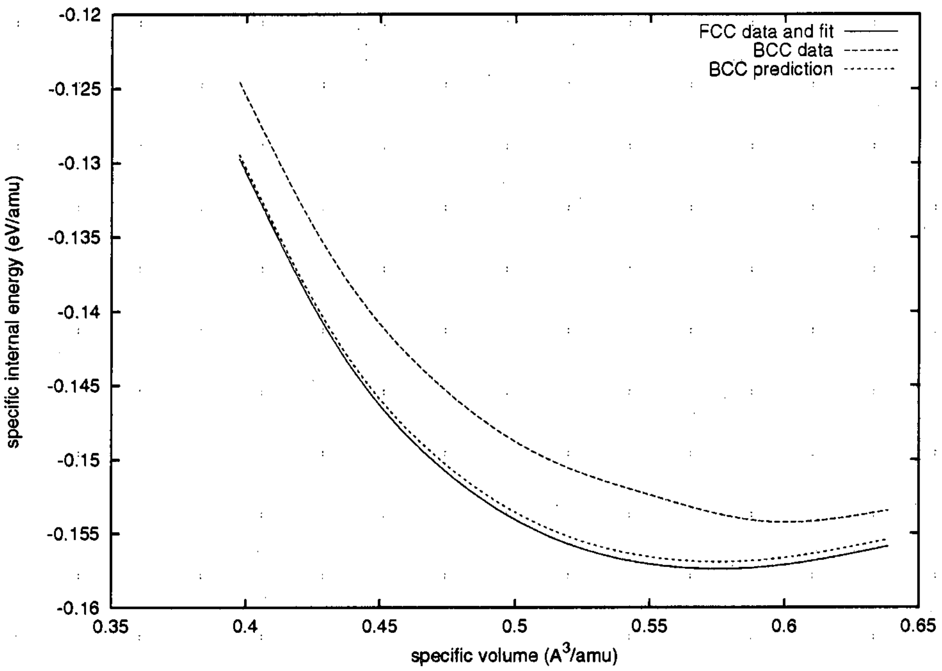

Figure 3.7. Transferability of inverse power potentials.

fitted to LDA FCC energies, with reciprocal powers of 4,6,8,10 and 12 in atomic separation. The density of states was calculated for a repeating cubic assemblage of FCC lattice cells. Because the inverse power potential has an infinite range, the calculations were repeated for an assemblage 3 and 4 FCC cells across, neglecting the potential if atoms were separated by more than 10 and 15 Å respectively for a lattice parameter of 4.0 Å. This procedure explored the degree of convergence in finite simulations. The density of states varied only slightly between these calculations.

The density of states predicted from the inverse power potential had the same shape as the experimental results, but was stretched out in the frequency axis. (Fig. 3.8).

Attempts were made to determine 'long-range' Morse parameters, and also parameters for the more common first or second nearest neighbour Morse interactions. The Morse potential becomes negligible at a finite radius. The significance of 'long-range' parameters is that the number of neighbours rises with compression, as more atoms are pushed within the range of the potential. By contrast, the nth nearest neighbour potentials can be considered as including a cutoff at

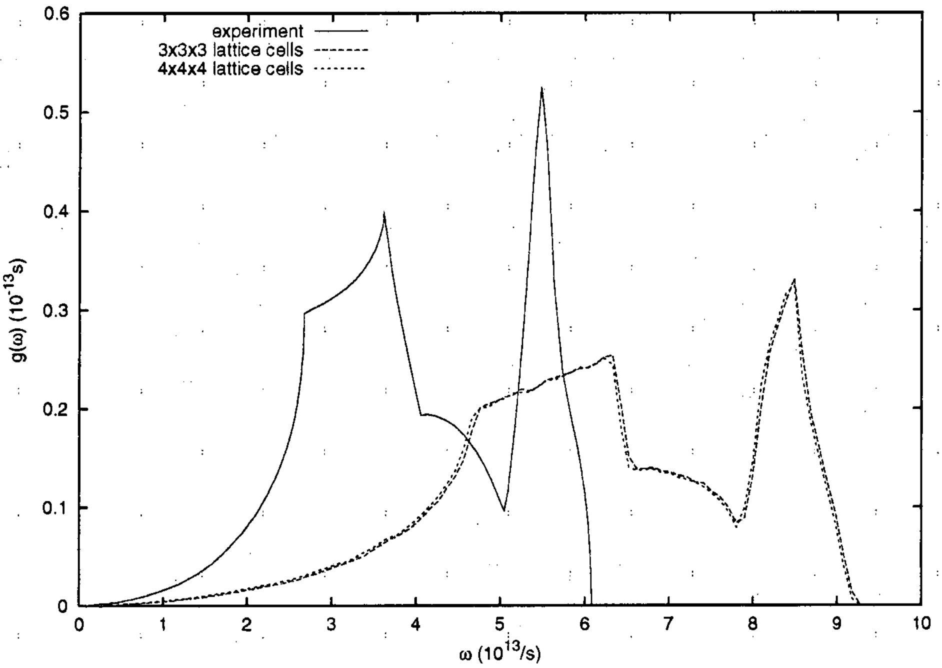

Figure 3.8. Phonon density of states for aluminium, deduced from an inverse power potential.

a finite radius r_c which contracts with compression, so that only the specified neighbours are in range. If the potential is to be used for simulations like molecular dynamics, a variable cutoff is inconvenient because the local density in the simulation can vary significantly over a short distance, leading to ambiguities as to the value of r_c to choose for each atom.

No 'long-range' parameters were found which could match more than the first part of the compression curve, to a few GPa. Potentials limited to first or second nearest neighbours could be matched accurately to the compression data for a single phase. The match to FCC was better than that to BCC. No potential was found which could match the energies for both FCC and BCC phases together: the Morse potential consistently gave too great an energy difference between FCC and BCC. This effect was worse for nearest neighbours only. If the potential was allowed to extend beyond second nearest neighbours, the accuracy started to reduce. The results for GGA were exactly analogous to those for LDA. (Fig. 3.9.)

The parameters which accurately reproduced single-phase cold curves are listed in Table 3.4. The number of neighbours was expressed as a cutoff radius, calculated using the neighbour distances for the FCC structure. The same cutoff radius was then used for the BCC structure.

Table 3.4. Morse parameters for aluminium.

fit to	neighbours	D (eV)	α (1/Å)	r_0 (Å)
LDA FCC	first	0.707858	1.10589	2.79103
LDA FCC	first + second	0.562595	1.06324	2.91452
LDA BCC	first	1.03763	1.09652	2.74014
LDA BCC	first + second	0.619085	1.04627	2.87176
GGA FCC	first	0.639054	1.14083	2.79069
GGA FCC	first + second	0.511498	1.09918	2.90785
GGA BCC	first	0.932628	1.13836	2.73625
GGA BCC	first + second	0.558105	1.08706	2.86544

For calculating the density of phonon states, Morse potentials extending to nearest or next-nearest neighbours and reproducing LDA and GGA FCC energies were used. In all cases, the calculated density of states had the correct shape, but was stretched out in the frequency axis compared with the experimental data, though less so than the densities of states obtained from the inverse power potentials. LDA and GGA gave very similar densities of states. (Fig. 3.10.)

The convergence properties of the Finnis-Sinclair parameters were investigated by choosing as initial values previously suggested parameters for aluminium [56] and also from a set of zeros. The iterations converged to essentially the same values in both cases. The iterative program allowed fitting to all data simultaneously, with a weight associated with each data point. In this way parameters could be optimised to match the data for a single phase independently, or for all sets together.

When starting from a set of zero parameters, a systematic procedure was followed to first determine the values of the 'outermost' (large r) terms in the spline fits. The weights for the data points corresponding to greater compressions (and therefore sampling smaller values of r) were reduced so that the fitting procedure did not choose a 'compromise fit' between different values of r. The fitting procedure was then progressively extended to smaller r, increasing the weights at higher compressions as extra spline terms were activated.

Considering the FCC energies for LDA or GGA alone, the Finnis – Sinclair form could be optimised to reproduce it quite accurately. The potential then reproduced the BCC energies quite well at small compressions, but under-predicted

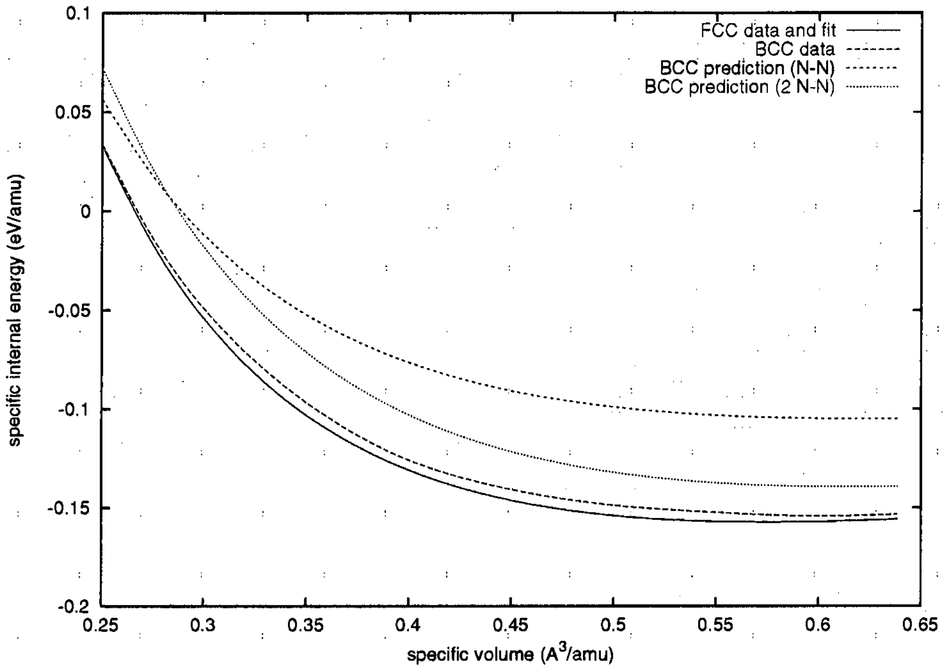

Figure 3.9. $e(v)$ curves for aluminium from Morse fit to LDA FCC.

the difference as compression increased. Fitting both data sets together produced a compromise fit at intermediate compressions. (Tables 3.5 and 3.6, and Figs 3.11 and 3.12.)

The optimised Finnis-Sinclair parameters gave phonon modes which were almost all imaginary. Understating the case, these gave inaccurate predictions of the phonon density of states. The reason for this behaviour is probably that fitting a multi-body potential such as the Finnis-Sinclair to isotropic pressure data is an ill-conditioned process. In order to fit anything more sophisticated than a radial pair potential, non-isotropic data, such as shear modulus or force on displaced atoms, should be included.

A set of empirical parameters was also tried [56]. These reproduced some of the observed properties of aluminium close to ambient conditions, e.g. the compressibility and elastic constants. They did not match the predicted absolute potential energy or the behaviour under large compressions. This potential reproduced more of the observed features of the density of states, although the maximum frequency and the positions of important features were not quite correct. (Fig. 3.13.)

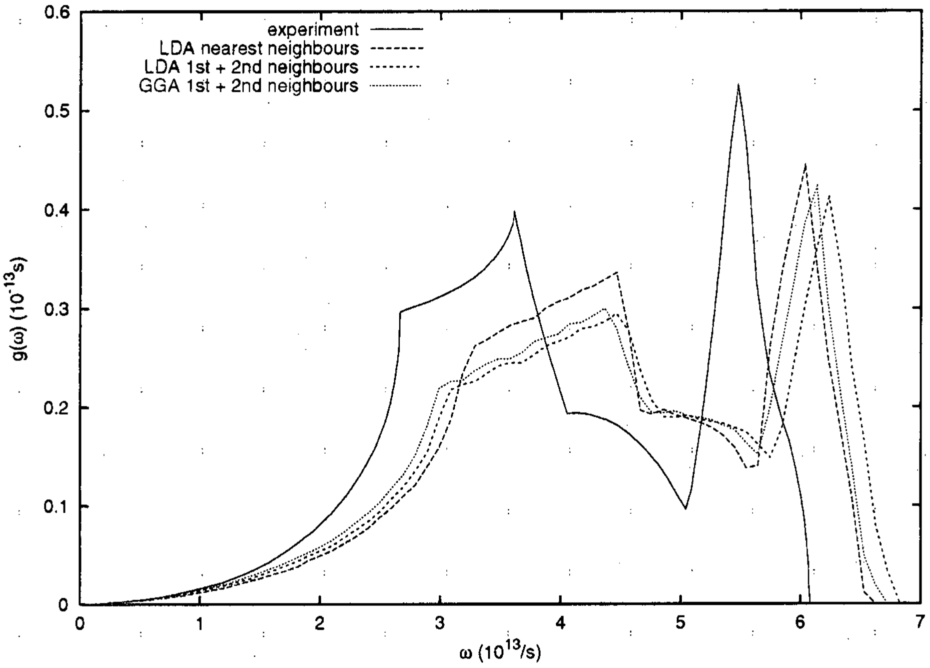

Figure 3.10. Phonon density of states for aluminium deduced from Morse potentials.

Inverse power, Morse and Finnis – Sinclair potentials were all capable of predicting densities of states of roughly the correct shape. The density of states from the radial potentials was stretched out in the frequency axis, the inverse power form more so that the Morse. This discrepancy may be caused partly by fitting different functional forms to discrete energy – volume data, where higher derivatives exaggerate differences in the functional forms and goodness of fit. (The density of states is obtained from the second derivative of the potential function.) This probably explains the difference between inverse power and Morse, where the Morse was found to be a more natural fitting form. The remaining difference (between Morse and experiment) is probably caused by multi-body or angular terms, and the contribution of the conduction electrons, which depends to a first approximation on the density only. In accordance with this explanation, Finnis-Sinclair potentials appeared more promising, but a better optimisation scheme would have to be found to prevent lattice instabilities.

Table 3.5. Finnis-Sinclair parameters for aluminium, fitted to FCC cold curve. (Knots are expressed in units of an 4.0503 Å, which is the experimental lattice parameter for FCC aluminium, but acts as an arbitrary parameter as far as the potential function is concerned. The coefficients are in units such that the potential is calculated in eV.)

	k					
	1	2	3	4	5	6
r_k	1.225	1.18	1.09	1.050	0.9	0.70710678
LDA a_k	2.21834	8.48597	2.29947	1.40611	-5.21263	-4.06097
GGA a_k	2.4189	8.63738	2.4279	0.984752	-6.12842	-8.04307

	k	
	1	2
R_k	1.200	0.93
LDA A_k	66.074	452.965
GGA A_k	59.2106	484.569

Ab initio restoring forces

Restoring forces were predicted from electron ground state calculations using a periodic cell of 2 × 2 × 2 FCC cells. This should give reasonable forces out to second nearest neighbours.

The atom at $(0, 0, 0)$ was displaced by a small amount in the x-direction. The calculations were otherwise the same as for the perfect FCC structure, except that for the 1000 reciprocal space points used to sample the electron states, the reduced symmetry meant that 150 were unique rather than the 35 used previously. Calculations were performed with the cubic lattice parameter a from 3.0 to 4.0 Å, at intervals of 0.1 Å. The restoring force on the displaced atom was adjusted slightly so that the sum of the forces was zero to a higher accuracy. (The sum of the 'raw' forces from CASTEP typically gave a net residual force which, although very small, could be made several orders of magnitude smaller by subtraction.) The force on atoms equidistant from the perturbed atom and its images was scaled by the number of equidistant perturbed atoms. The force on atoms closer to one of the perturbed images than to the perturbed atom at the origin was set to zero. (This procedure is described in more detail in Section 2.3.) The symmetry operations were then used to fill in the rest of the stiffness matrix.

A large number of iterations toward the ground state was required in order to

Table 3.6. Finnis-Sinclair parameters for aluminium, fitted to FCC and BCC cold curves simultaneously.
(Knots are expressed in units of an 4.0503 Å, which is the experimental lattice parameter for FCC aluminium, but acts as an arbitrary parameter as far as the potential function is concerned. The coefficients are in units such that the potential is calculated in eV.)

	k					
	1	2	3	4	5	6
r_k	1.225	1.18	1.09	1.050	0.9	0.70710678
LDA a_k	1.02131	-7.40168	-13.9145	57.9952	-26.4662	67.6306
GGA a_k	0.995125	-5.50213	-18.196	60.7362	-28.7832	63.4892

	k	
	1	2
R_k	1.200	0.93
LDA A_k	18.8918	327.34
GGA A_k	16.1369	347.002

obtain a smooth density of states – the forces converged much less quickly than the total energy. If non-converged forces were used then additional, unphysical structure was predicted in the density of states, and a significant proportion of the phonon modes were imaginary. However, considering the density of states as a distribution, the first few moments (mean, skewness etc) obtained from non-converged calculations were similar to those of the experimental density of states. The lattice-thermal energy is found by integrating over the density of states, so it is useful to compare moments. It was found that the lattice-thermal energy was far less sensitive to the degree of convergence. The density of states was reproduced with reasonable accuracy using converged forces (Fig. 3.14), but non-converged forces were used in predicting the equation of state.

Sensitivity to statistics of phonon model

The effect of the number of frequency bins was investigated by performing calculations with between 10 and 5000 over the full frequency range at one compression. The lattice-thermal energy did not appear to converge unless well over 100 bins were used. This result suggests that the Debye or Einstein models of the lattice modes [40] are unlikely to be adequate for detailed EOS work.

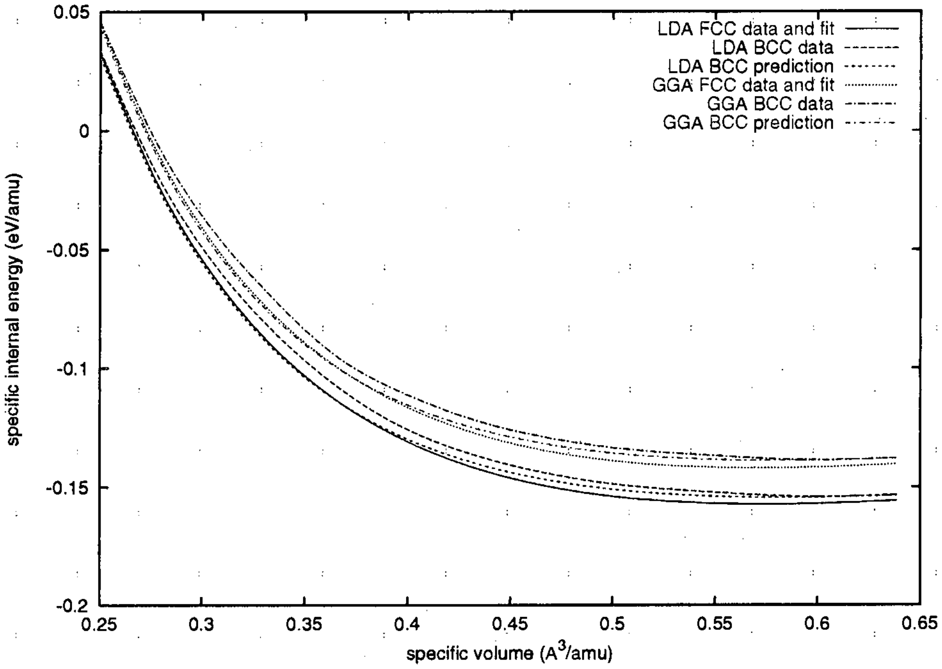

Figure 3.11. $e(v)$ curves from Finnis-Sinclair potentials fitted to FCC ground state energies.

The effect of the number of wavevectors was investigated by performing calculations with between 125 and 8000 wavevectors. Graphs of the lattice-thermal energy were indistinguishable once about 1000 wavevectors were included.

Anharmonic contribution

The importance of phonon – phonon interactions was estimated by performing Monte-Carlo (MC) calculations of the equilibrium configuration of atoms in a group of $5 \times 5 \times 5$ FCC cells. These calculations used the second-neighbour Morse potential fitted to the GGA cold curve for the FCC structure.

Calculations were performed for a variety of densities, and a variety of temperatures at each density. In each calculation, the MC algorithm was applied until the energy had converged, typically ~500 to 1000 iterations, then the total energy was averaged over an equal number of iterations.

For each density, calculations were performed at increasing temperatures, using the final configuration from the previous temperature as the starting configuration for the next temperature. The effect of this procedure was to allow the

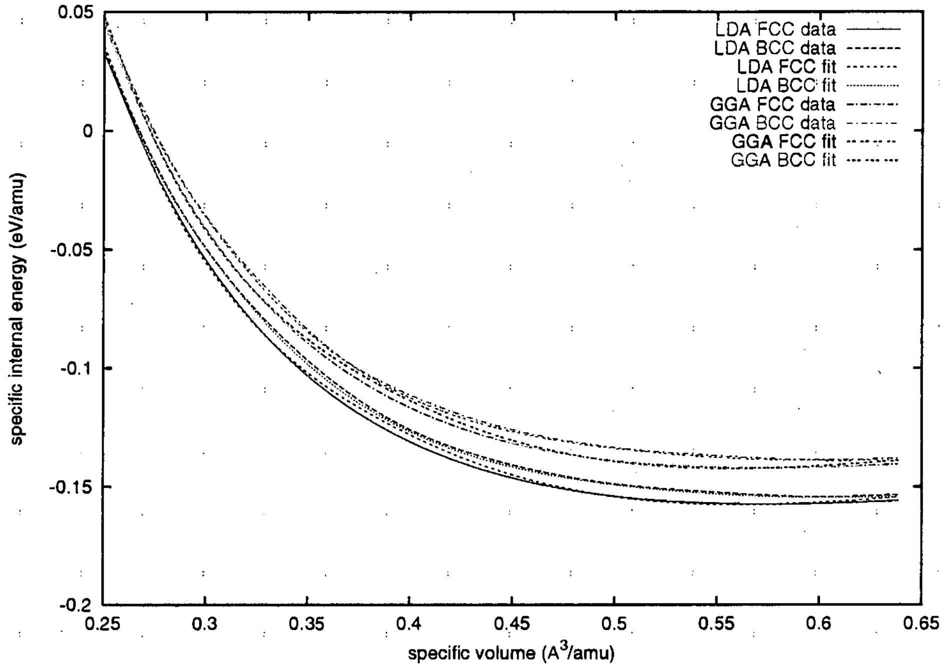

Figure 3.12. $e(v)$ curves for aluminium from Finnis-Sinclair potentials fitted to FCC and BCC ground state energies.

simulations at higher temperatures more time to approach equilibrium.

The anharmonic contribution was estimated for each simulation by comparing the potential energy with the kinetic energy – the latter assumed to be $\frac{3}{2}k_BT$. Neglecting quantum mechanical effects (specifically, the freezing out of modes below their excitation energy), the energies should be equal if each atom sees an effective potential which is quadratic.

For the range of states considered (covering temperatures up to 2000 K), the anharmonic contribution was smaller than the statistical noise in the potential energy.

3.1.4 Electron-thermal contribution

The electron-thermal contribution was estimated using the simple model based on Sommerfeld theory but using the predicted $E_F(\rho)$, and also with the full band structure predicted using CASTEP.

For the band structure, twice as many energy levels were considered at each k-point as would be required on average to contain all the electrons in the ground

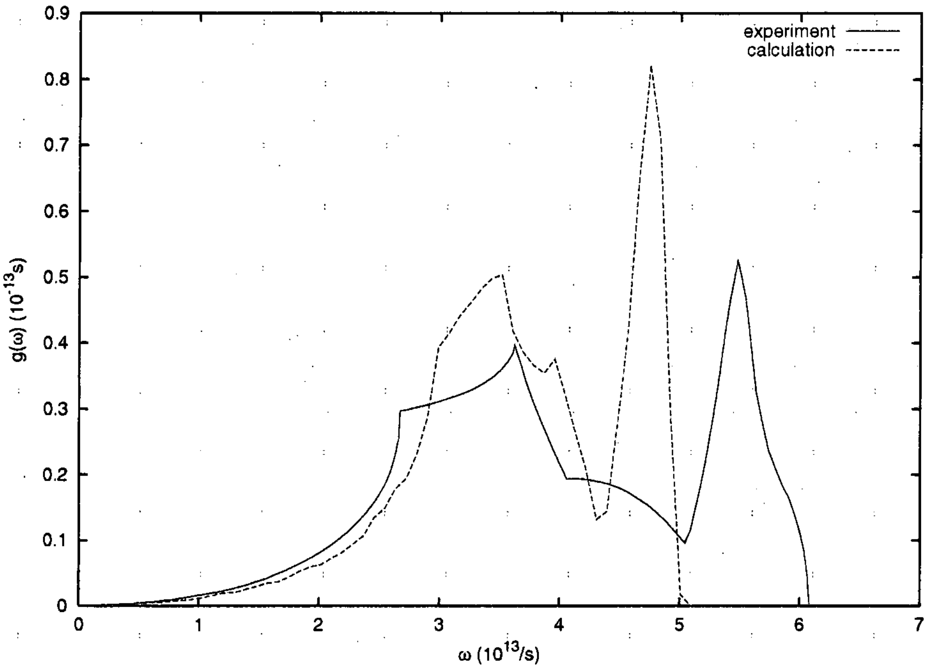

Figure 3.13. Phonon density of states for aluminium deduced from an empirical Finnis-Sinclair potential.

state. This provided a reasonable number of levels above the Fermi energy for the excitation of electrons at finite temperatures. By predicting the occupation of each level as a function of temperature, it was found that the range of levels deduced should be adequate for temperatures up to 2 eV or so. Above this level, it was predicted that appreciable numbers of electrons would be excited to near the top of the levels, so the finite number of levels would start to influence the result. However, in reality once a significant proportion of electrons become excited, the energy levels alter, so the whole treatment loses validity.

The predicted density of levels exhibited a square root dependency, in agreement with free electron theory [40]. Compared with the free electron model, the energy scale is displaced by the binding energy of the electrons with respect to the ions, as represented by the pseudopotentials. For illustration, the densities of levels were modelled by a function of the form

$$g(E) = \sqrt{\frac{1}{\beta}(E - E_0)}, \tag{3.3}$$

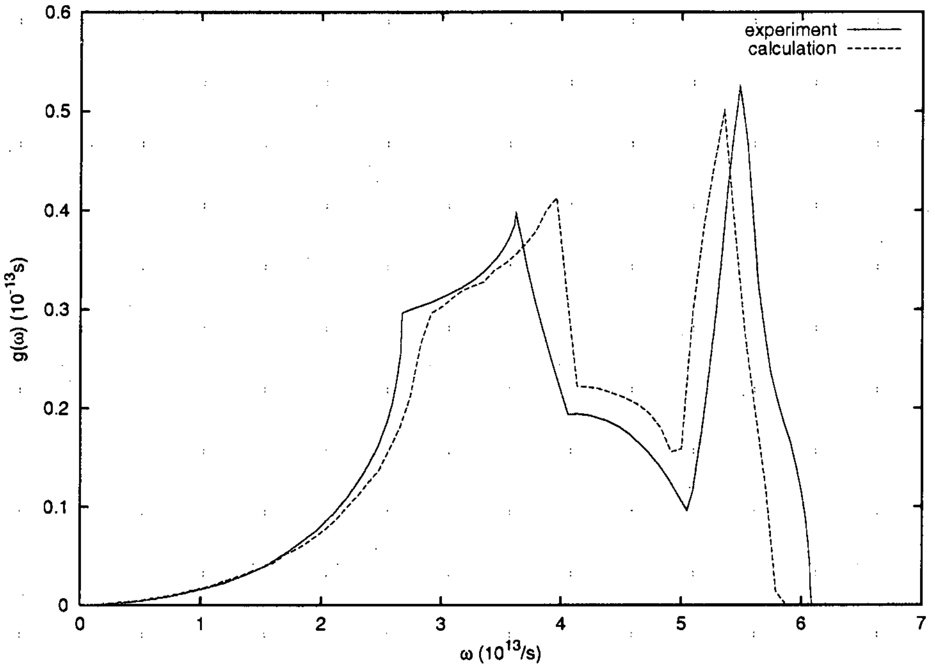

Figure 3.14. Phonon density of states for aluminium deduced from atomic perturbation in the *ab initio* ground state calculation.

where E_0 and β were found by least squares fitting of the band structure calculations in (g^2, E) space (Table 3.7). Subtracting E_0 so that the lowest energy level is zero as for free electrons, the Fermi energy at $a = 4.0$ Å becomes ~11.37 eV, in reasonable agreement with the free electron value of 11.7 eV [40]. (Fig. 3.15)

As one would expect for a simple metal, the modified Sommerfeld model predicted electron-thermal energies which were close to the band structure calculations. This result served as a check that the band structure calculation and associated software were performing correctly, and could be used with confidence

Table 3.7. Free electron fitting parameters for predicted density of electron energy levels in aluminium.

a (Å)	E_0 (eV)	β (1/eV$^{3/2}$.Å^3)
4.0	-2.15	4.25
3.8	-1.43	5.90
3.6	-0.70	8.13

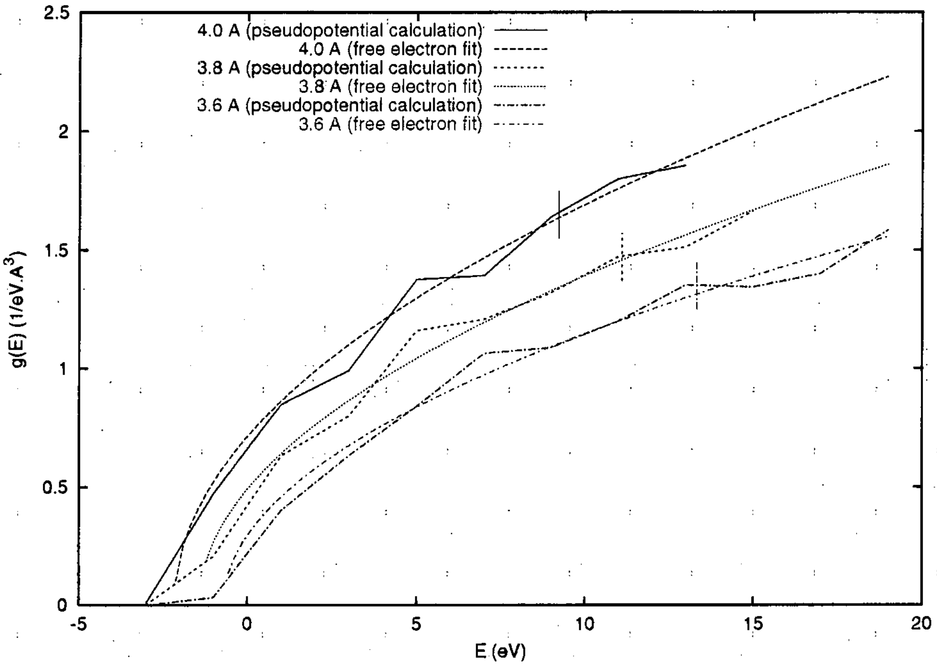

Figure 3.15. Density of electron energy levels in aluminium deduced from the band structure, over a range of values of the FCC lattice parameter.
Vertical bars show the Fermi energy for each density of levels. Free electron fits to the band structure calculations are also shown. (Note: the ragged shape reflects the relatively coarse resolution of k-points.)

for more complicated band structures. (Fig. 3.16.)

3.1.5 Equations of state

An *ab initio* EOS for the FCC structure was generated by combining the frozen-ion cold curve, the phonon modes and the electron-thermal energy as described above.

The pressure offset necessary to correct bring the *ab initio* EOS into agreement with the equilibrium density observed at 293 K was 3.4 GPa (3.3 GPa using linear interpolation). Accordingly, an *ab fere initio* EOS was generated with the specific internal energy tilted by an amount corresponding to 3.4 GPa.

Rigorously, the energy tilt should also manifest itself as a modification to the force constants in the interatomic interaction, and hence to the phonon modes. The effect of this extra correction was evaluated by predicting the lattice-thermal

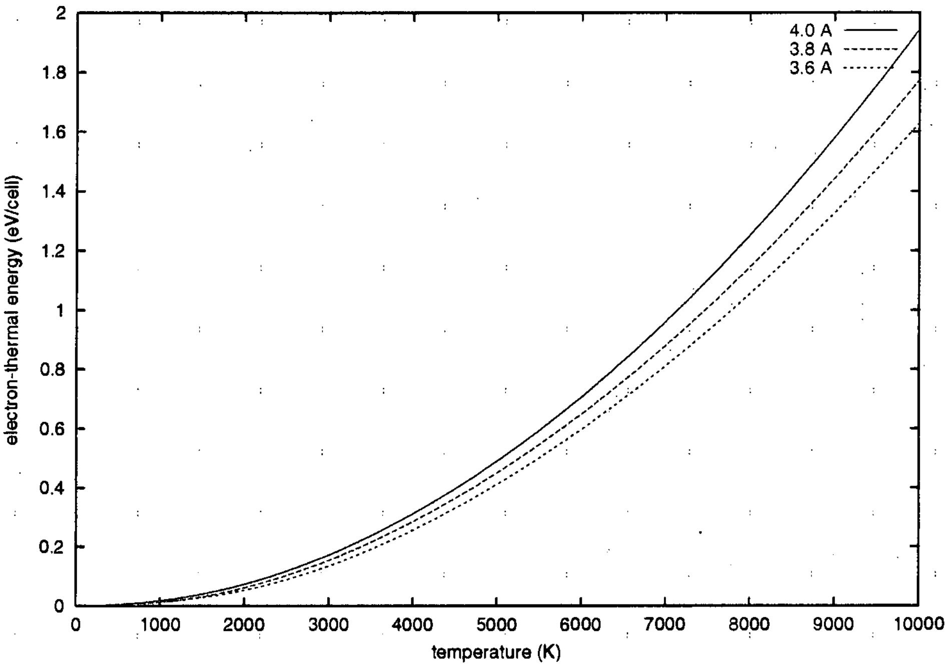

Figure 3.16. Electron-thermal energy predicted from band structure.

energy with and without the correction by calculating the phonon modes using interatomic potentials fitted to the *ab initio* cold curve or to the adjusted cold curve. The difference was found to be negligible.

3.1.6 Comparison with mechanical data

The EOS were evaluated by comparison with mechanical data.

Density at standard temperature and pressure

The STP density of aluminium is observed to be 2.70 Mg/m^3[40]. The *ab initio* EOS gave a density of 2.76 Mg/m^3: ~2% too large, or 0.7% too large in lattice parameter. At 2.75 Mg/m^3, the *ab fere initio* EOS was in good agreement with the experimental value. One would hope to see perfect agreement since this EOS was adjusted to improve the match; the agreement observed confirms that the adjustment was made correctly and reflects the uncertainties caused by interpolation and truncation errors.

The accuracy of the *ab initio* EOS was in line with expectations from the use of the local density approximation.

293 K isotherm

The 293 K isotherm was extracted from each EOS, using quadratic interpolation between the ordinates. The isotherm was compared against diamond anvil measurements [71] (Fig. 3.17).

The *ab fere initio* EOS was significantly closer to the data than was the *ab initio* EOS, particularly at lower pressures. The agreement became poorer at densities over ~4.5 Mg/m^3. This may be a deficiency in the calculations, as the electrons are forced progressively into the core region as the density increases, making the pseudopotential model less accurate. It is also possible for systematic errors to occur in the diamond anvil results, as it becomes more difficult to ensure an isotropic stress at higher pressures.

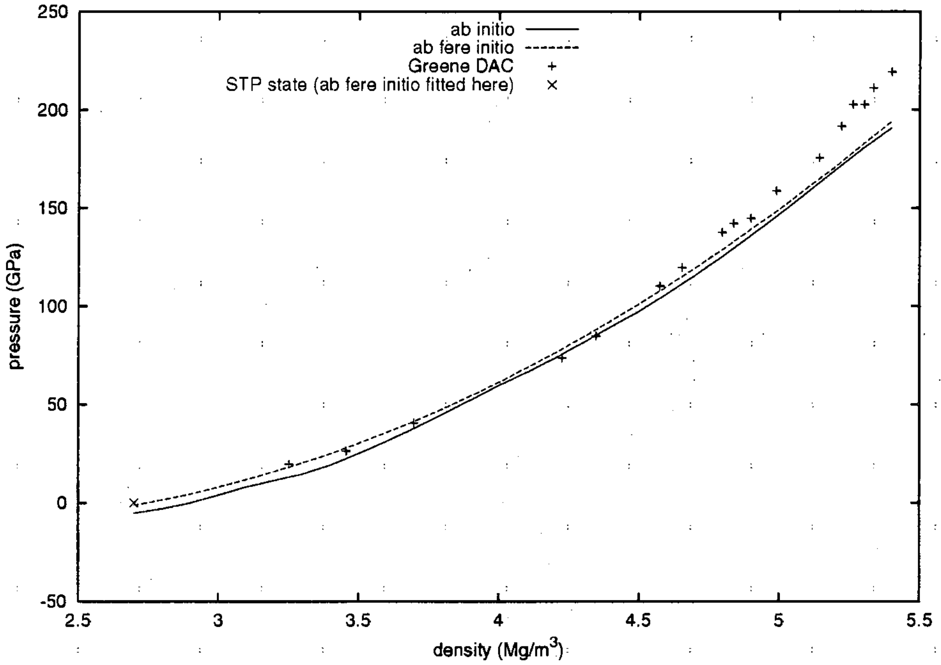

Figure 3.17. 293 K isotherm for aluminium.

Shock Hugoniot

The shock Hugoniot for each EOS was predicted and compared with published data on aluminium alloy 1100 [72, 73], which is close to the pure element.

The Hugoniot was calculated in terms of the relations between shock speed u_s and particle speed u_p (Fig. 3.18), and pressure p and density ρ (Fig. 3.19).

As with the 293 K isotherm, the *ab fere initio* EOS was significantly closer to the data than was the *ab initio* EOS. The agreement became slightly poorer at densities over ~4.0 Mg/m^3, first falling below the experimental data and then rising above. This behaviour is characteristic of an EOS which is too soft: eventually the thermal energy increases rapidly. Once the theoretical EOS begins to depart from the experimental cold curve (or 293 K isotherm), Hugoniot temperatures predicted by the EOS are likely to be increasingly inaccurate.

The bilinear interpolation scheme is the probable cause of the undulations which show up most clearly in the $u_s - u_p$ relation. Linear interpolation should not cause a cumulative error at higher pressures, because Hugoniot states are found from a root-finding process involving the EOS (see Appendix F). This is in contrast to the situation in finding isentropes, which require an integration of $-p\,dv$, and which therefore would accumulate a contribution from the use of linear interpolation. The EOS surface is convex ($\partial^2 p/\partial\rho^2 > 0$) for FCC Al, so linear interpolation is an approximation which always lies on or above $p(\rho)$ and hence overestimates the integral of p with ρ or v.

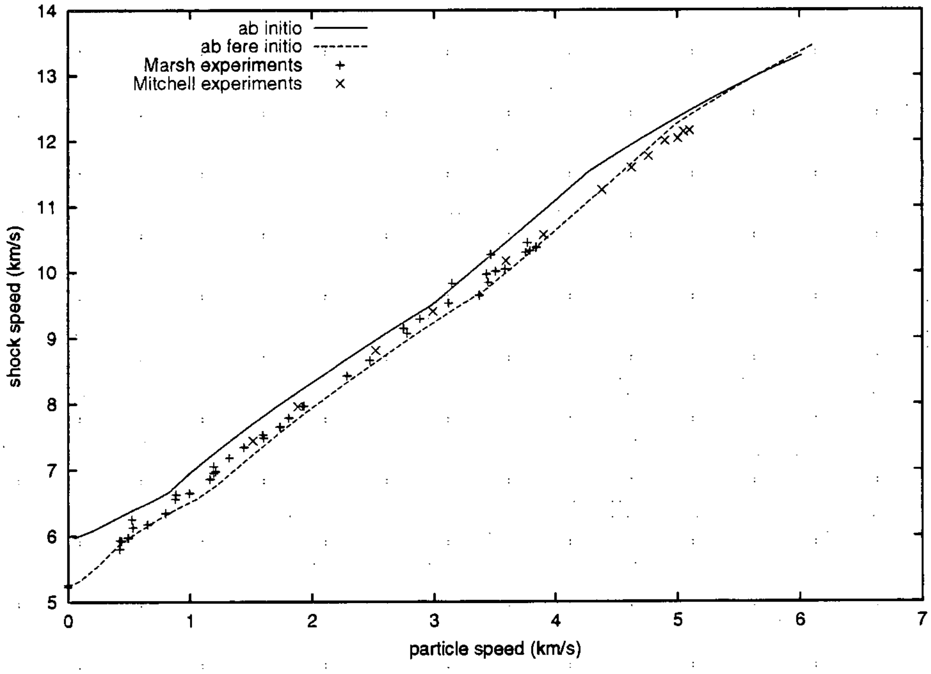

Figure 3.18. Shock speed – particle speed relation for aluminium.

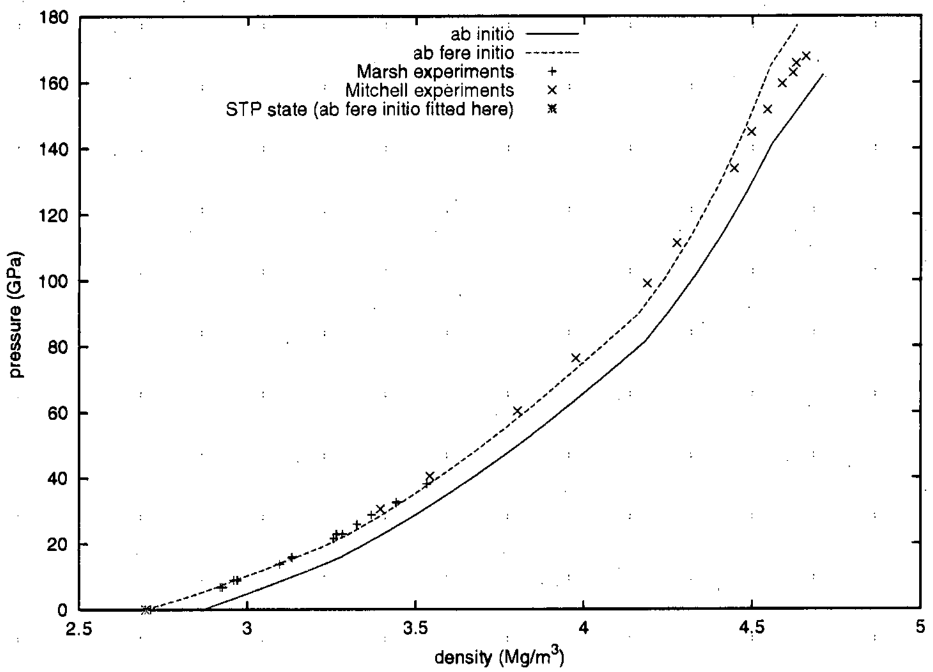

Figure 3.19. Shock pressure – density relation for aluminium.

3.2 Silicon

Various current research programmes are developing methods for measuring the changes in Bragg reflections of X-rays as a shock wave passes through a crystalline material, and hence inferring the mass density [74, 75]. Silicon is a convenient material, because large crystals are available relatively cheaply from the semiconductor industry. In order to relate the density to the pressure, the EOS must be known.

The objective was to calculate the EOS up to the relatively modest compressions likely in experiments with shock states up to a few tens of GPa – a few tens of percent change in lattice parameter. The 'core overlap' problem discussed above should not occur in this regime.

3.2.1 Diamond structure

Cold curve

CASTEP calculations were made on a single lattice cell of the diamond structure. Plane waves with energies up to 400 eV were used in the basis set. Electron states were calculated at 1000 evenly-distributed points in reciprocal space. The cubic symmetry of the diamond structure means that only 35 of these points were unique, reducing the amount of calculation required. The Pulay stress, previously found to be negligible for aluminium (with a similar pseudopotential and a smaller plane wave cutoff), was ignored.

Electron ground state calculations were performed for values of the cubic diamond lattice parameter a between 4 and 5.5 Å, at intervals of 0.1 Å. The result was a set of discrete points (ρ, e_c) on the cold curve (Fig. 3.20). These results have an unknown absolute energy offset which could be found by performing a calculation with a single isolated atom. The offset was not found in this initial study, as it has no effect on any aspect of the EOS in compression.

Lattice-thermal contribution

The lattice-thermal contribution $e_l(\rho, T)$ was determined by deriving the quasi-harmonic phonon modes for the lattice. Attempts were made to obtain radial pair potentials, but although it was found possible to optimise the parameters in a Morse potential to reproduce the cold curve, the resulting potential predicted

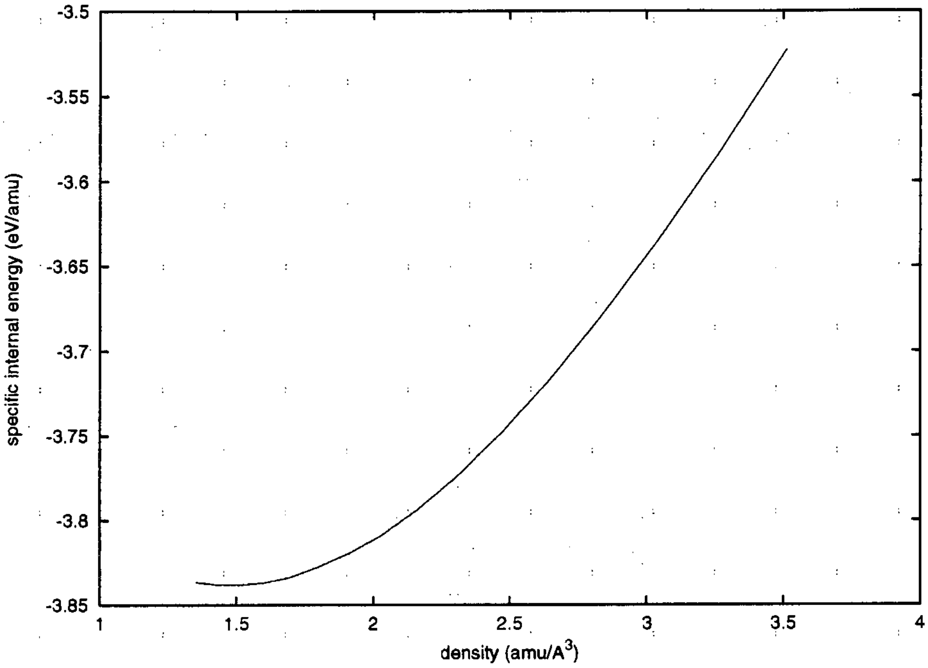

Figure 3.20. Calculated frozen-ion cold curve.

many imaginary phonon frequencies for the diamond structure. In principle, this might indicate that an atomic structure is unstable with respect to certain displacements of the atoms, and thus that the structure is not the equilibrium one. However, non-close-packed structures are stabilised by angular contributions to the interatomic potential. Radial pair potentials cannot model these contributions, and are therefore not suitable for non-close-packed structures. (An extreme example is the whole of organic chemistry, where molecules can be visualised as a set of atoms connected by localised, directional bonds.) Potential functions such as Stillinger – Weber and Tersoff potentials [101] are designed to model directional bonds. However, as was demonstrated above in the attempts to deduce *ab initio* interatomic potentials of aluminium, the problem of defining all the parameters in an multi-body or angular potential from calculations of isotropic stress is ill-defined.

Rather than using empirical angular or n-body potentials for silicon, phonons were obtained from direct evaluation of the forces when an atom is perturbed from its equilibrium in the electron ground state calculations.

CASTEP calculations were performed with the atom at $(0, 0, 0)$ displaced by a small amount in the x-direction. The calculations were otherwise the same as for the perfect diamond structure, except that for the 1000 points in reciprocal space, the reduced symmetry meant that 150 were unique. Calculations were performed at intervals of 0.5 Å in the cubic lattice parameter a between 4.0 and 5.5 Å, and also at intervals of 0.1 Å between 5.0 and 5.5 Å. The force on the displaced atom was adjusted slightly so that the sum of the forces was zero to a high accuracy. (The sum of the 'raw' forces from CASTEP typically gave a small net residual force.) The force on atoms equidistant from the perturbed atom and its images was scaled by the number of equidistant perturbed atoms. The force on atoms closer to one of the perturbed images than to the perturbed atom at the origin was set to zero. (This procedure is described in more detail in the previous chapter.) The symmetry operations were then used to fill in the rest of the stiffness matrix, as described in Section 2.3.

Solving the phonon frequency eigenproblem for each $\vec{k}$ gave 24 frequencies ω (three for each atom in the lattice cell). These were collected into regularly-spaced bins, giving an approximation to the density of states $g(\omega)$ (Fig. 3.21).

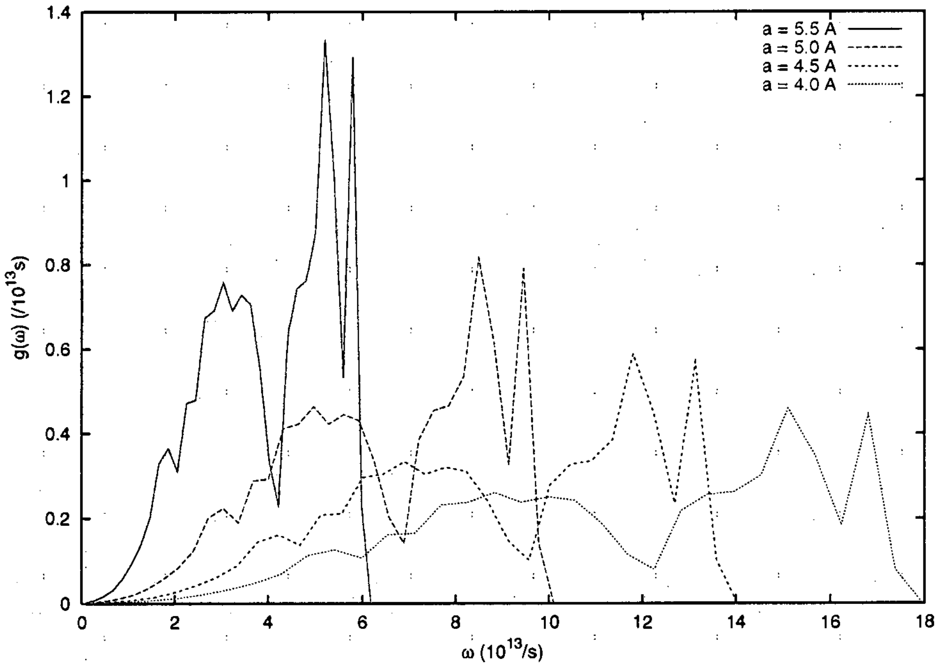

Figure 3.21. Phonon densities of states (displacement of 0.01 a).

It was found that – unlike the radial pair potential tried initially – the ground state force calculations predicted phonon modes which were entirely real, except for a small proportion (about 3 to 4%) at the highest compression, a =4.0 Å. These were long wavelength modes, signalling the onset of a static instability of the diamond structure in favour of the BCT [78].

A table of e_l values was calculated from each density of phonon states $g(\omega)$ for $0 \leq T \leq 10000$ K. 201 values of T were used, with intervals varying geometrically according to a 1% expansion factor between adjacent intervals.

Above about 1000 K, e_l is very close to linear in T. The specific heat capacity is by then very close to the Dulong and Petit value of $3k_B$ per atom [40].

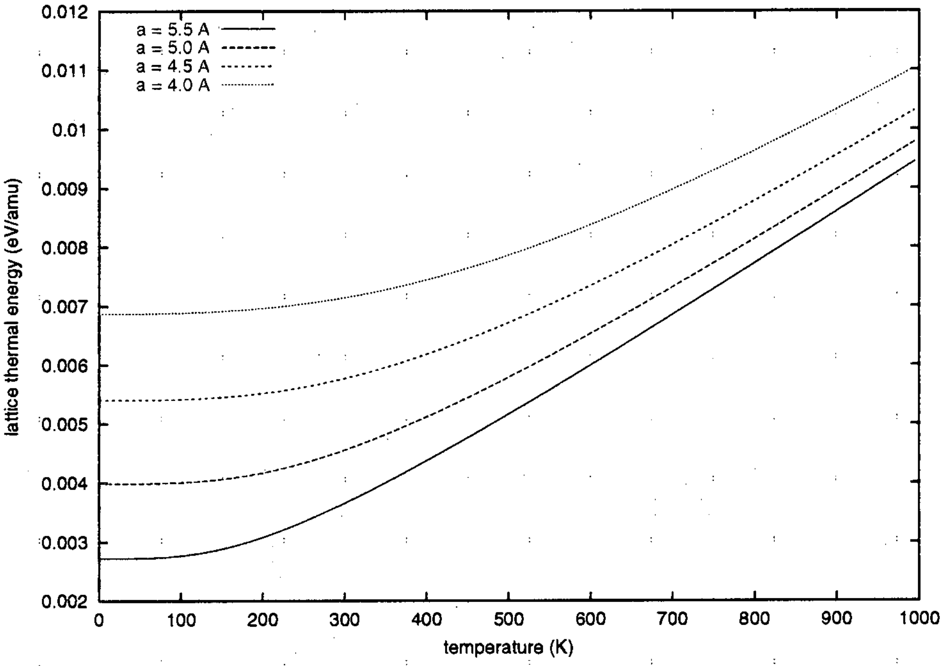

Figure 3.22. Variation of lattice-thermal energy with temperature (displacement by 0.01 a).

The variation of $e_l(T)$ with v (Fig. 3.22) demonstrates the significance of the zero-point lattice energy, increasing with compression as the phonon frequencies increase. Also demonstrated is the greater temperature necessary for a given rise of e_l over the zero-point value, as the Boltzmann occupation factor for a given mode at a given temperature decreases with increasing ω. To the compressions considered, the effect is significant well over 1000 K.

Anharmonicity: effect of different displacements

For each value of the lattice parameter a, forces were calculated for displacements of the atom at $(0, 0, 0)$ by 0.001, 0.01 and 0.02 of the side of the $2 \times 2 \times 2$ supercell, in the x-direction. The stiffness α, defined in Section 2.3, was calculated for each lattice parameter and displacement. The potential well was found to be slightly soft, i.e. the stiffness reduced slightly with displacement (Fig. 3.23).

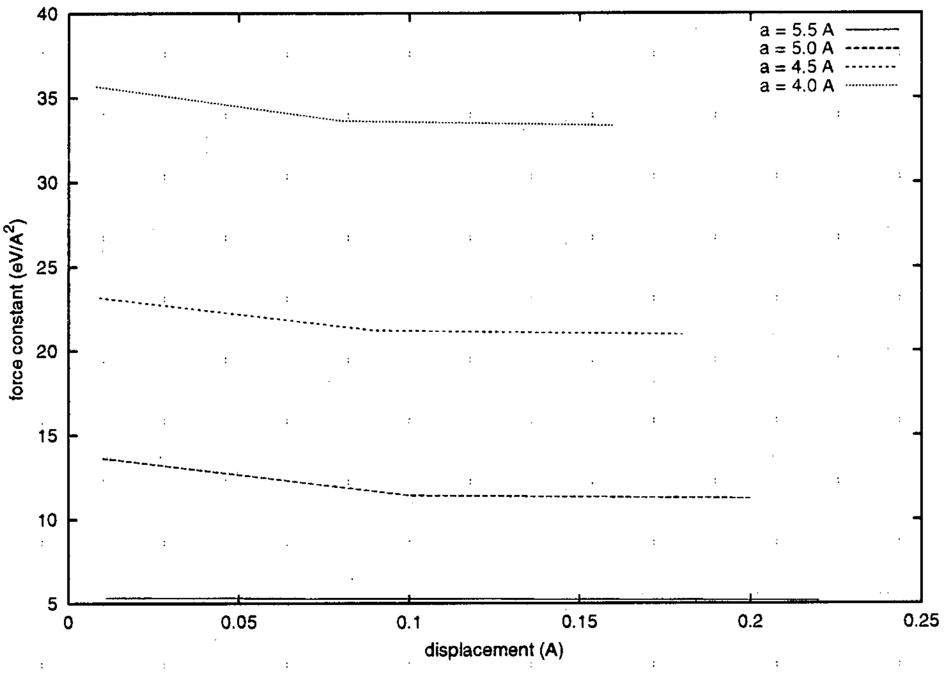

Figure 3.23. Variation of the lattice force constant or stiffness with displacement, for different values of the diamond cubic lattice parameter a.

Using the stiffnesses for a displacement of 0.01 times the lattice parameter, the magnitude of the displacement was predicted as a function of temperature (Table 3.8). These are overestimates where the heat capacity is less than the classical value, i.e. where the variation of e_l with T is curved. The calculations indicate that displacements of the order used fall roughly in the same range as those anticipated at temperatures in the range of interest.

The different stiffness matrices gave slightly different densities of states (Fig. 3.24, calculated with a relatively coarse set of wavevectors for illustration). The resulting variation of e_l with T was quite similar (Fig. 3.25).

It should be noted that this approach provides at best an indication of the

Table 3.8. Estimated atomic displacements in Å.

a (Å)	T (K)		
	273	1000	10000
5.5	0.049	0.093	0.294
5.0	0.033	0.063	0.198
4.5	0.024	0.045	0.143
4.0	0.019	0.036	0.113

sensitivity to anharmonic effects. Rigorously, the variation of stiffness with displacement should be treated as an interaction between quasiharmonic phonons, altering the effective frequency of each mode [38]. The population of each mode should still be found from Maxwell-Boltzmann statistics, because the modes are assumed to be in equilibrium with a heat bath of infinite capacity [2].

Interpolating densities of states

To investigate the accuracy of the interpolation schemes and the resulting accuracy of the lattice-thermal energy, interpolations were made over two different intervals of lattice spacing and compared with the frequencies $\{\omega_i\}$ from actual ground state calculations. The phonon frequencies calculated at $a = 4.5$ and 5.5 Å were interpolated to compare the frequencies at $a = 5.0$ Å. The density of states obtained by interpolating phonon modes over an interval of 1 Å was a poor match to the rigorous calculation. This partly reflects the choice of interpolating function, but is mainly caused by mode incoherence on changing the lattice parameter. The latter effect explains the absence of low frequency modes ($0 < \omega < 5$ THz) in the interpolated density of states, even though the calculations at both ends had modes in this range. The infrequently-observed modes in this range have been averaged with modes from the more common higher frequencies.

Similarly, the phonon frequencies calculated at $a = 5.0$ and 5.5 Å, and at 5.2 and 5.3 Å were interpolated to compare the frequencies at $a = 5.25$ Å. The density of states obtained by interpolating phonon modes over an interval of 0.5 Å or less was in much better agreement with the rigorous calculation.

Comparing the lattice-thermal energy even the poor density of states produced a curve which was almost indistinguishable from the rigorous calculation. This is evidence that the density of states or phonon dispersion relation are extremely

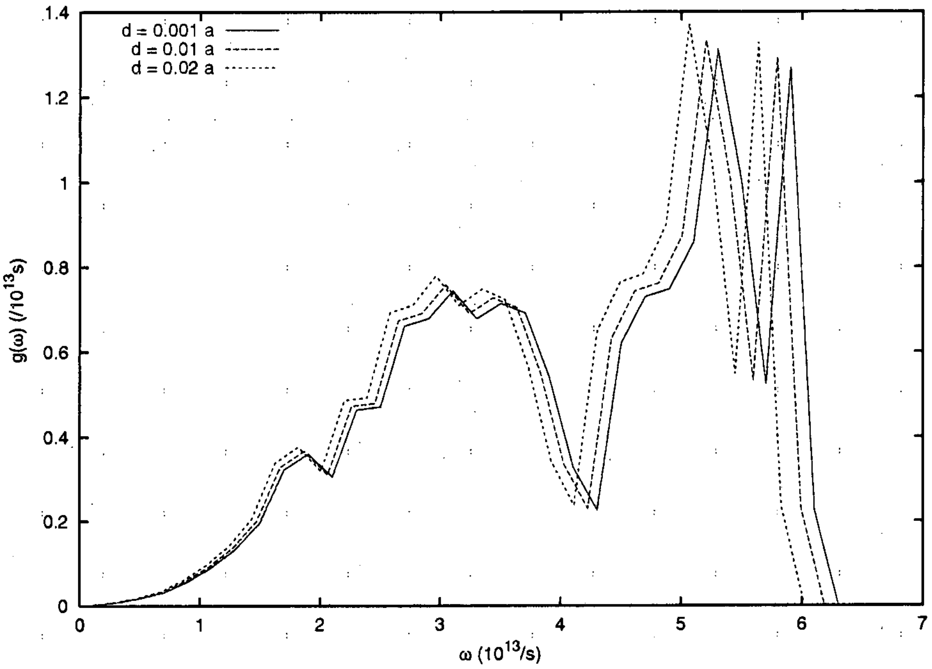

Figure 3.24. Sensitivity of the density of phonon states to the atomic displacement. (a = 5.5 Å)

sensitive tests of the lattice-thermal contribution to the EOS: the EOS appears insensitive to surprisingly gross changes in the phonon density of states. Other aspects such as phase boundaries show a greater sensitivity.

For the purposes of calculating an EOS, the linear interpolation scheme was used to predict approximate phonon frequencies for lattice spacings at intervals of 0.1 Å between the ground state calculations, which were at intervals of 0.5 Å for $a \leq 5.0$ Å. This was deemed reasonable, as an accurate equation of state at higher pressures should take polymorphism into account anyway. Phonon interpolation was not used at the lower densities, as ground state force calculations were made for each point on the cold curve.

Equations of state

An *ab initio* EOS was produced with the lattice-thermal contribution deduced from atom displacements of 0.01 a. The equilibrium (p = 0) density was predicted to be 2.426 Mg/m^3 at 0 K and 2.422 Mg/m^3 at 293 K. These densities were deduced using quadratic interpolation between the table ordinates. Using linear

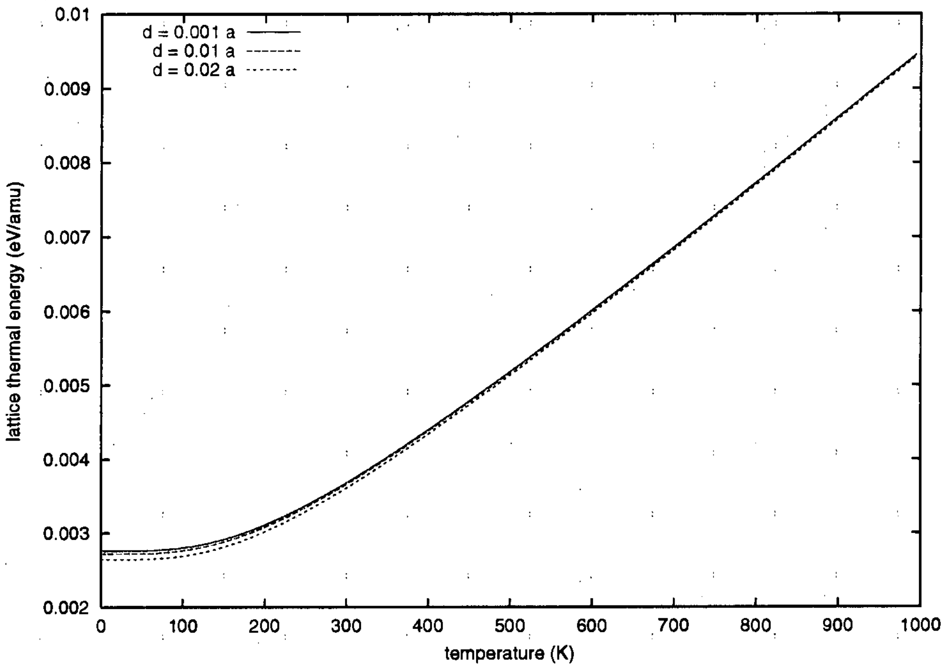

Figure 3.25. Variation of lattice-thermal energy with temperature for different displacements. (a = 5.5 Å)

interpolation, the densities were 2.424 Mg/m^3 at 0 K and 2.419 Mg/m^3 at 293 K. (The density of condensed matter at 1 atmosphere is practically identical to that at zero pressure.)

The sensitivity to the magnitude of the atom displacement was investigated by repeating the procedure with the lattice-thermal data generated with displacements of 0.001 a. The predicted values of equilibrium density were 2.424 Mg/m^3 at 0 K and 2.421 Mg/m^3 at 293 K, using quadratic interpolation.

The observed density of silicon at ambient temperatures is 2.33 Mg/m^3[40], so the *ab initio* densities were just under 4% too large. This is similar to the accuracy obtained for aluminium, and has been attributed to deficiencies in the use of the LDA for the electron ground state.

The pressure at the expected equilibrium density was -3.7 GPa using quadratic interpolation and -3.6 GPa using linear interpolation. Using interpolated phonon modes instead of using the rigorous force calculations between 5.0 and 5.5 Å, these pressures were -3.4 and -3.3 GPa respectively. This demonstrates further that the EOS is relatively insensitive to variations in the detail of the density of phonon states.

The pressure offset necessary to correct bring the *ab initio* EOS into agreement with the equilibrium density observed at 293 K was 3.7 GPa (3.6 GPa using linear interpolation). Accordingly, an *ab fere initio* EOS was generated with the specific internal energy tilted by an amount corresponding to 3.7 GPa. The normal density then agreed with the desired value to better than 0.1%.

The physical models which were used resulted in a plausible-looking EOS for silicon in the diamond structure. The number of ground state calculations performed meant that the coverage in the range 0 to 20 GPa or so was relatively sparse, so the detailed representation of the regime where the diamond structure is known to exist is rather crude.

The correction which had to be applied to the *ab initio* EOS was significant in comparison with this pressure range. The zero-point energy of the lattice was about half as large – an important addition, and evidence of the importance of a quantum-mechanical treatment of the lattice-thermal energy. The predicted difference between the isotherms at 0 and 273 K was comparatively small. Similarly, the difference between the shock Hugoniot and the isotherm passing through the initial state was predicted to be quite small up to 10 GPa or so. However, the Hugoniot was calculated to diverge significantly from the isotherm above a few tens of GPa. (Figs 3.26 and 3.27.)

3.2.2 Body-centred tetragonal structure

The body-centred tetragonal (β-Sn) structure can be obtained from the diamond by compressing along one of the coordinate directions. It has been observed in diamond anvil experiments, and predicted by LMTO and pseudopotential calculation to be energetically favourable under compression at $T = 0$ [78]. At low temperatures, the Imma phase is predicted to be more stable than BCT.

CASTEP calculations were performed with plane waves up to 300 eV and 1000 k-points (75 unique under symmetry operations). Calculations were performed with a cutoff of 400 eV at the extremes of lattice parameter used in order to estimate the significance of this difference, i.e. the importance of more localised states. The difference in energy and stress was found to be quite small, suggesting that 300 eV was an adequate cutoff to use with this pseudopotential to obtain results that were converged over plane waves.

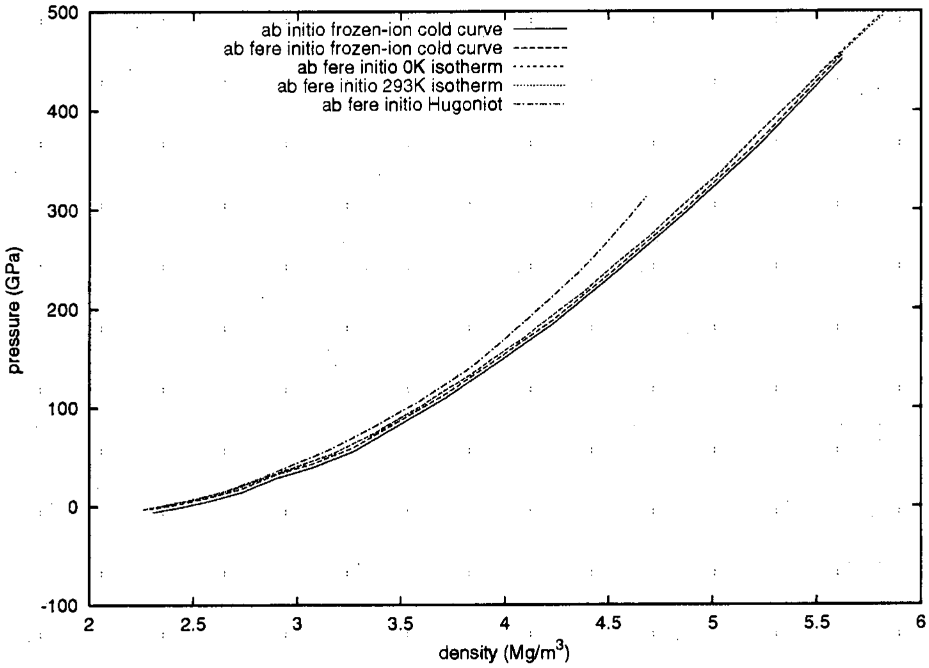

Figure 3.26. Comparison between frozen-ion cold curves, isotherms and Hugoniot for silicon in the diamond structure.

Eliminating the internal degree of freedom

The equation of state desired in this work is that relating pressure to density. The BCT structure has an internal degree of freedom – the c/a ratio. For a given density, there is likely to be a single value of c/a which gives an isotropic stress. This value cannot be predicted *a priori*, so it is necessary to find it in some way.

Ideally, a ground state calculation would be performed at the desired c/a ratio for each density. In practice, calculations are performed for a range of c/a ratios close to the desired value, according to various possible schemes described below. These may include a calculation which is sufficiently close to the desired state to be used directly, but in practice the robust (simplistic) procedures used over most of the density range tended to bracket states of isotropic stress in an interval over which the variations in energy, stress and/or lattice parameters were significant.

The preferred method was to determine the ground state energy and stress over a sufficient range of values of c/a that the value giving an isotropic stress could be found by interpolation. The energy corresponding to the same interpolated point can then be obtained easily.

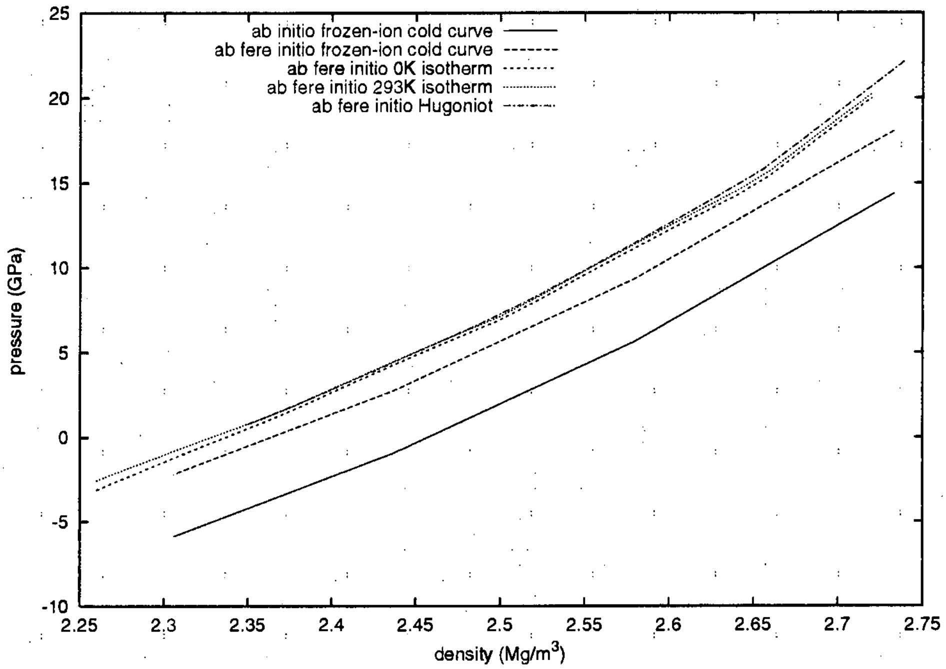

Figure 3.27. Comparison between frozen-ion cold curves, isotherms and Hugoniot for silicon in the diamond structure (detail near $p = 0$).

An alternative method can be used if the result need not be so accurate, e.g. in regions of the state which the material is not expected to reach, for instance if it is known that a different state has a lower free energy. The method assumes that the elastic constants are all equal (or that the relation between them is known). In this case, a series of ground states of non-isotropic stress (e.g. with any internal degrees of freedom held constant) can be corrected to predict the isotropic states by estimating the strain necessary to relieve the stress deviator, estimating the elastic constants from the bulk modulus.

If a series of calculations was performed over a range of density with the internal degrees of freedom held constant, it is straightforward to estimate the variation of bulk modulus with density from the pressure – volume or energy – volume relation. The method then works as follows:

1. Calculate the ground state energy e and stress σ for each of a set of lattice parameters related to density ρ according to some simple prescription for the internal degrees of freedom, e.g. hold them constant.

2. Calculate the mean pressure

$$\bar{p} = -\mathrm{Tr}\,\sigma/3 \tag{3.4}$$

   and stress deviator

$$s = \sigma + \bar{p}I. \tag{3.5}$$

3. Estimate the bulk modulus at each density:

$$B = -v\frac{\partial \bar{p}}{\partial v} \tag{3.6}$$

   using smooth functions to interpolate between the set of ground states.

4. Estimate the elastic constants c from the bulk modulus, using any knowledge available about their relative magnitudes, e.g.

$$B \simeq \sum_i \beta_i [c]_i \tag{3.7}$$

   where the β are constants of proportionality for the elastic constants c. The assumption about the relative magnitudes might be expressed as a constraint that the $\alpha_i[c]_i$ are constant for some α_i. If several states were obtained at or near the same density, with different values of the free parameters, then the elastic constants can be determined quite rigorously.

5. Predict the elastic strain required to relieve the stress deviator, e.g. assuming small independent displacements:

$$[\epsilon]_i = -[s]_i/[c]_i. \tag{3.8}$$

6. Apply the strain ϵ, predicting the density, pressure and energy at the new state.

This method was tried out as a way of correcting a series of calculations of constant c/a performed at lower densities than were likely to be required (where the diamond structure was predicted to be the equilibrium phase) in order to extend the range of the equation of state for the BCT phase without requiring much computation.

The more general method of deducing elastic constants from stress – strain calculations is to use the relation

$$[c]_{ij} = \frac{\Delta\sigma_j}{\Delta\epsilon_i}, \tag{3.9}$$

where now the elastic constants form a tensor of rank 2 [76, 77].

Calculations with constant c/a

An initial set of calculations was made, varying a for several values of a constant c/a ratio. The results demonstrated that the c/a ratio is quite sensitive to a. Coincidentally, a single state of roughly isotropic stress was obtained from one of the calculations.

At larger volumes, the states with $c/a = 3/4$ had the lowest energies. These states were used to estimate the ground state with isotropic stress, using the procedure described above and assuming (unrealistically) that the isotropic elastic constants are equal ($\alpha_i = 1 : i \in \{11, 22, 33\}$) and that the second Lamé constant $\mu = 0$. Then $B = c_i : i \in \{11, 22, 33\}$ [76]. The bulk modulus was estimated from the variation of mean pressure with density using piecewise quadratic interpolation. Linear and cubic interpolation produced similar results. The final point (a =5.0 Å) was over-sensitive to the interpolation scheme, and was disregarded. (Fig. 3.28.)

Box-moving calculations for constant hydrostatic pressure

A box-moving algorithm was used at lower pressures to obtain $c(a)$ for an isotropic applied stress. This method assigns an effective mass to the lattice parameters, and allows the lattice cell to deform dynamically under the action of the net force (difference between applied stress and ground state stress at the instantaneous configuration) and a damping term. As well as being an efficient labour-saving device, this algorithm can converge arbitrarily close to states of the desired isotropic stress.

The algorithm was found to be unreliable at higher pressures – the optimisation procedure is a tradeoff between minimising the energy for a given structure, and reducing the energy by altering the structure. At higher pressures the algorithm tended to allow the structure to change too rapidly, resulting in unstable oscillations in a and c between successive iterations. Altering the box-moving

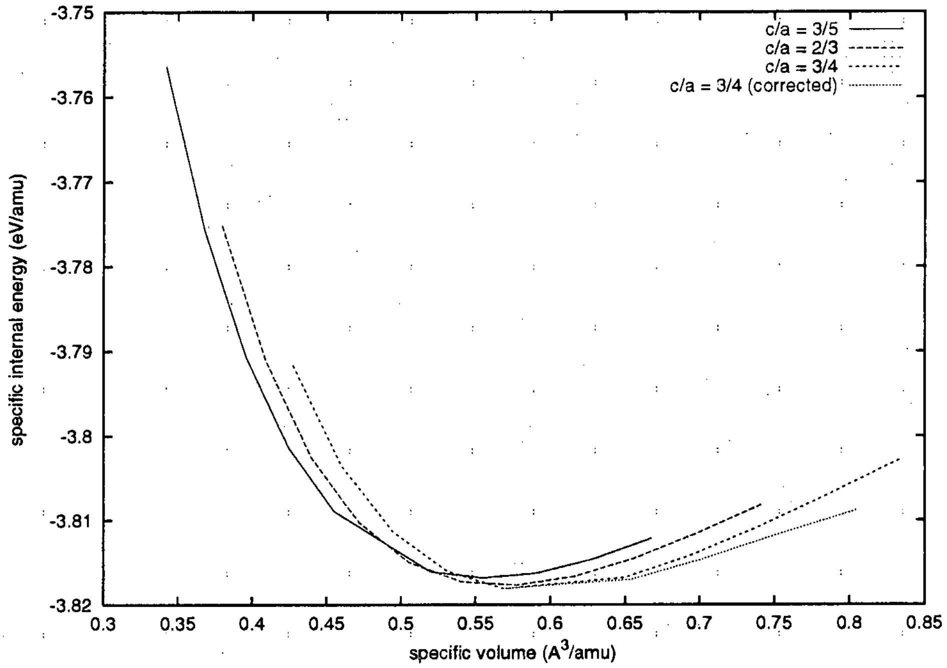

Figure 3.28. Cold curves for BCT Si, with a constant c/a ratio.

timestep, damping factor and effective mass did not provide a realistic solution – a lengthy series of these rather time-consuming calculations appeared necessary to find suitable values for these pseudo-physical parameters. It was found more efficient to simply perform a series of static box calculations covering the expected range of c for a given a.

Exploring discrete (a, c) states

The most reliable method used to eliminate the internal degree of freedom was simply to hold a constant and perform a series of calculations with different values of c, until the structure of isotropic stress was bracketed or determined to a suitable accuracy.

Bracketing was used in most cases. The value of c for which the stresses were isotropic was estimated by linear and quadratic interpolation. Linear interpolation was found to give the same result as quadratic (to ~0.01% in pressure) so long as the values of c chosen bracketed the case of isotropic stress to within a few %.

There was a little difference between the ground state energies obtained with a cutoff of 400 eV compared with 300 eV, though very small in comparison with the range of the data. The difference in stress was also small. The difference in ground state energies between calculations with different values for the cutoff was roughly constant with compression, suggesting that the difference in wavefunctions occurred close to the atomic nuclei in low-energy states that were not affected significantly by changes in the interatomic spacing. As one would expect, the ground state energy was slightly lower with a larger cutoff.

Summary of equilibrium states

Combining the equilibrium states predicted by the various methods described above, the resulting energy – volume curve appears reasonably smooth (Fig. 3.29). This indicates that energies calculated using the different methods are consistent with one another.

The variation of c/a with v is much more sensitive to the method of calculation (Fig. 3.30) and looks less smooth. Around the energy minimum, the elastic constants for the BCT structure are quite small and may have different signs. The uncertainty in c/a reflects the sloppiness of the structure in this regime.

Stress and energy derivatives

The equation of state based on these ground state calculations requires the pressure to be calculated. At $T = 0$, this can be either $\bar{p}$ as determined from the stress or $-de/dv$. Both methods are equivalent in principle, but are expected to be the same to some finite accuracy in numerical work.

To investigate the accuracy, the stress σ calculated by the alternative methods described above was compared with estimates of de/dv. To obtain de/dv, groups of adjacent (v, e) ground state points were fitted with polynomials,

$$e \simeq P_n(v). \tag{3.10}$$

These piecewise continuous fits were then differentiated to give de/dv.

Linear fits systematically over-predicted de/dv at smaller volumes, compared with smoother functional forms. This is what one would expect for a bulk modulus which increases monotonically with pressure. Nearer the energy minimum, the linear fits exhibited significant oscillations to either side of σ. These may be connected with the way the results from different methods of calculating the

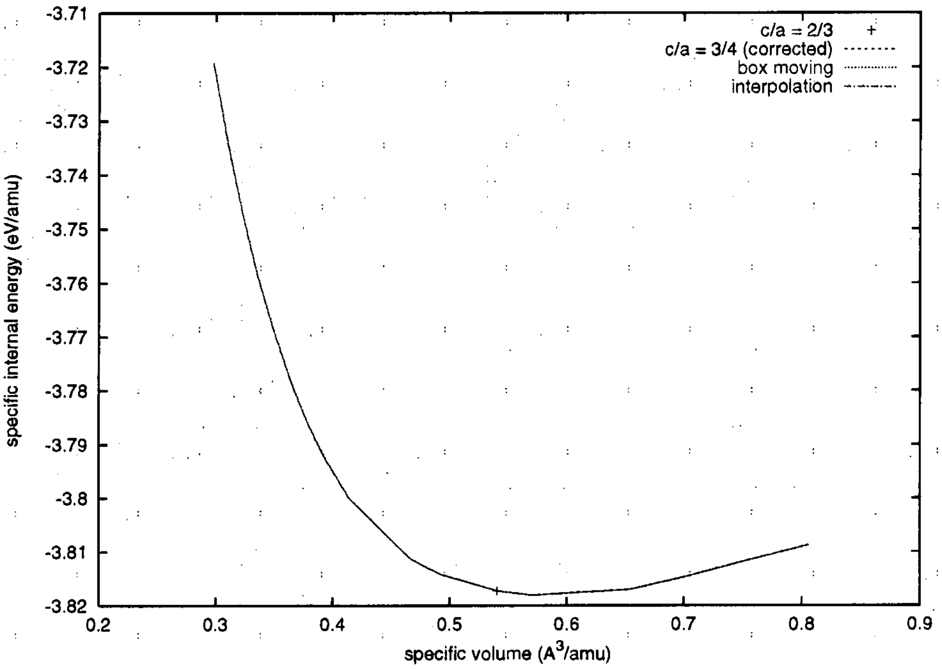

Figure 3.29. Summary of Si BCT ground state energies. (Slightly different methods were used to find the equilibrium c/a ratio in different compression regimes; this graph summarises the results from all the methods.)

ground state were combined. The result from linear fits was sensitive to the number of points included in the fit (2, 3 or 4). The derivatives from quadratic and cubic fits were much closer to σ, and were less sensitive to the number of points included in each group. At high volumes, the results from all the polynomial fits fell below σ. This demonstrates the finite accuracy of the procedure used to remove the residual components of the stress deviator. (Fig. 3.31.)

It can be concluded that quadratic or cubic polynomials are an accurate way of calculating the pressure from the energy in constructing the equation of state. Linear fits should be avoided for the Si BCT table as they generally under-predict the pressure and exhibit some unphysical oscillations in the Si BCT cold curve. The ground state calculations are close enough together in v to enable polynomial fits to be used with reasonable accuracy.

These results also indicate that the CASTEP calculation of stress is consistent with the ground state energies.

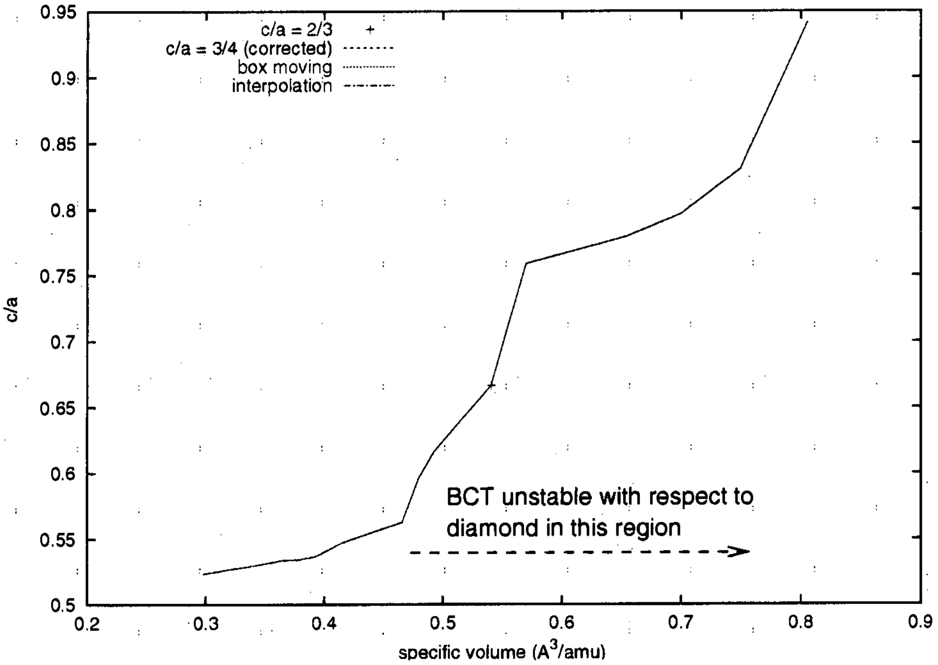

Figure 3.30. Predicted variation of c/a with specific volume v for BCT Si.

Grüneisen model of the lattice-thermal contribution

The BCT structure has a lower symmetry than cubic, and required a large computational effort to predict the phonons. The lattice-thermal contribution was also estimated by a Grüneisen model, in order to investigate the accuracy with which a phase boundary might be predicted without accurate phonons.

A couple of variants of the basic BCT equation were tried, in order to investigate the sensitivity to the thermal model and the correction previously used to improve the accuracy of a single-phase equation of state. The parameter varied was the effective number of degrees of freedom in the atomic motion. If all vibrational and translational degrees of freedom are excited then the lattice contribution to the specific heat capacity would be $3\,k_B$ (6 degrees of freedom) per atom. This is the maximum possible value, and may be reduced if phonon modes are quenched at temperatures which are too low to cause excitations. In order to investigate the sensitivity to this effect, EOS were generated for the BCT phase with 6 and 3 degrees of freedom.

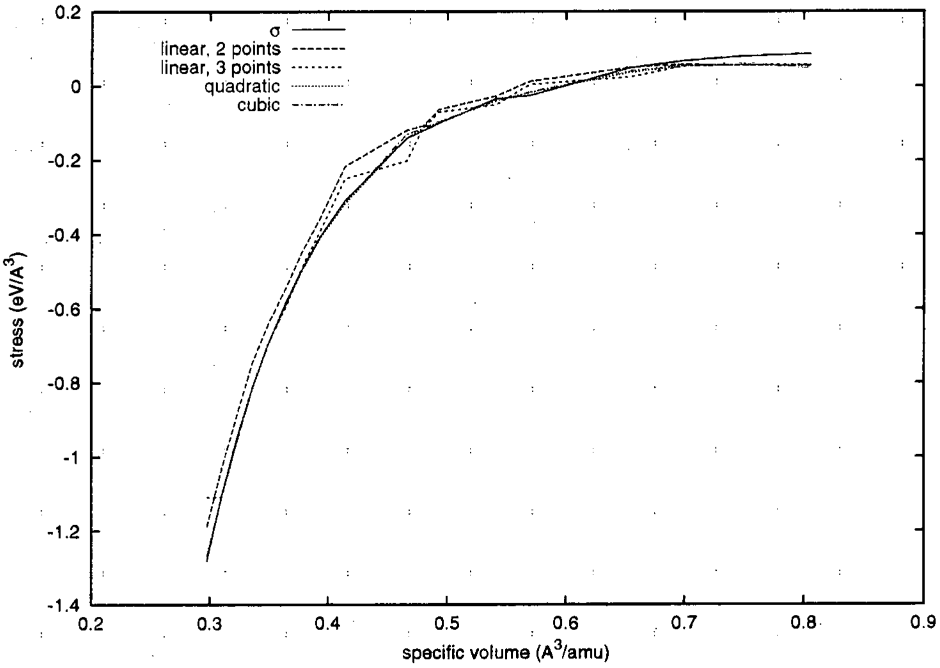

Figure 3.31. de/dv for BCT Si from σ and polynomial fits to $e(v)$.

Ab initio phonons

Ab initio phonons were calculated by estimating the stiffness of the lattice by performing electron ground state calculations with atoms displaced from their equilibrium positions, as described in Sections 2.3 and 2.6. To obtain the BCT forces, two atoms were displaced in turn, parallel with each of the lattice vectors. The BCT calculations required far more computer time than the diamond, so the lattice-thermal contribution was generated from restoring forces obtained at only three densities. The density of phonon states was estimated at intermediate densities by interpolating the frequency of each phonon using a Grüneisen-like variation [40],

$$\omega \propto \rho^{1/3}. \qquad (3.11)$$

Equations of state

EOS were generated by combining the cold curve at isotropic stress with the electron-thermal energy and the atom-thermal energy, as described in Section 2.6. A separate EOS was generated for each of the Grüneisen variants and for the *ab initio* phonons. The same energy tilt was applied to the *ab initio* EOS as was applied for the diamond structure, giving a consistent *ab fere initio* EOS across the two phases.

3.2.3 Diamond/BCT equation of state

The diamond and BCT EOS were analysed to predict the equilibrium phase diagram, as described in Section 2.6. The *ab initio* EOS with no adjustment to match the STP density was not considered.

The *ab fere initio* EOS over-predicted the pressure of the diamond/BCT transition at low temperatures, compared with previous estimates [78]. In order to allow a more meaningful comparison with experiment, a modified version of the quasiharmonic EOS was generated with an additional energy offset in the BCT phase to bring the transition pressure into better agreement.

Phase diagram

The phase diagram of silicon has been investigated previously by static press experiments, shock wave studies and various calculational schemes [78]. The uncertainty in the data is ~3 GPa, reflecting hysteresis in the experimental results (transition pressure on release compared with compression) and variation between different calculational techniques.

With a constant lattice specific heat capacity of $3k_B$ per atom (6 degrees of freedom) in the BCT phase, the diamond/BCT boundary was significantly steeper than the accepted result and tilted the opposite way with increasing temperature. Using a Grüneisen equation of state with a lattice specific heat capacity of $1.5k_B$ per atom (3 degrees of freedom), the slope of phase boundary moved towards the accepted value, though the tilt was still in the opposite sense. The transition pressure at $T = 0$ appeared to be greater than the expected value in all cases. The slope obtained with quasiharmonic phonons was significantly more accurate, and the additional energy offset applied to the BCT EOS brought the *ab fere initio* EOS into closer agreement. The effect of including the electron-thermal contribution was almost negligible. (Fig. 3.32.)

Shock Hugoniot

The shock Hugoniot was calculated for each equation of state. The calculated Hugoniot curves were compared with published experimental data for single

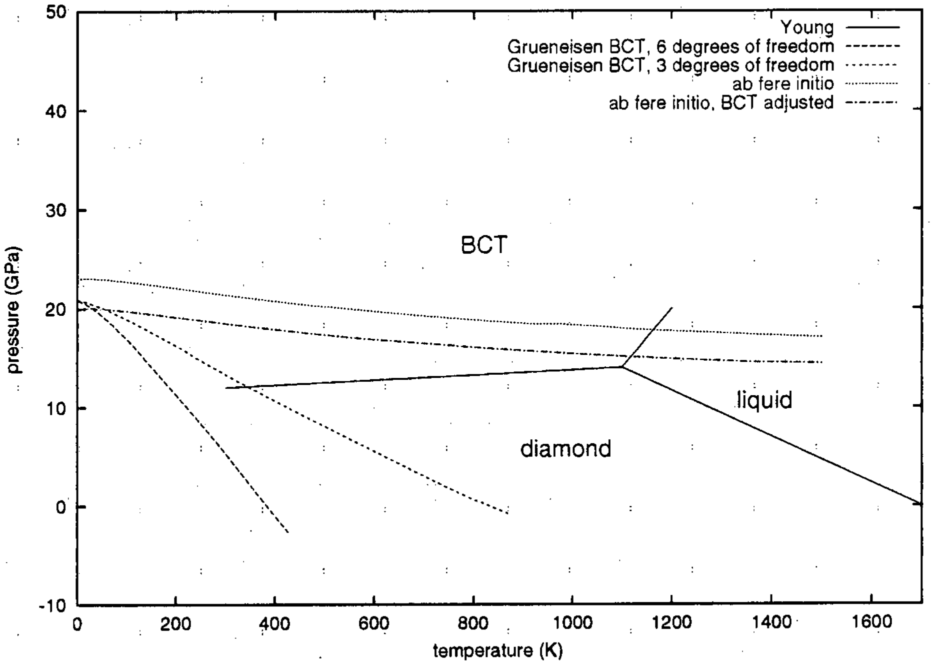

Figure 3.32. Predicted diamond/BCT phase boundaries compared with accepted results for silicon.
All lines are diamond/BCT phase boundaries, except for the additional steeper lines showing the diamond/liquid and BCT/liquid phase boundaries from Young [78].

crystals of silicon [79].

For the Grüneisen EOS, the variations in the lattice heat capacity had little effect. The equilibrium equations of state based on the Grüneisen models of the BCT lattice-thermal energy matched the data quite well at pressures up to ~100 GPa. The *ab fere initio* EOS matched the experimental data equally well, but the *ab fere initio* EOS for which the BCT part was adjusted to match the phase diagram more accurately was found to match the Hugoniot data slightly less accurately. This discrepancy suggests that there might be some uncertainty in either the phase boundary data or the Hugoniot data. (Fig. 3.33.)

The deviation of the theoretical equations of state above ~100 GPa could be caused by another phase transition. Silicon is reported to change to the FCC structure in this region [78]. (In fact, silicon has other phase changes at intermediate pressures, but these have similar densities and compressibilities to the BCT structure.) At ~80 GPa, the compression is great enough for doubts

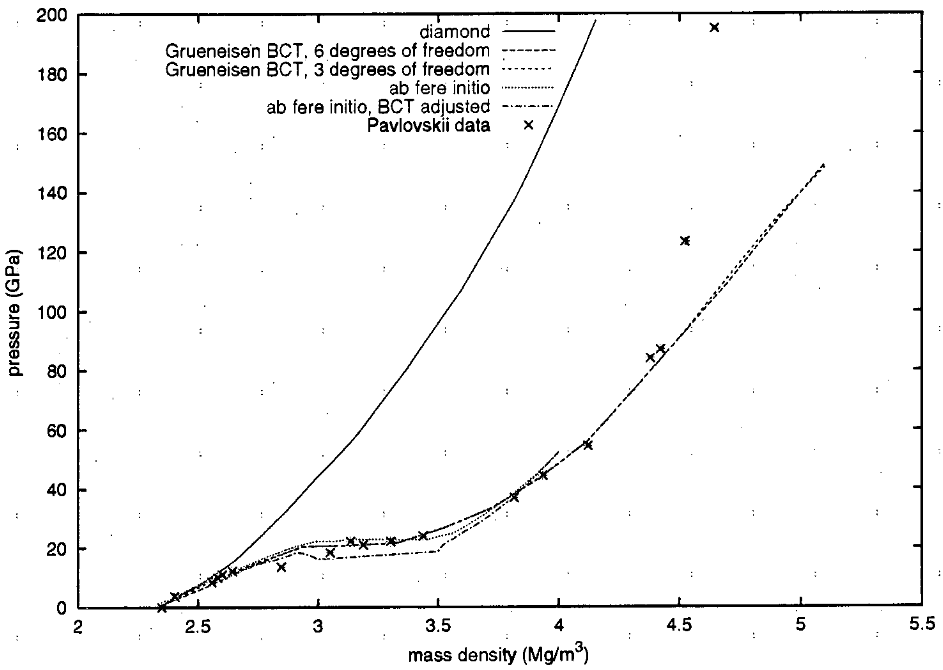

Figure 3.33. Hugoniot curves (pressure – density plane) compared with experimental data.

to be raised about the pseudopotential treatment used; it is possible that outer electrons would be forced into the core region at compressions of this order.

3.3 Beryllium

Calculations were performed of beryllium in the HCP (STP [40]), BCC and FCC phases. HCP and BCC are observed experimentally to be the equilibrium phases below melt, up to at least 6 GPa [78]. *Ab initio* studies have been performed previously at a higher pressure range [80].

The simulations used CASTEP V5.1.2. The pseudopotential was an *ab initio* form defined in reciprocal space, modelling the inner 2 electrons. In all cases, the code was allowed to perform 45 iterations towards a converged ground state. The convergence history of the ground state energy was checked; after this number of iterations it barely altered between iterations.

3.3.1 HCP

At STP, Be has the HCP structure, with $c/a = 1.567$ and $a = 2.29$ Å [40], or $c/a = 1.56$ [78]. The atomic mass of Be is 9.012180 amu.

The HCP structure can be described in several ways, choosing the a and b directions 60° or 120° apart in a 2-atom lattice cell, or defining an orthorhombic cell containing 4 atoms (see Appendix B).

Example calculations were performed using each description in order to investigate the sensitivity of the ground states deduced to the model. This largely tests the stability of the computer program, but also provides an indication of the importance of finite size effects in that the orthorhombic cell can contain electron modes of longer wavelength.

The K290 symmetry program did not recognise the full symmetry of the hexagonal lattice, so the CASTEP calculations were performed with more freedom than the desired structure warranted. A perfect calculation should produce the same result when assuming lower symmetry, but unnecessary work is done in calculating duplicated data. The orthorhombic model also includes an extra degree of freedom, since a and b are specified separately.

The models were found to give the same variation of energy with density, though with a slightly different absolute offset. The components of stress in the x and y directions were quite close for the orthorhombic model, but significantly less so for the hexagonal. For this reason, the orthorhombic model was used as standard for generating the EOS.

Ground states were calculated using the orthorhombic model for a constant

c/a ratio of 1.567. $10 \times 10 \times 10$ k-points were used, with a plane wave cutoff of 750 eV. The stresses deduced were not isotropic, indicating that the c/a ratio for compression at isotropic stress is not constant (Fig. 3.34). The stresses in the x and y directions were very similar, indicating that the orthorhombic calculation reproduced the hexagonal symmetry to an acceptable accuracy.

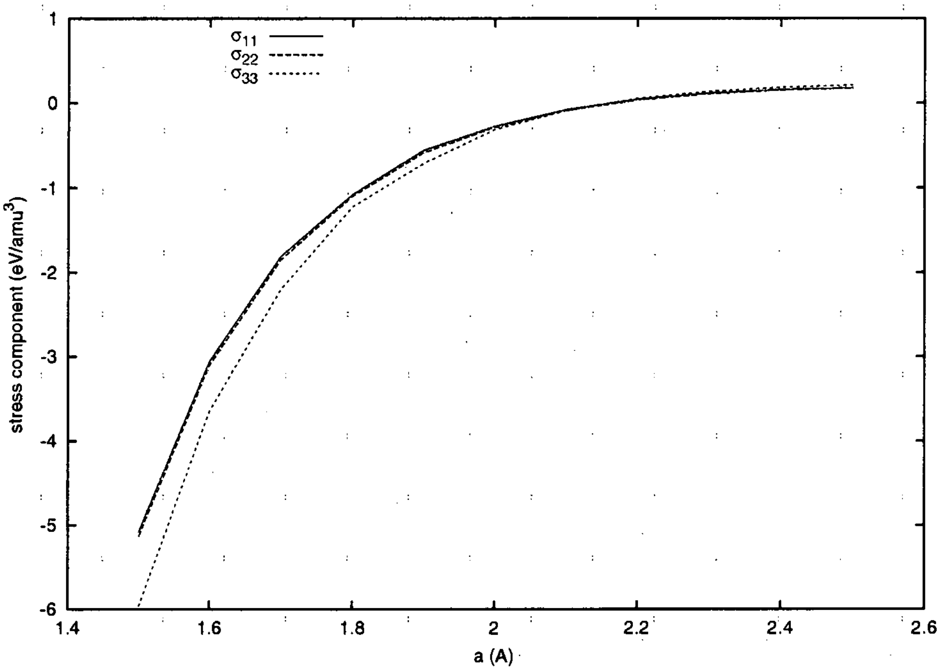

Figure 3.34. Predicted stresses on the cold curve for HCP Be with $c/a = 1.567$.

For each of a representative set of values of a, the ground state stress was calculated as c was varied. In this way, the variation of c with a for isotropic stress was estimated (Fig. 3.35). The calculated variation was in good agreement with the observed c/a ratio. The c/a ratio was predicted to reach the ideal value (1.63 [40]) under compression. Although the difference in ground state energy with a was significant (Fig. 3.36), the variation of ground state energy with mass density was much smaller (Fig. 3.37). This suggests that it may not be necessary to find the state of isotropic stress in order to calculate the frozen-ion cold curve.

Electron band structure

The electron band structure was estimated by collecting the band energies calculated using CASTEP at each weighted $\vec{k}$ point, as described in Sections 2.4

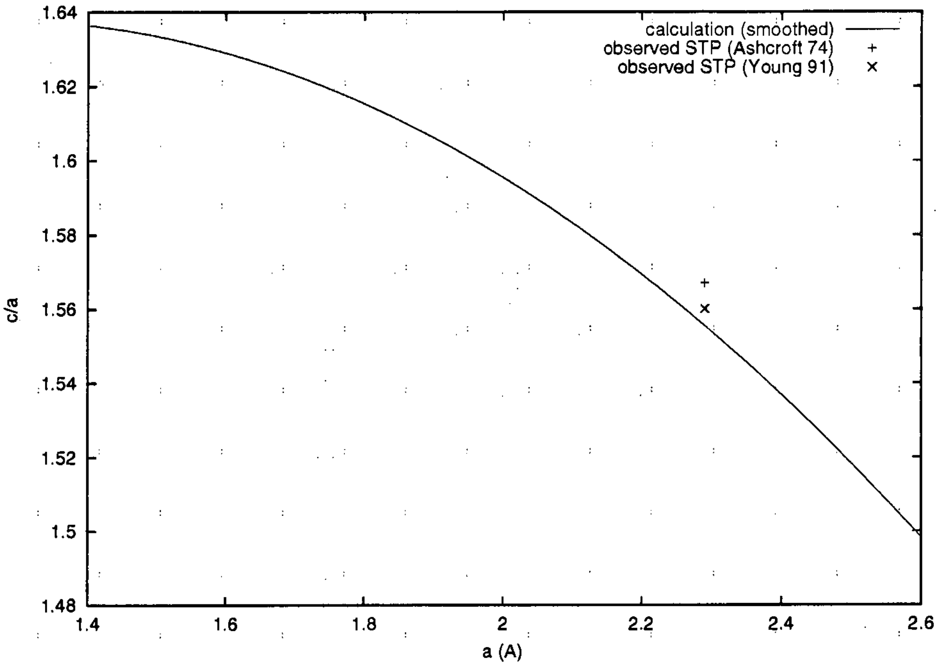

Figure 3.35. Predicted variation of c/a with a for HCP Be.

and 2.6. The density of levels exhibited two distinct peaks, quite unlike the free electron-like structure for aluminium (Fig. 3.38). This result is consistent with previous predictions using explicit orbitals [81], though with a lower resolution. The Fermi energy fell close to the minimum between the peaks. The chemical potential and electron-thermal energy should thus have a more complicated dependence on temperature than for a simple metal.

The band structure at each density was used to predict the chemical potential and hence the electron-thermal energy (Figs 3.39 and 3.40).

Phonon modes

Ab initio phonon modes were deduced by displacing the atom at 0,0,0 and using the restoring forces predicted from the ground states in CASTEP, as described in Sections 2.3 and 2.6.

It was sufficient to calculate forces from two perturbations: the x-direction and the z-direction. Atoms in the same $x-y$ plane are related to each other by translational symmetry. Adjacent atoms in different $x-y$ planes are related by translation and inversion.

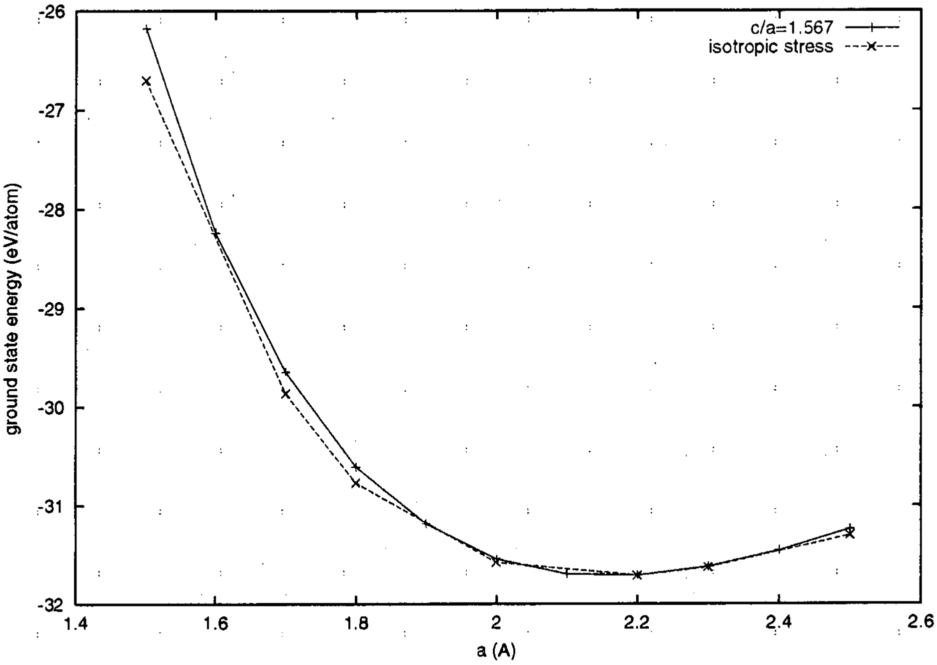

Figure 3.36. Predicted variation of ground state energy with a for HCP Be.

With a varying c/a ratio, the lattice no longer has rotational symmetry about the $\langle 11\bar{1}\bar{3}\rangle$ direction through lattice points, but it still has 3-fold symmetry about the $\langle 001\bar{1}\rangle$ direction through any atom. The 3-fold symmetry can be used to convert forces for a perturbation in the x-direction (parallel with the $\vec{u}$ lattice vector) to a perturbation parallel to the $\vec{v}$ lattice vector. Here, a orthorhombic lattice cell was used, for which the x perturbation cannot be mapped directly into the y-direction by symmetry. However, the effect of a perturbation in the y-direction can be calculated quite accurately by assuming that the restoring forces $\vec{f}$ from a sum of displacement vectors $\delta\vec{r}_i$ is the same as the sum of the forces from each separate displacement:

$$\vec{f}\left(\sum_i \delta\vec{r}_i\right) \simeq \sum_i \vec{f}(\delta\vec{r}_i). \qquad (3.12)$$

This is the same as taking the first term in a Taylor expansion for $\vec{f}$ with respect to each $\delta\vec{r}_i$, and is true for small $\delta\vec{r}_i$ – which is also the condition for quasiharmonic phonons. Then, if we calculate $\vec{f}(\delta\vec{r}_x) = \vec{f}(\delta\vec{r}_u)$ (the forces from a displacement in the x or $\vec{u}$ direction) and, by symmetry, $\vec{f}(\delta\vec{r}_v)$ (the forces from a displacement

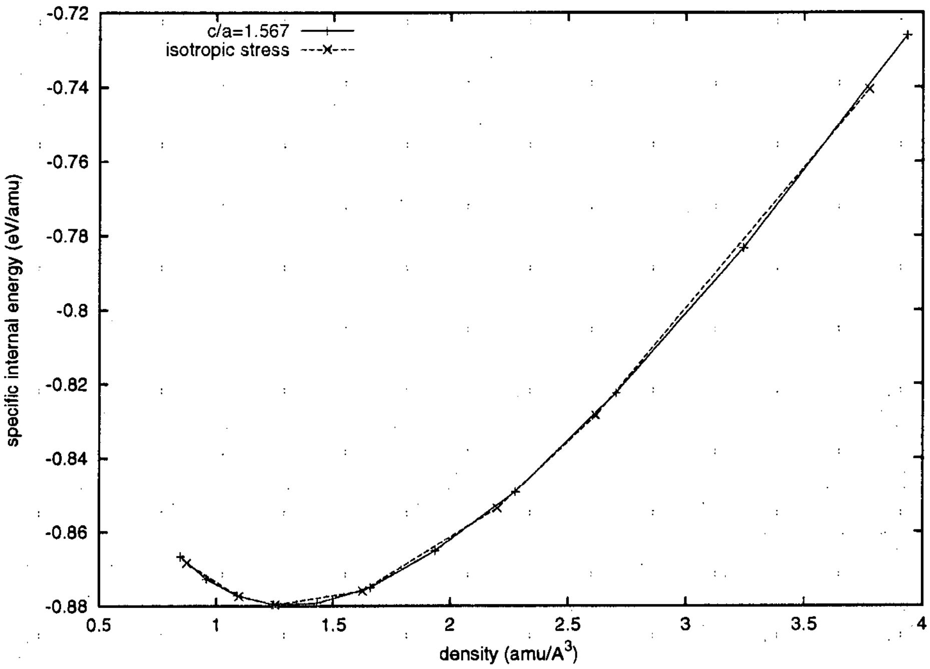

Figure 3.37. Predicted variation of ground state energy with mass density for HCP Be.

in the $\vec{v}$ direction), then we can calculate $\vec{f}(\delta\vec{r}_y)$ from

$$\vec{f}(\delta\vec{r}_y) \simeq \alpha_u \vec{f}(\delta\vec{r}_u) + \alpha_v \vec{f}(\delta\vec{r}_v) \quad : \quad \vec{r}_y = \alpha_u \vec{r}_u + \alpha_v \vec{r}_v. \tag{3.13}$$

Choosing $|\vec{r}_y| = |\vec{r}_u| = |\vec{r}_v|$, we find that $\alpha_u = -1/\sqrt{3}$ and $\alpha_v = 2/\sqrt{3}$.

3.3.2 BCC

Under compression and heating, Be changes to the body-centred cubic (BCC) phase [78]. Predictions were made of the frozen-ion cold curve, representing the BCC structure by a cubic lattice cell with atoms at 0,0,0 and 1/2, 1/2, 1/2. The range of the lattice parameter a was chosen to cover the density at STP.

The cold curve was calculated with $5 \times 5 \times 5$ or $10 \times 10 \times 10$ k-points over the Brillouin zone and a plane wave cutoff of 500 or 750 eV, according to the procedure described in Section 2.6. The results were very similar at specific volumes less than about 0.6 Å^3/amu, but varied significantly at greater volumes. The energies obtained with $5 \times 5 \times 5$ k-points exhibited undulations around the

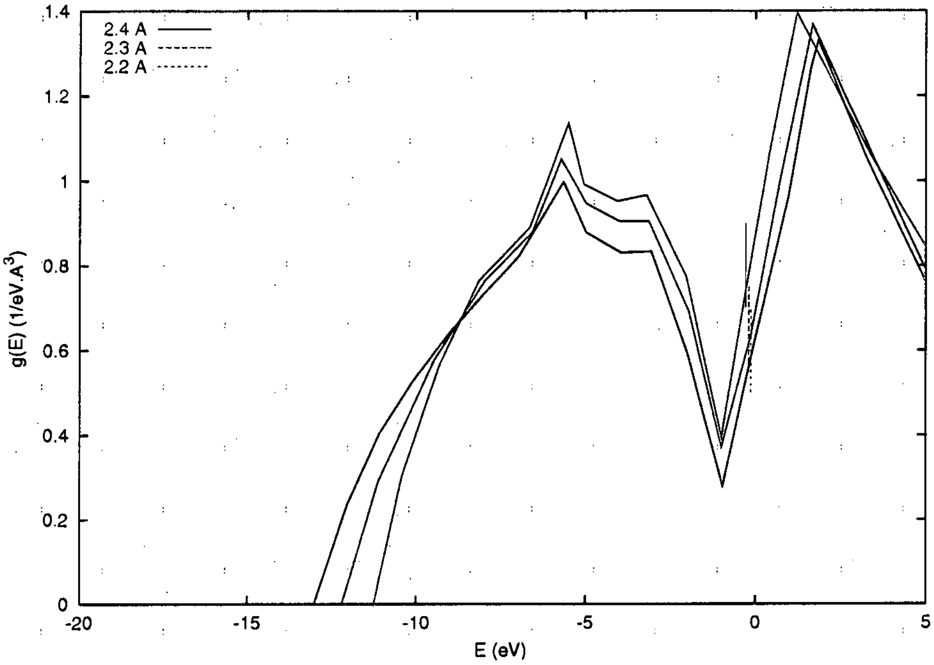

Figure 3.38. Density of electron energy levels calculated for HCP Be. Vertical bars show the Fermi energy for each density of levels.
(Note: the ragged shape reflects the relatively coarse resolution of k-points.)

equilibrium volume. The undulations were reduced by including the Pulay stress. Interestingly, the stress predictions did not show such a pronounced variation. (Figs 3.41 and 3.42.)

Electron band structure

The electron band structure was estimated by collecting the band energies calculated using CASTEP at each weighted $\vec{k}$ point as described in Sections 2.4 and 2.6. As for the HCP structure, the density of levels exhibited two distinct peaks. The Fermi energy fell close to the minimum between the peaks. (Fig. 3.43). The band structure at each density was used to predict the chemical potential (Fig. 3.44) and hence the electron-thermal energy (Figs 3.45 and 3.46).

Phonon modes

Ab initio phonon modes were deduced by displacing the atom at 0,0,0 and using the restoring forces predicted from the ground states in CASTEP, as described

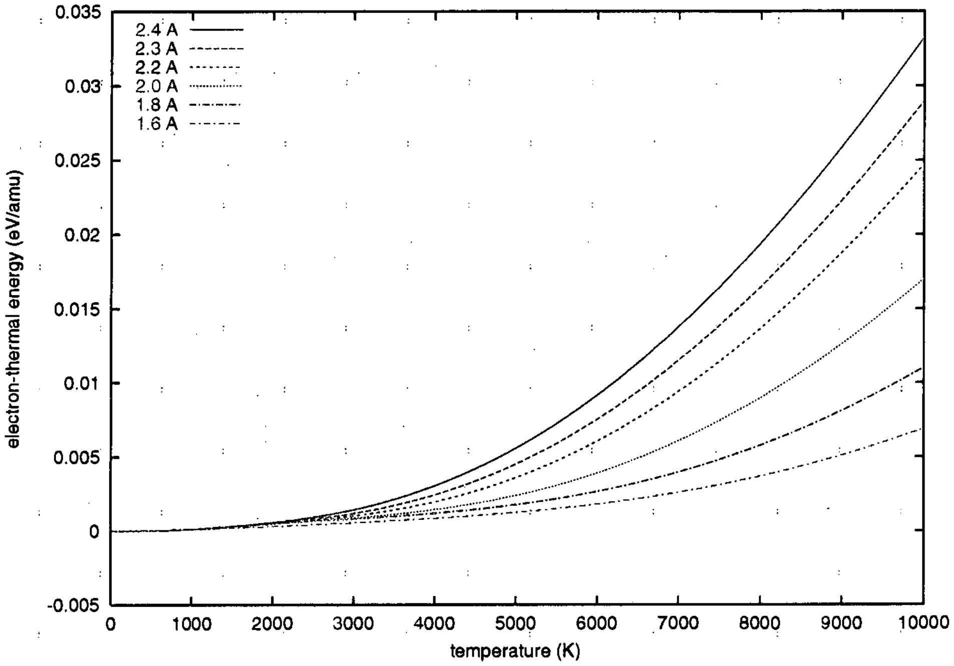

Figure 3.39. Electron-thermal energy from band structure for HCP Be.

in Sections 2.3 and 2.6.

It was sufficient to calculate forces from a single perturbation in the x-direction. Atoms at $0, 0, 0$ and $\frac{1}{2}, \frac{1}{2}, \frac{1}{2}$ are related to each other by translational symmetry.

3.3.3 FCC

The frozen-ion cold curve for the FCC structure was calculated in order to verify that the code and pseudopotential reproduced the experimental observation that FCC is *not* the stable structure. These calculations used a plane wave cutoff of 750 eV, and $10 \times 10 \times 10$ points in k-space, and were performed according to the method described in Section 2.6.

3.3.4 Equilibrium phase diagram

Considering the frozen-ion cold curves, the first principles calculations predicted that HCP is the stable structure around $p = 0$ (minimum in the energy curves), in agreement with observations. The curves became very close under compression, and suggest that BCC and then FCC may become stable at a few hundred GPa.

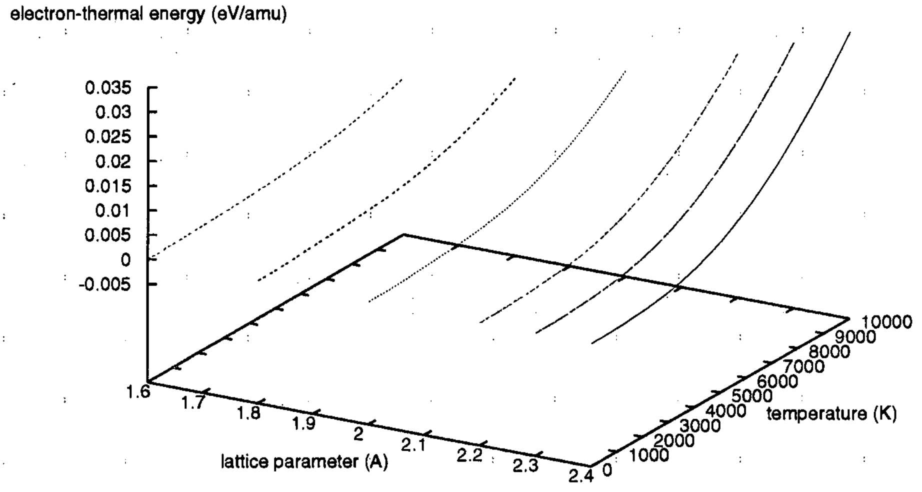

Figure 3.40. Electron-thermal energy from band structure for HCP Be.

(Fig. 3.47.)

Using the thermodynamically complete EOS for the HCP and BCC phases, the classical phase boundary was predicted. This was in reasonable agreement with results from other workers [78]. The phase boundary was found to be extremely sensitive to all contributions to the EOS. Great care was required in order to generate a phase boundary that agreed to this accuracy. In particular, the phonons had to be calculated from forces over a finely-spaced density mesh, in comparison with the mesh that had been found necessary in order to predict reasonably accurate shock Hugoniots. Electron-thermal effects made a negligible contribution to the phase boundary. (Fig. 3.48.)

3.3.5 Density at STP

The *ab initio* EOS predicted an STP density of 1.682 Mg/m^3, compared with ~1.84 Mg/m^3 from experiment [40, 78]. The density for p =1 atm at T = 0 K was ~1.69 to 1.70 Mg/m^3, using linear and quadratic interpolation respectively – about 8% higher than the observed value.

For 273 K at ~1.84 Mg/m^3, the *ab initio* EOS gave a pressure ~5.58 GPa using

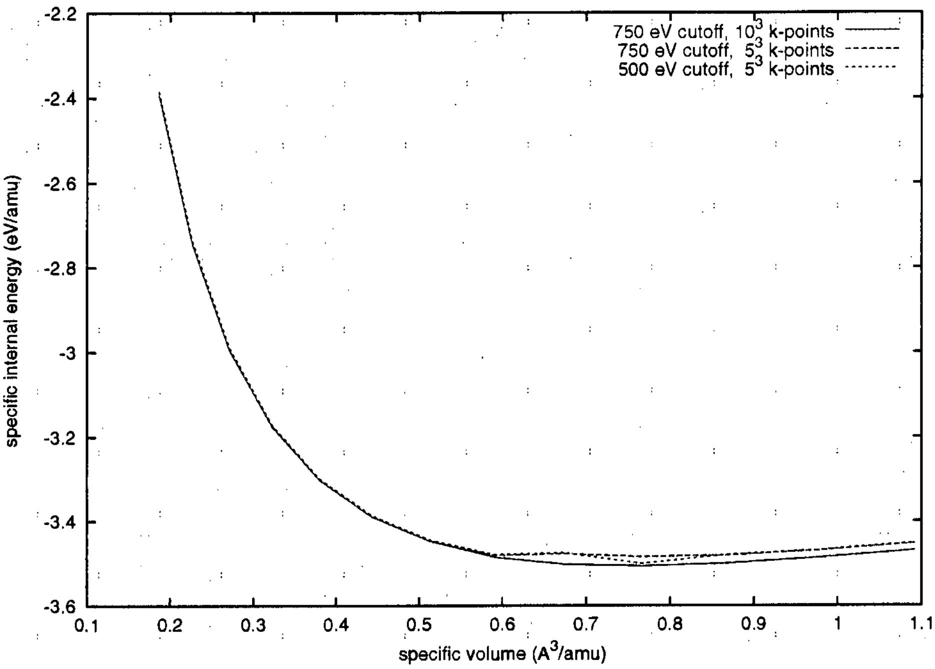

Figure 3.41. Sensitivity of ground state energy to plane wave cutoff and k-point density for BCC Be.

linear interpolation, or 5.24 GPa using quadratic interpolation. An *ab fere initio* EOS was generated by tilting the specific internal energy by this pressure. The STP density was then 1.847 Mg/m^3, and the density at 0 K became 1.853 Mg/m^3.

3.3.6 Shock Hugoniot

The shock Hugoniot was predicted using the *ab initio* and *ab fere initio* EOS. The HCP/BCC transition made no significant difference to the shock Hugoniot in the range of pressures considered.

The hydrodynamic shock Hugoniot was in poorer agreement with the experimental data [72] than for aluminium or silicon. This was probably because beryllium has a relatively high shear modulus – ~150 GPa, compared with ~27 GPa for aluminium [70]. The longitudinal sound speed c_L in an elastic material is given by

$$\rho c_L^2 = K + \frac{2}{3}\mu, \tag{3.14}$$

where K is the bulk modulus and μ the shear modulus [76]. The effect of shear

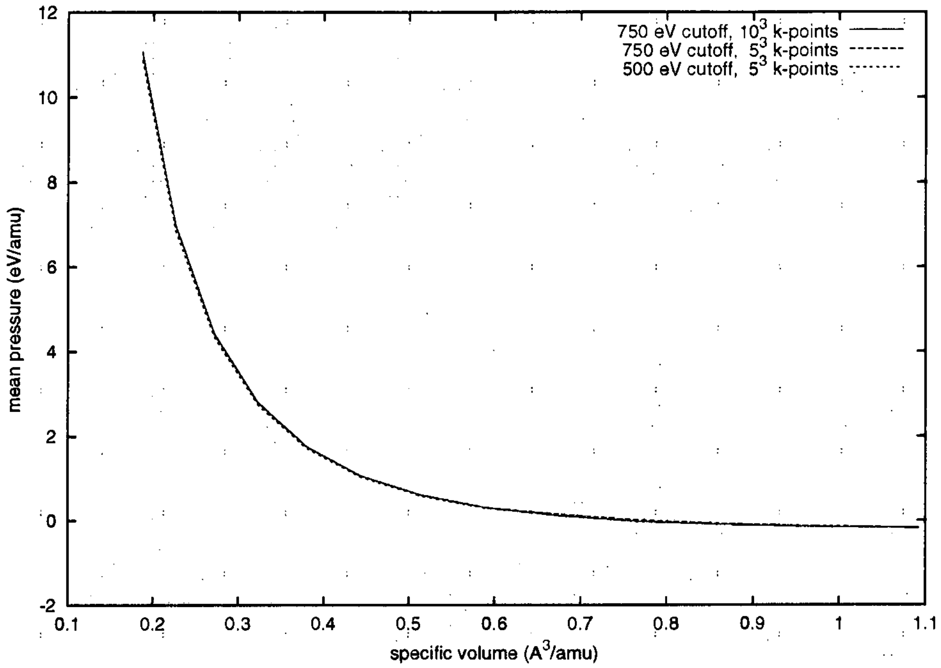

Figure 3.42. Predicted stresses on the cold curve for BCC Be.

modulus on the shock Hugoniot was estimated by including its contribution, scaled by $1 - p/Y$ where p is the hydrodynamic pressure on the Hugoniot and Y the yield stress. The resulting Hugoniot was significantly closer to the experimental data. The predicted Hugoniots exhibited slight undulations; these were caused by the relatively coarse tabulation in mass density. (Figs 3.49 and 3.50.)

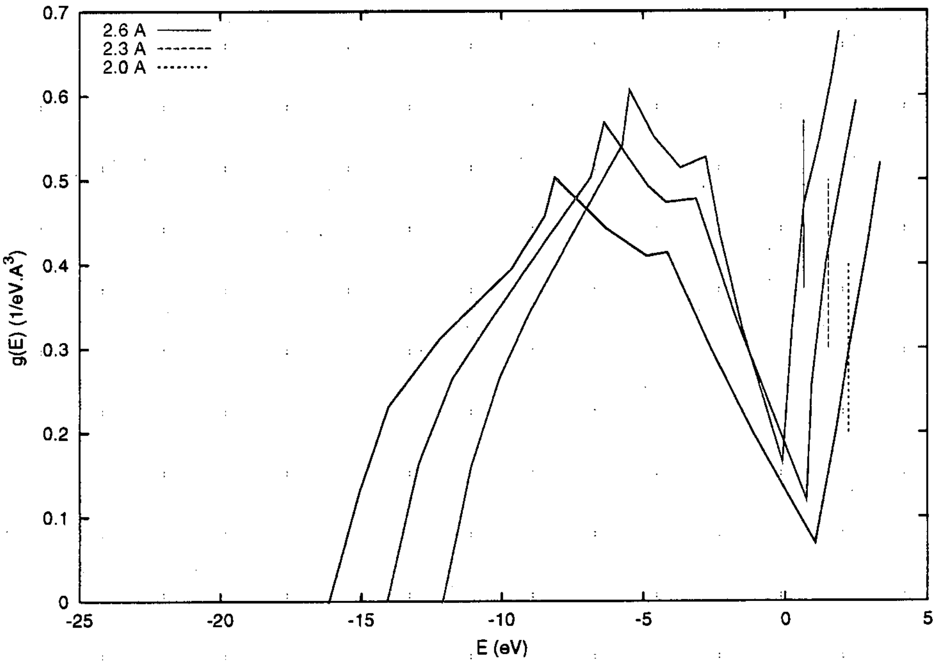

Figure 3.43. Density of electron energy levels calculated for BCC Be. Vertical bars show the Fermi energy for each density of levels.
(Note: the ragged shape reflects the relatively coarse resolution of k-points.)

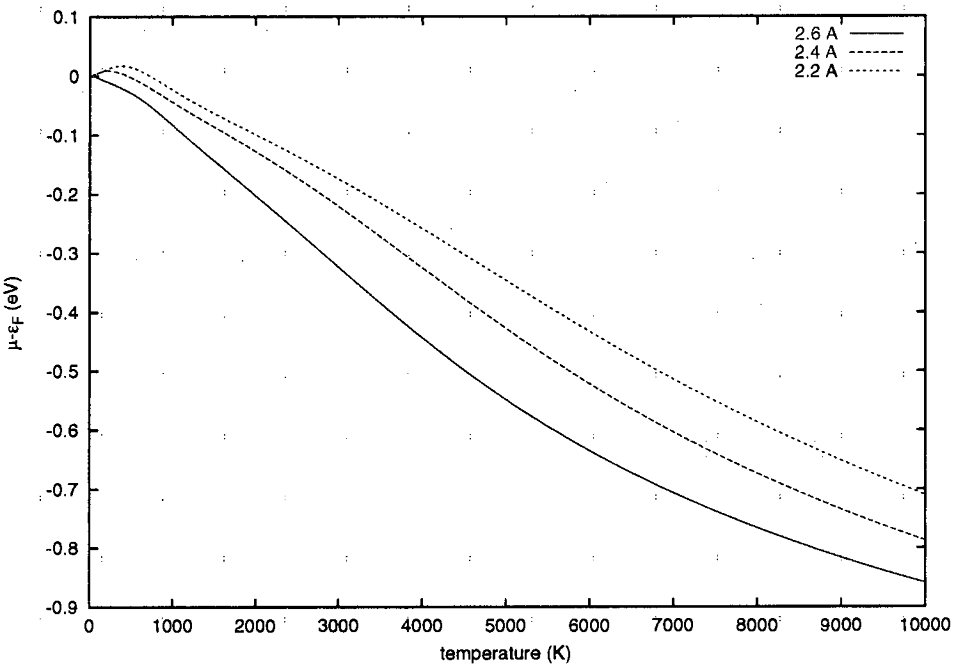

Figure 3.44. Difference between chemical potential and Fermi energy (chemical potential at $T = 0$) for BCC Be.

Note: the chemical potential near $T = 0$ is calculated with an uncertainty of the order of the spacing between electron energy levels.

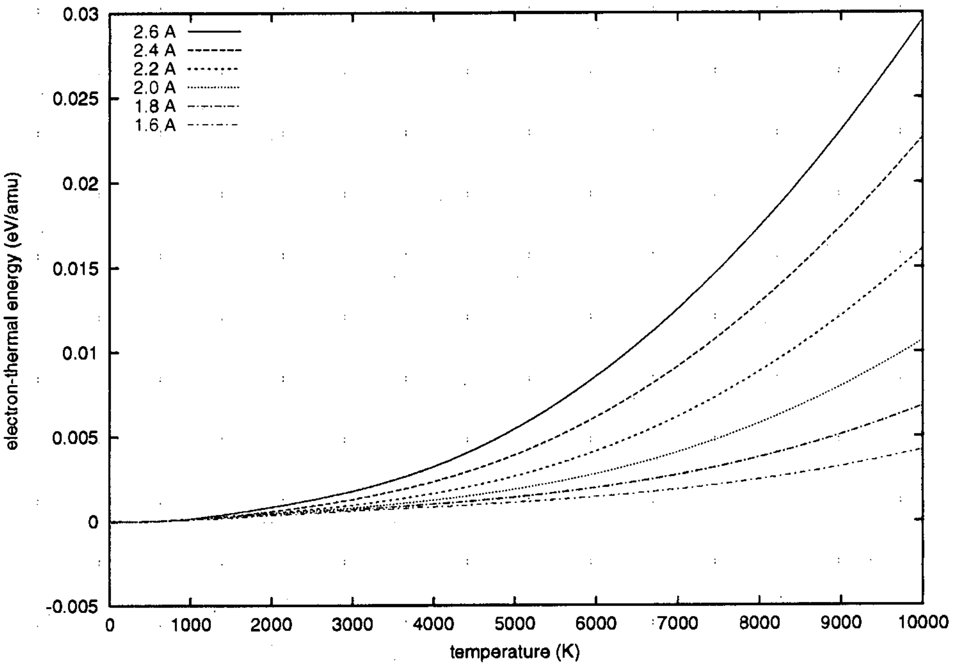

Figure 3.45. Electron-thermal energy from band structure for BCC Be.

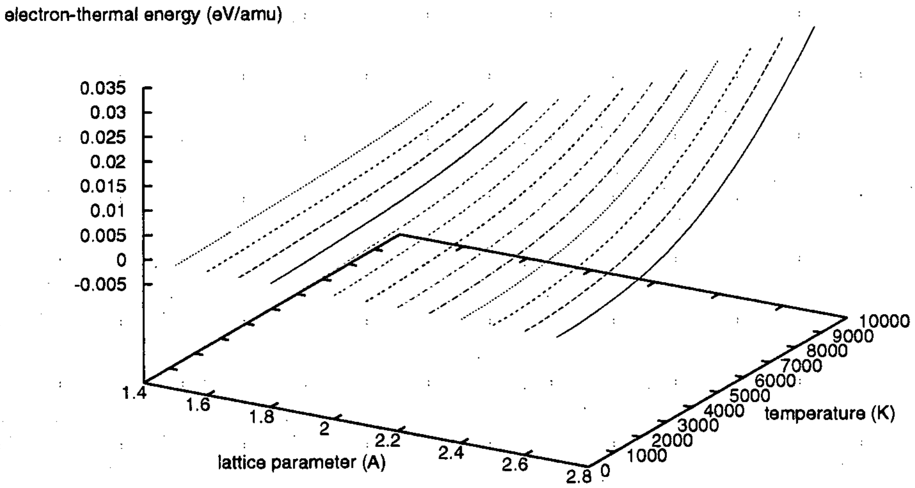

Figure 3.46. Electron-thermal energy from band structure for BCC Be.

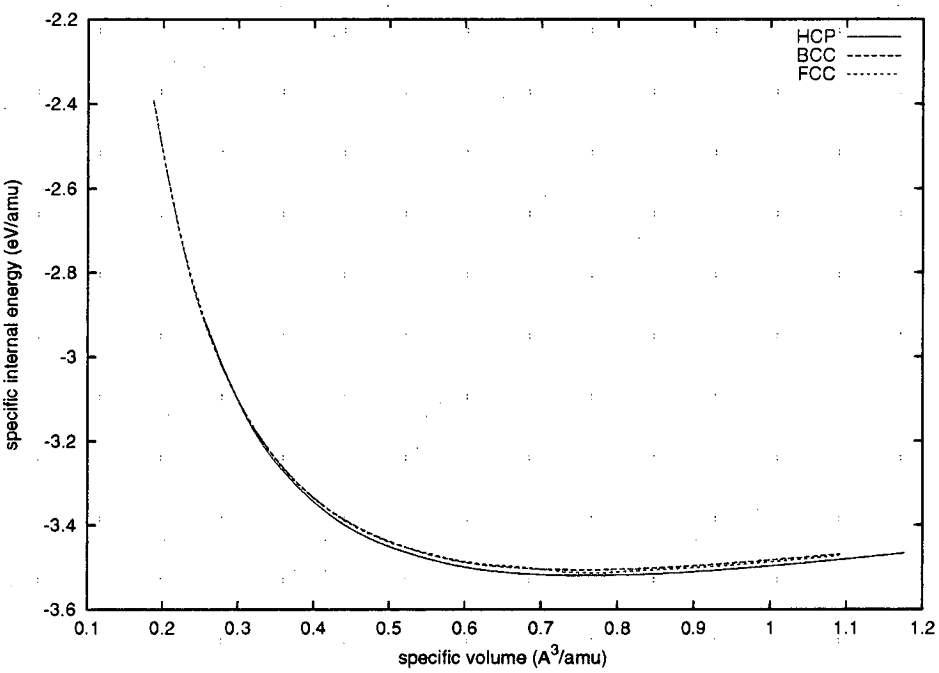

Figure 3.47. Comparison between cold curve energies for HCP, BCC and FCC Be.

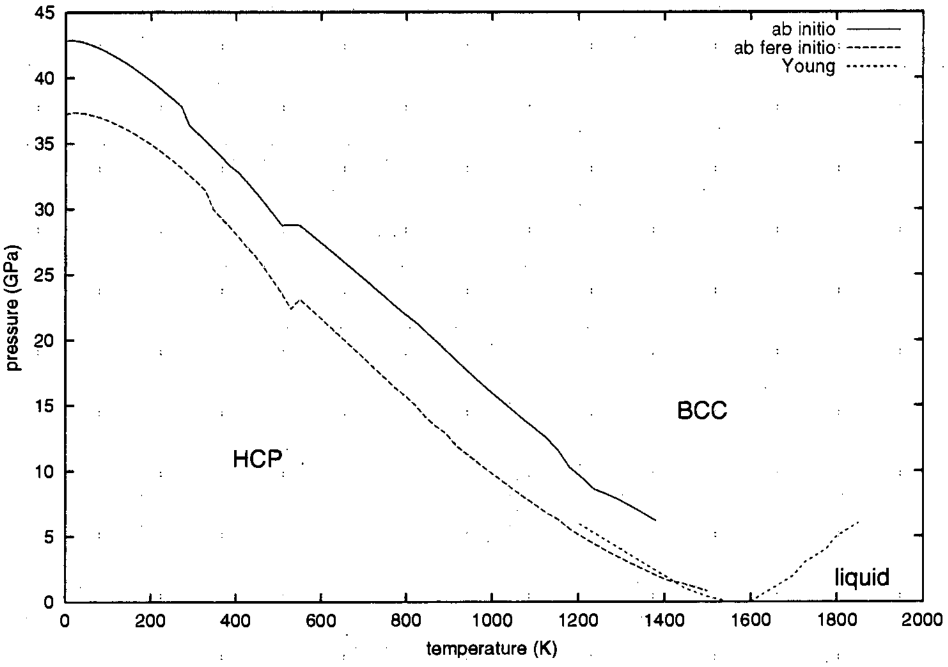

Figure 3.48. Phase diagram for Be.
All lines are HCP/BCC phase boundaries, except for the additional line showing the BCC/liquid phase boundaries from Young [78].

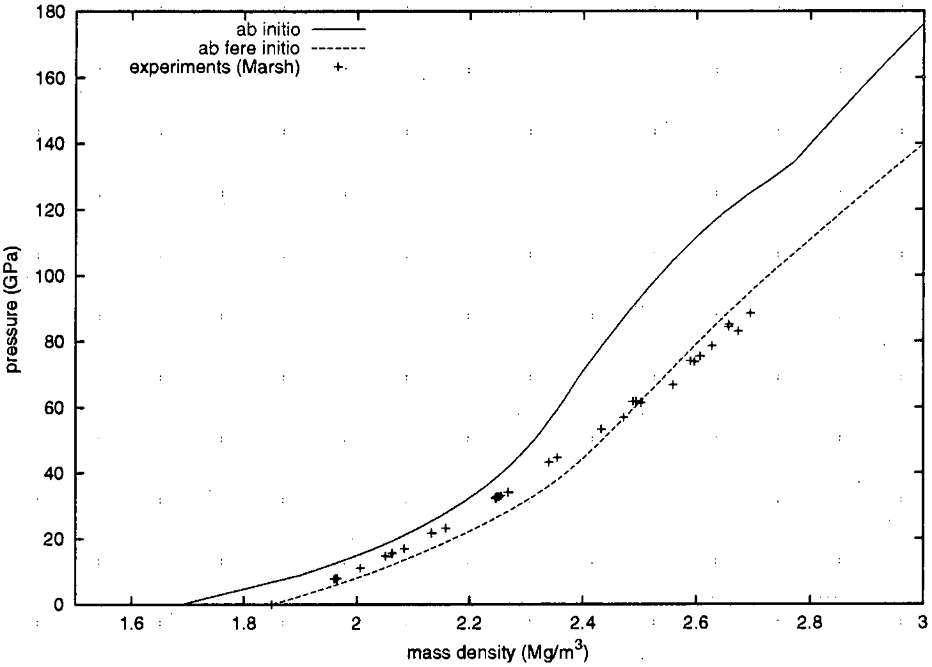

Figure 3.49. Shock Hugoniot for Be (pressure – density space).

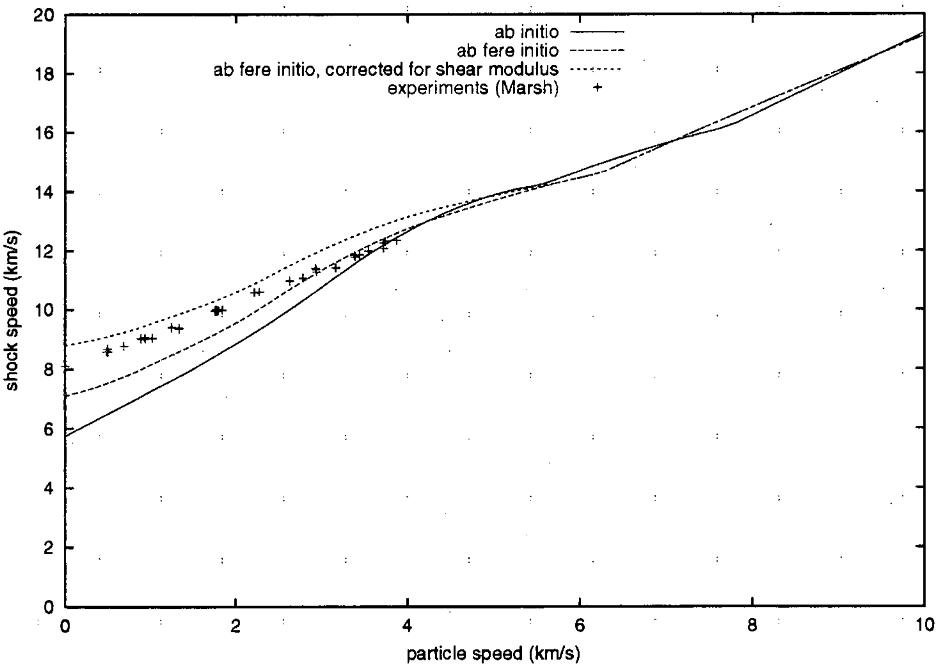

Figure 3.50. Shock Hugoniot for Be (shock speed – particle speed space).

Chapter 4

Use in hydrodynamic simulations

4.1 1D hydrocode

To demonstrate the application of the EOS tables, and to investigate the effect of complicated EOS surfaces on hydrodynamic waves, a 1D continuum mechanics program [82, 83] (often known as a 'hydrocode') was written to simulate the time evolution of a multi-material system. A library of material properties functions was written, allowing the EOS tables to be used in the hydrocode.

4.1.1 Continuum equations

Consider the mechanical response of a material in the continuum approximation. Conservation of mass, momentum and energy are expressed by the following relations:

$$\begin{aligned} \frac{\partial \rho}{\partial t} + \mathrm{div}(\rho \vec{u}) &= 0 \qquad (4.1) \\ \frac{\partial \rho \vec{u}}{\partial t} + \mathrm{grad}(\rho u^2 + p) &= 0 \qquad (4.2) \\ \frac{\partial e}{\partial t} + \mathrm{div}(e\vec{u}) + \frac{p}{\rho}\mathrm{div}\,\vec{u} &= 0, \qquad (4.3) \end{aligned}$$

where the mass density ρ, material velocity $\vec{u}$, pressure p and specific internal energy e are functions of position $\vec{r}$ and time t. These are a version of the Euler equations [82, 83, 84] describing the motion of material under the action of hydrodynamic forces and in the absence of thermal conduction. The stress – strain and conduction contributions are straightforward to add, but are outwith the scope of the present work.

The continuum equations are closed using the equation of mechanical state $p(\rho, e)$. They form a hyperbolic initial value problem: given the fields $\rho(\vec{r})$, $e(\vec{r})$ and $\vec{u}(\vec{r})$ at some time t_0, the variation with time can be found over some region $\{\vec{r}\} \in \mathcal{R}$ given the variation of $\vec{u}$ or p on the boundary $\delta\mathcal{R}^N$, where N is the number of spatial dimensions considered. In the absence of stress terms (and other physical processes such as chemical reactions), the Euler equations have characteristics [85] moving at $\vec{u}$ – corresponding to the local flow – and $\vec{u} + c\hat{v}$ – a cone of characteristics corresponding to signals travelling at the local sound speed in all directions (values of the unit vector $\hat{v}$).

In many cases, and certainly when developing new material models, it is convenient to solve the continuum equations in a coordinate system which moves with the material. This (Lagrangian) formulation makes it easier to add models of physical processes and associated spatial fields, without having to include in the continuum equations contributions from spatial variations in the fields when material is not stationary [82, 84]. The Lagrangian formulation of the continuum equations is

$$\begin{aligned}
\frac{D\rho}{Dt} + \rho\,\mathrm{div}\,\vec{u} &= 0 && (4.4)\\
\rho\frac{D\vec{u}}{Dt} + \mathrm{grad}\,p &= 0 && (4.5)\\
\rho\frac{De}{Dt} + p\,\mathrm{div}\vec{u} &= 0, && (4.6)
\end{aligned}$$

where D/Dt is the Lagrangian derivative operator,

$$\frac{D}{Dt} \equiv \frac{\partial}{\partial t} + \vec{u}.\mathrm{grad}. \qquad (4.7)$$

The Lagrangian derivative is the time derivative in the frame moving locally with the material, i.e. following the characteristics which travels at $\vec{u}$.

4.1.2 Discrete representation

The continuum equations were represented in discrete form using a staggered mesh [83, 84]. The spatial domain was divided into finite **cells**, each defined by a set of **nodes** on its boundary. Values of the position $\vec{r}$ and material speed $\vec{u}$ are carried on the nodes; values of the mass density ρ and specific internal energy e are carried in the cells. (A **material type** is also defined for each cell, for multi-material problems. A separate equation of state is defined for each material type.) (Fig. 4.1.)

The staggered mesh formulation is convenient in allowing mass and energy to be conserved exactly. Since the spatial domain in split into cells with well-defined boundaries, and the mass and energy are defined at the cell centres, then it is possible to ensure that the products of density and volume, and of specific energy, density and volume, are kept constant in the numerical integration.

In the Lagrangian staggered mesh formulation, the continuum equations become

$$\frac{d\rho_i}{dt} = -\rho_i\,\mathrm{div}_i\{\vec{u}_j\} \qquad (4.8)$$

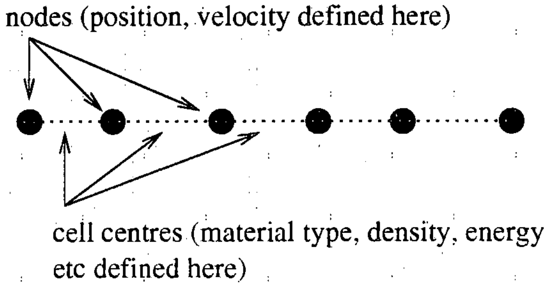
nodes (position, velocity defined here)
cell centres (material type, density, energy etc defined here)

Figure 4.1. Staggered mesh scheme for defining continuum fields.

$$\frac{d\vec{u}_j}{dt} = -\frac{1}{\bar{\rho}_j}\text{grad}_j\{p_i\} \tag{4.9}$$

$$\frac{de_i}{dt} = -\frac{1}{\rho_i}p_i\,\text{div}_i\{\vec{u}_j\} \tag{4.10}$$

$$\frac{d\vec{r}_j}{dt} = \vec{u}_j, \tag{4.11}$$

where the index i runs over the cells and j over the nodes. $\bar{\rho}_j$ is the mass density associated with node j, calculated as the volume-weighted average density in all cells containing the node. The differential operators return the divergence of a nodal field at the centre of the cell defined by the nodes, and the gradient of a cell-centred field at the node associated with the cells.

4.1.3 Numerical integration

The continuum equations were integrated using a predictor-corrector scheme [83]. An overall second order update over a time interval δt is obtained from a first order forward-time prediction over $\delta t/2$ followed by a leapfrog step over the full δt:

$$\tilde{f}(t+\delta t/2) = f(t) + \frac{\delta t}{2}\dot{f}[f(t)] \tag{4.12}$$

$$f(t+\delta t) = f(t) + \delta t\dot{f}[\tilde{f}(t+\delta t/2)] \tag{4.13}$$

where f is the set of field variables, $\tilde{f}$ their first order estimates and $\dot{f}$ the Lagrangian derivatives calculated using the field values at the appropriate time level.

If a simulation includes processes which take place on different time scales, it can be more efficient to integrate the different sets of equations separately, the parameters in each set taken as constant while the other set is updated. For example, it might be possible to split the state parameters π into a collection of subsets $\{\pi_i\}$. To integrate the state from known conditions at a time t over a time interval δt, the rate equations

$$\dot{\pi}(t) = C(\pi(t)) \tag{4.14}$$

are split in the same way, so

$$\dot{\pi}_i(t) = C_i(\pi_i(t), \tilde{\pi}_{j\neq i}(\tilde{t})), \tag{4.15}$$

where $\tilde{\pi}_j$ and $\tilde{t}$ signify that the π_j may have been updated already, before the π_i, and hence may refer to a different (constant) time between t and $t + \delta t$. Each set of parameters π_i can be updated using whatever numerical scheme is appropriate.

This approach, known as operator-splitting, makes it simple to model processes with different time scales in the same simulation. There are drawbacks, since the integration may not be as accurate as if the variation of all π_j during the time interval were taken into account when integrating the π_i. Thus, simulations using operator-splitting may be more sensitive to mesh resolution than fully time-accurate simulations.

Consider a continuum model with a mixture of materials, each of which has a state defined by its density and specific internal energy (ρ_i, e_i) (with corresponding equations of state for p, T etc) and a volume fraction f_i. With operator-splitting, the hydrodynamic equations are solved for constant $f_i(\vec{r})$, and phase change rates, mechanical and thermal equilibration and the changes to any other internal state parameters (e.g. defect concentrations) applied separately to adjust the state at each position $\vec{r}$.

There are many possible ways of applying operator-splitting to the different processes. For instance, the phase change could be integrated over the whole time step δt, and then equilibration applied over the whole time step. Alternatively, the phase change process could be split into substeps $\delta t' = \delta t/N$ for some N, and equilibration applied over each substep $\delta t'$.

4.1.4 Treatment of shock waves

The numerical scheme described above is fairly standard for continuum mechanics, and is unstable to the build up of shock waves. As is also standard, an **artificial viscosity** was used to stabilise the scheme for the calculation of shock waves [83, 85]. In the continuum representation, shock waves are spatial discontinuities in the state. The artificial viscosity spreads a shock wave over a few cells. In the energy update, the pressure p_i in each cell is replaced by

$$p_i' \equiv p_i + q_i \tag{4.16}$$

where q is a pseudo-viscous pressure calculated from the divergence of material velocity:

$$q = \alpha^2 \rho \, du^2 + \beta \rho c \, du \tag{4.17}$$

where

$$du \equiv |\min(dr\,\mathrm{div}\vec{u}, 0)|, \tag{4.18}$$

dr is a scale length associated with each cell and α and β are global parameters. It can be seen that q is nonzero only during compression. The values of α and β are essentially free; 'standard' values are common for different applications, depending (among other considerations) on the amount of shock smearing which can be tolerated.

4.1.5 Time step constraints

The predictor-corrector is an explicit scheme, so the allowed values of the time step δt are constrained in order to remain stable. The time step constraints are

$$\delta t_{ci} < dr_i/c_i \quad \text{(Courant condition)} \tag{4.19}$$

$$\delta t_{di} < 1/\mathrm{div}_i\{\vec{u}_j\} \quad \text{(divergence condition)} \tag{4.20}$$

$$\delta t_{qi} < dr/(4\alpha^2 \, du_i + 2\beta c) \quad \text{(artificial viscosity condition).} \tag{4.21}$$

The Courant condition ensures that causality is preserved (no hydrodynamic signal can travel faster than the local sound speed); the other conditions are necessary for accuracy. The maximum time step for each cell is

$$\delta t_i = \min(\gamma_c \delta t_{ci}, \gamma_d \delta t_{di}, \gamma_q \delta t_{qi}) \tag{4.22}$$

where the γ are safety factors, necessary because the system of equations is non-linear. Common choices for the γ are ~ 0.7.

In principle, the continuum equations could be integrated with different local time steps. Again in accordance with normal procedures, a global time step δt was found at each time t,

$$\delta t = \min_i(\delta t_i). \tag{4.23}$$

Although less efficient than local time steps, this method is simpler to implement and makes it more convenient to calculate the fields at subsequent times.

4.1.6 Software implementation

The hydrocode was written in C++ using the WXC++ class library. An object-oriented structure was adopted; this seemed slightly more awkward to implement than a traditional Fortran-style structure, but proved significantly easier to modify to add alternative material models and spatial fields.

The hydrocode was 1D in the definition of position and velocity (scalar values rather than vectors) and in the calculation of the field derivatives. However, the structure was designed to be relatively straightforward to generalise to 2D and 3D, e.g. allowing a list of nodes for each cell. (This comment does not imply that the hydrocode would not be prone to multidimensional instabilities without further modifications to the numerics.)

A C++ materials package was developed for use with the hydrocode and other scalar codes used in this work, e.g. for predicting Hugoniot and isentrope loci. This is fairly general, containing a range of equation of state forms.

Boundary conditions were defined as a list of nodes at which the velocity history was prescribed according to an arbitrary function. These nodes are normally the outermost in the region of simulation. Pressure boundary conditions can be simulated by defining a region of constant 'applied pressure' (one of the material types allowed).

Part of the data is an arbitrary scale function of position $\sigma(r)$, used in calculating the divergence:

$$\mathrm{div}_i u = \frac{\sigma(r_{i+\frac{1}{2}})u_{i+\frac{1}{2}} - \sigma(r_{i-\frac{1}{2}})u_{i-\frac{1}{2}}}{(r_{i+\frac{1}{2}} - r_{i-\frac{1}{2}})(\sigma(r_{i-\frac{1}{2}}) + \sigma(r_{i+\frac{1}{2}}))/2} \tag{4.24}$$

instead of the plane geometry

$$\mathrm{div}_i u = \frac{u_{i+\frac{1}{2}} - u_{i-\frac{1}{2}}}{x_{i+\frac{1}{2}} - x_{i-\frac{1}{2}}}. \tag{4.25}$$

By different choices of σ, the simulation can be altered between plane, cylindrical and spherical symmetry, or indeed to '1D' flow in a profiled duct:

$$\sigma(r) = \begin{cases} 1 & \text{plane} \\ 1/r & \text{cylindrical} \\ 1/r^2 & \text{spherical} \end{cases} \tag{4.26}$$

4.1.7 Validation

The hydrocode was tested by simulating Sod's problem [84, 86]. This consists of a notional bursting diaphragm between regions of high and low pressure. A shock is driven into the region of low pressure; a rarefaction runs into the region of high pressure. Between the shock and the rarefaction lies a contact surface. Across this, the pressure and material velocity are constant, but the low pressure material has been shocked and the high pressure material released to reach the same mechanical state. The density and internal energy at either side are therefore different. (Fig. 4.2.)

The material was a perfect gas with $\gamma = 1.4$; the high pressure region had $\rho = 1$ and $e = 2.5$; the low pressure region had $\rho = 0.125$ and $e = 2$. [1] The results were compared with analytic values. The discrepancies are:

- Smearing of sharp features (such as the head of the rarefaction fan) by the discrete representation of the fields.
- Additional smearing of the shock wave by the artificial viscosity.
- Errors in states where the shock started to form (overheating).

The accuracy was as satisfactory as would be expected with the numerical scheme described above. (Figs 4.3, 4.4, 4.5, 4.6 and 4.7.)

Simulations were also performed of a piston driving a steady shock into a region. The speed and states behind the shock agreed with solutions to the Rankine – Hugoniot equations.

---

[1] For hydrodynamics with no time-dependent processes, the units are irrelevant so long as a consistent set is used. Alternatively, all quantities can be treated as dimensionless.

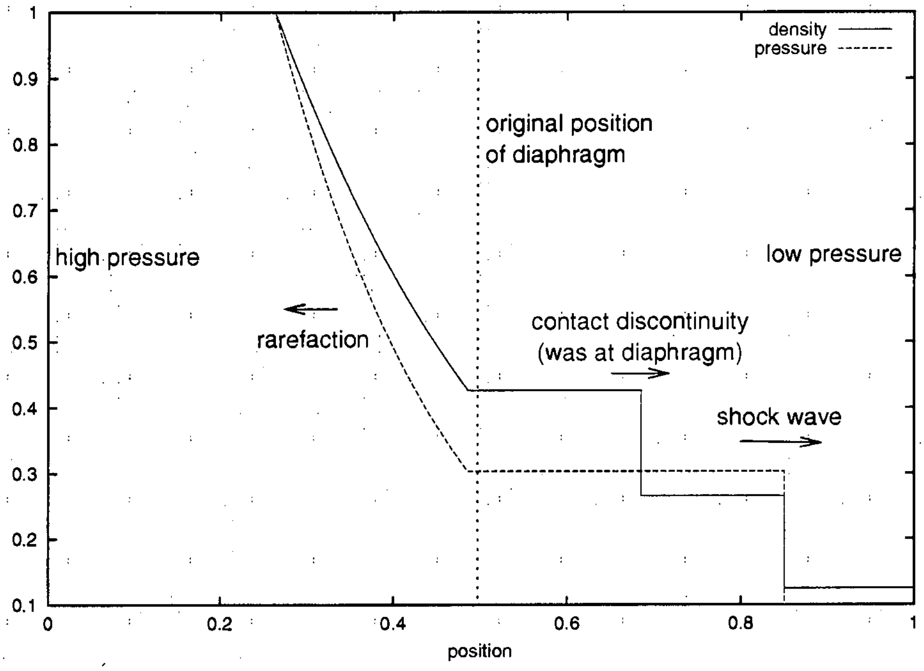

Figure 4.2. Sod's problem: snapshot during propagation of shock wave.

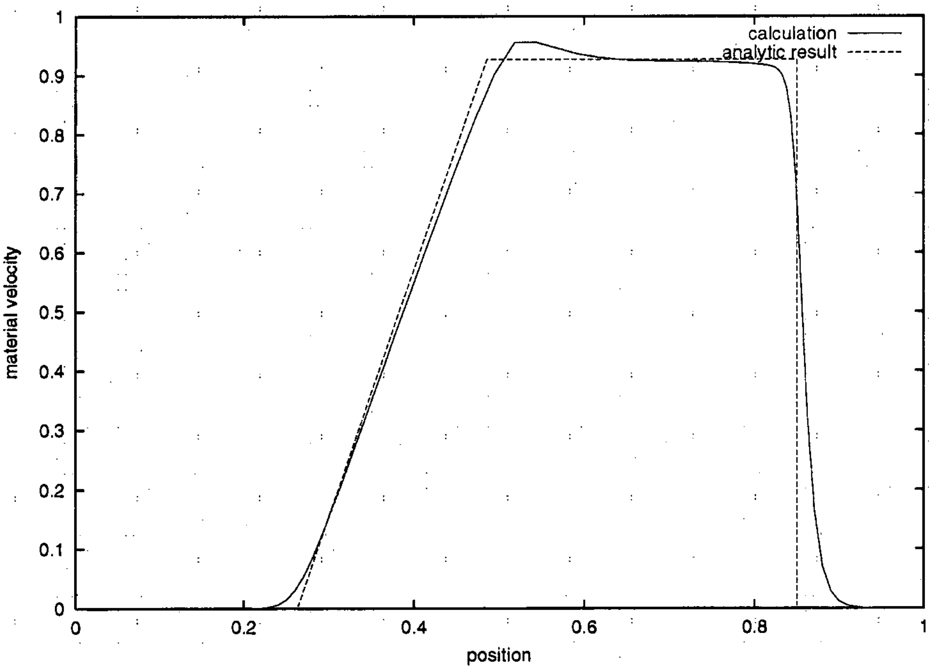

Figure 4.3. Sod's problem: particle velocity at time 0.2.

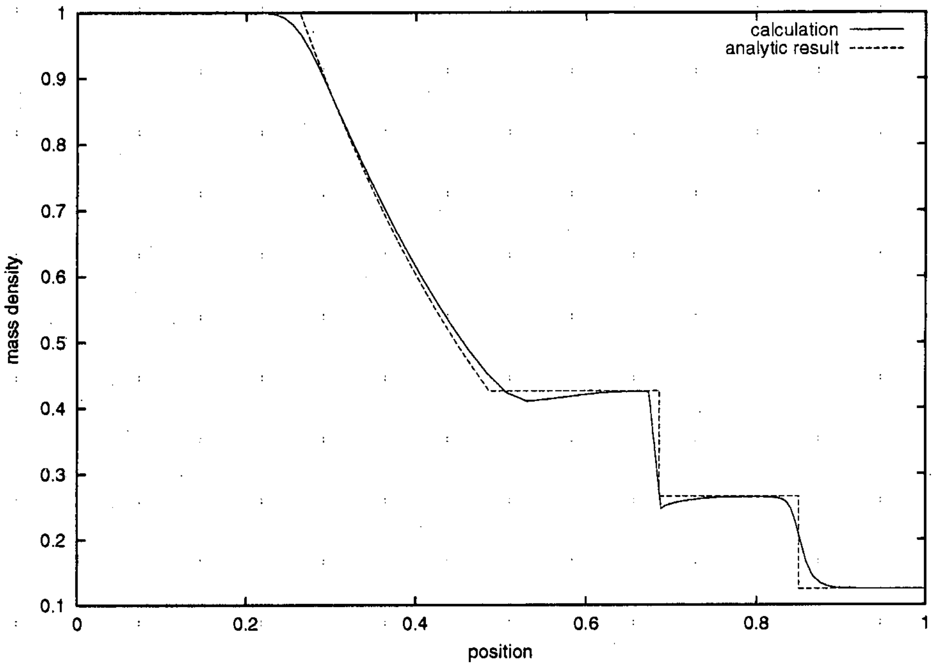

Figure 4.4. Sod's problem: mass density at time 0.2.

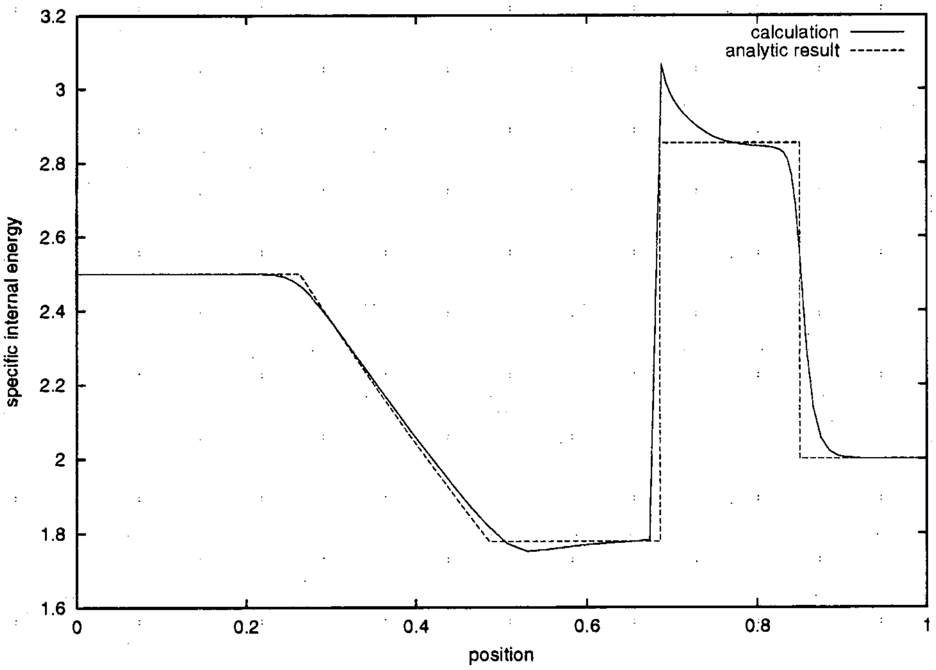

Figure 4.5. Sod's problem: specific internal energy at time 0.2.

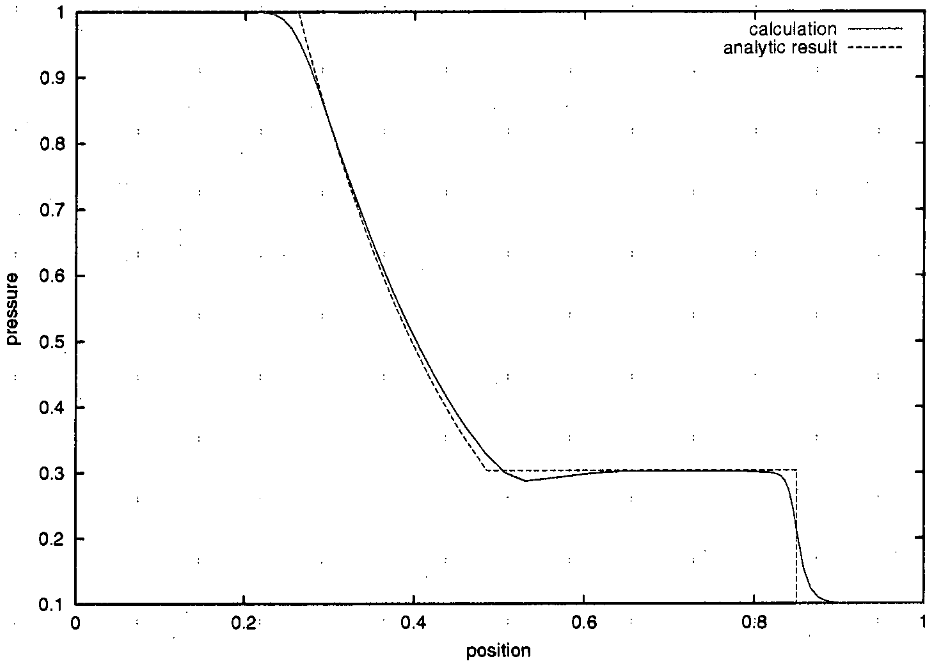

Figure 4.6. Sod's problem: pressure at time 0.2.

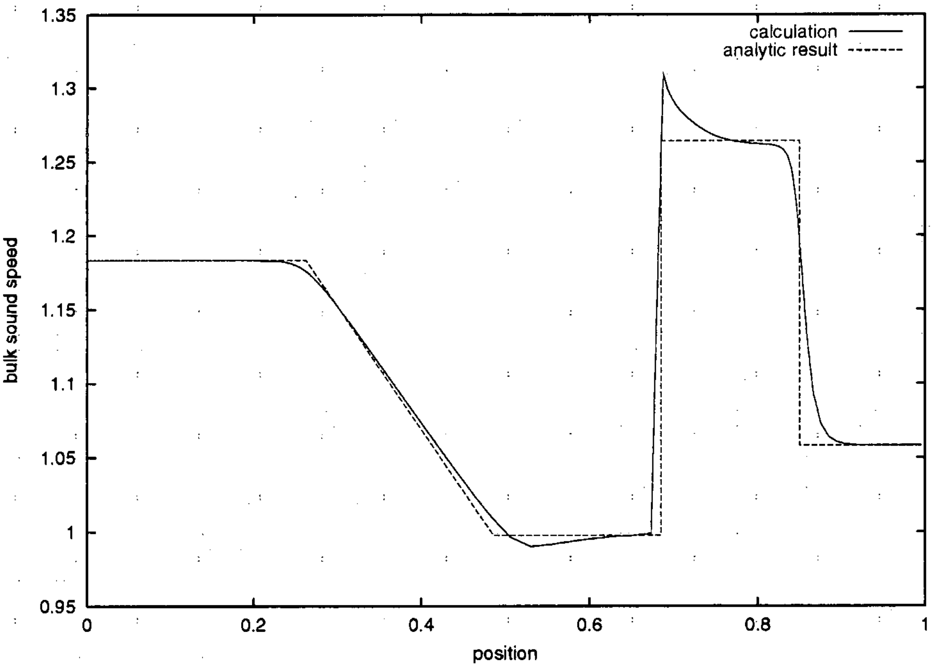

Figure 4.7. Sod's problem: bulk sound speed at time 0.2.

4.2 Double shock structure in silicon

4.2.1 Stability of shock waves in polymorphic materials

The speed u_s of a shock wave with respect to the undisturbed material ahead is given by the Rankine – Hugoniot equation [1]

$$u_s^2 = v_0^2 \frac{p - p_0}{v_0 - v}. \tag{4.27}$$

If $(p - p_0)/(v_0 - v)$ does not increase monotonically with ρ then a shock of some pressure p_1 may split into a pair of shocks, one moving at maximum speed for $p < p_1$, and with the corresponding pressure p_a say, and a second shock which completes the compression to p_1 (Fig. 4.8).

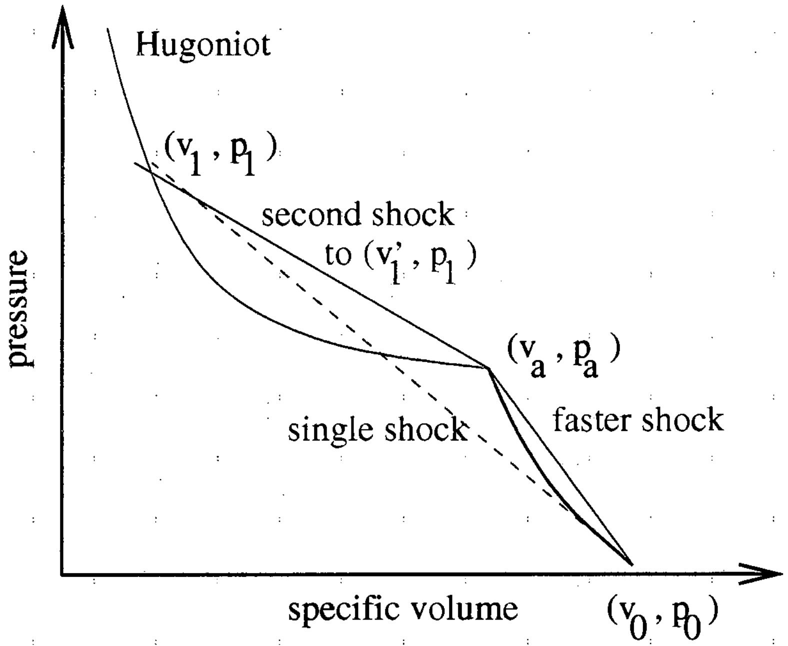

Figure 4.8. Shock splitting on a non-monotonic Hugoniot.

In principle, a single shock to p_1 could be maintained. In practice such a shock is unstable to perturbations in state. Any small variation of density, temperature or shock pressure could trigger the split. In real materials the variations might be impurities, defects, grain boundaries or even just thermal motion of the atoms.

Below the density corresponding to neutron star formation, the cold curve of any polymorphic phase stiffens with compression. Irrespective of phase transitions, the shock Hugoniot should ultimately stiffen. The region(s) of pressure over which a shock wave may split should therefore be bounded above and below. (Fig. 4.9).

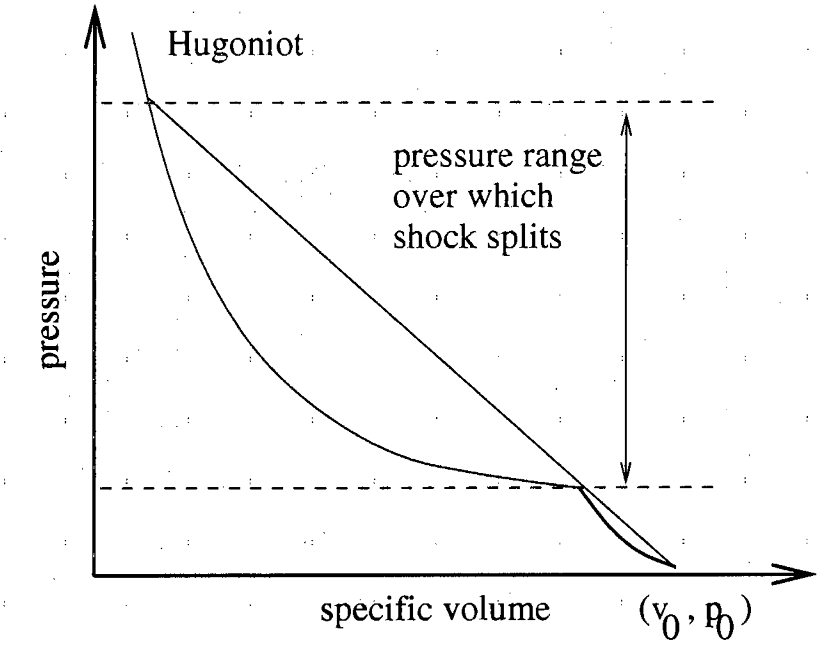

Figure 4.9. Range of shock strengths which may split.

The principal Hugoniot is the locus of states which may be reached by the action of a single shock on the undisturbed material. If a shock wave splits into a double shock structure, the second shock must produce a state lying on the secondary Hugoniot starting at the state generated by the first shock. In general, the state behind a double shock is different to the state behind a single shock with the same final pressure; (ρ_1', e_1') say, giving the same pressure p_1. In practice, the difference is small for many cases, because the Hugoniot starting at any state is in second order contact with the isentrope through that state.

An analogous effect occurs when the elastic constants and flow stress of a material are considered. Shock waves of some range of pressures may split into an elastic precursor and a plastic wave. The measured profiles of such waves have been used to investigate physical processes contributing to material strength [70].

Table 4.1. Shock splitting in silicon.

equation of state	pressure		speed of first shock (km/s)
	minimum (GPa)	maximum (GPa)	
Grüneisen	21	38	6.63
ab fere initio	23	40	6.78
ab fere initio, BCT adjusted	18	35	6.43

4.2.2 Equilibrium / isotropic predictions for silicon

Using the diamond/BCT equations of state calculated for silicon, the range of pressures leading to shock splitting was predicted. (Table 4.1 and Fig. 4.10.) The predictions fall in a fairly narrow range. The upper limit for splitting caused by this phase change may not be observed in practice because other phase changes may complicate the shock structure.

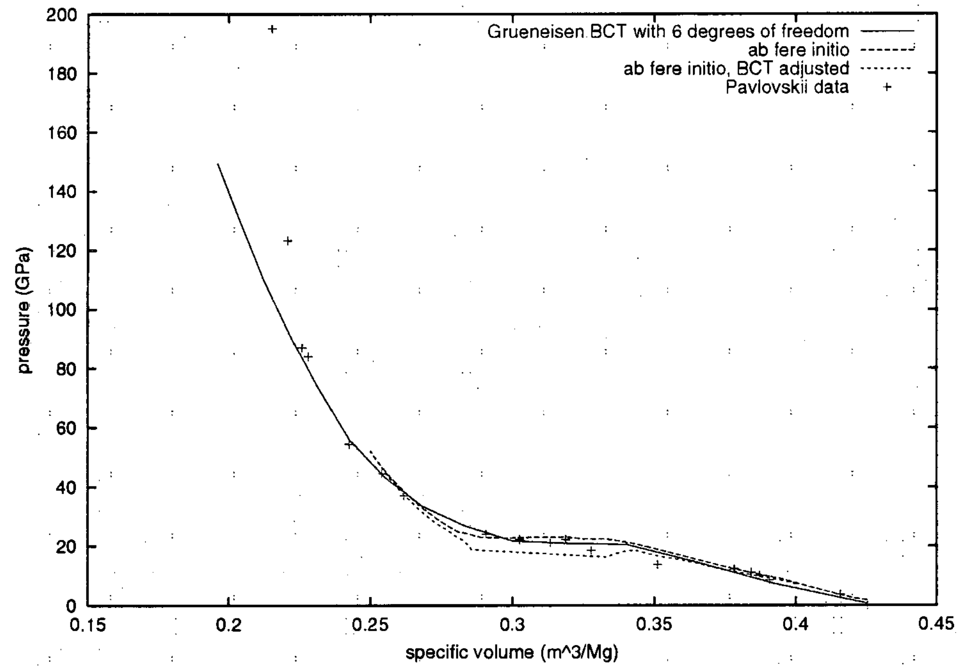

Figure 4.10. Hugoniots for silicon (pressure – specific volume plane).

The pressure range calculated here for shock splitting is based on applying the equilibrium equation of state to isotropic hydrodynamics. These calculations contain two important simplifications:

- Phase transitions are instantaneous with no energy barrier. In the case of the diamond-to-BCT, the transition can take place by compression of the diamond lattice in any of the coordinate directions, so this simplification should not matter.
- The stress – strain contribution to the response of the continuum is negligible.

These corrections are addressed in extensions to this work where the EOS techniques developed here are applied to the calculation of elastic and plastic properties [87, 88].

4.2.3 Comparison with transient X-ray diffraction data

Measurements of the shock wave structure have been made using a laser to drive a plane shock into a sample of material, and an intense source of X-rays to allow the time-variation of a crystal diffraction peak to be followed.

Lasers can also provide convenient source of X-rays. One problem is synchronising the X-rays with the shock; using a laser for both purposes allows accurate synchronisation. X-rays can be generated by focusing a laser beam onto a solid. The intense energy deposition can produce a hot, dense plasma which is a strong emitter of k-line radiation. Other X-ray sources are possible [89], as are different methods of generating the shock wave such as gas guns and high explosives.

Experiments have been performed on silicon using the 'TRIDENT' laser at Los Alamos National Laboratory, USA [90]. A single crystal of silicon was used, cut in the 001 direction. A thin layer of gold was deposited on the surface to be irradiated, in order to reduce electron preheat from the laser energy used to drive the shock. X-ray diffraction was measured in the Bragg and Laue configurations, using X-ray streak cameras which recorded the angular variation of a diffraction peak as a function of time. The diffraction peaks are generated through constructive interference of X-ray photons scattered from many atoms in the crystal, and the signal recorded at the detector is thus the integral over a finite depth of crystal. As the shock wave approaches the free surface, a peak corresponding to compressed material is seen before the undisturbed peak disappears, because the diffraction pattern probes a finite thickness of material. (Figs 4.11 and 4.12.)

The X-ray streak cameras operated by converting X-ray photons from a line in space into electrons, whose trajectory is deflected by a varying electric field to image the line across a rectangular region of film or CCD array. Some X-rays were able to penetrate the photoconvertor and expose the detector directly. As a result, the TXD images typically exhibit a vertical feature which is a projection of the camera slit (i.e. the imaging line) directly onto the streak image with no temporal variation.

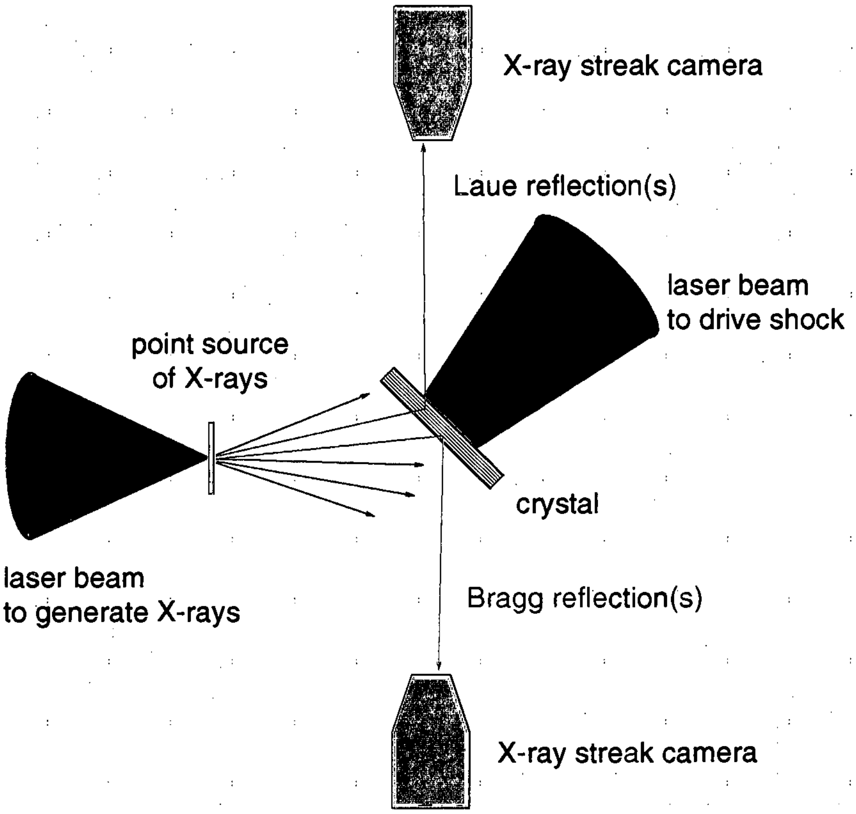
X-ray streak camera
Laue reflection(s)
laser beam
to drive shock
point source
of X-rays
crystal
laser beam
to generate X-rays
Bragg reflection(s)
X-ray streak camera

Figure 4.11. Schematic of transient X-ray diffraction experiments at 'TRIDENT'.

The X-ray source consisted of two closely-spaced lines from the helium-like α and β transitions in the plasma. These could be seen fairly clearly in some of the shocked and unshocked signals.

The pressure generated by the laser pulse was not known exactly, but estimates have been made given the laser output and calculations of energy deposition in the silicon. The laser pulse used to drive the shock wave had a roughly constant intensity for a few ns. A constant laser intensity does not necessarily generate a shock of constant pressure – some of the results were indicative of a shock whose pressure increased slightly with time, producing a ramping increase in the Bragg angle from the shocked material.

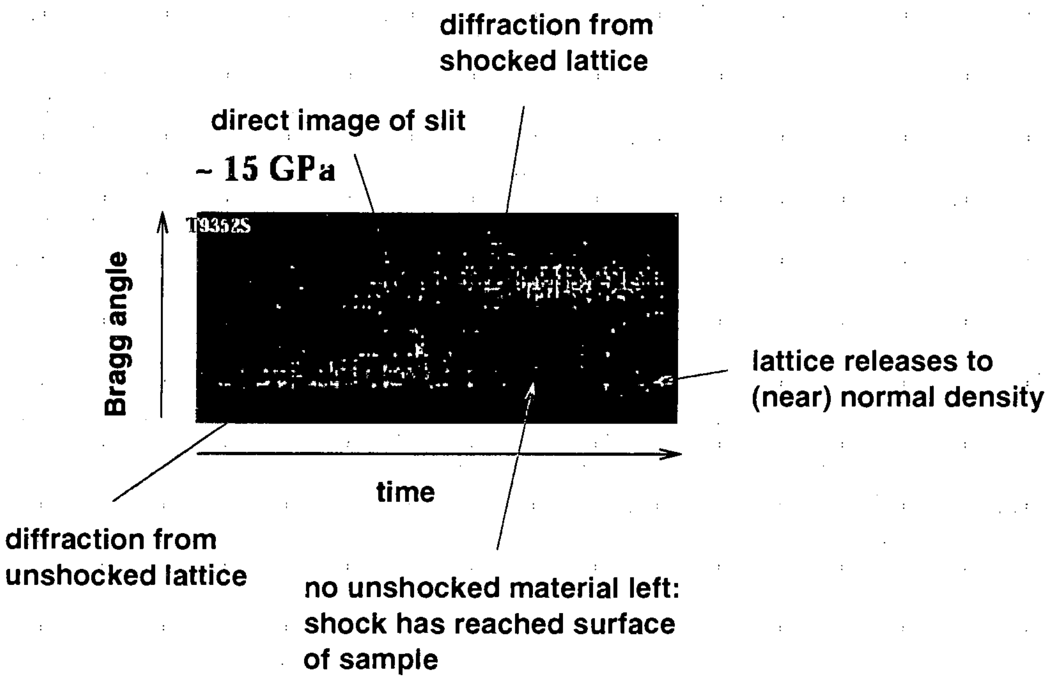
diffraction from
shocked lattice
direct image of slit
~ 15 GPa
T9352S
Bragg angle
lattice releases to
(near) normal density
time
diffraction from
unshocked lattice
no unshocked material left:
shock has reached surface
of sample

Figure 4.12. Sample result from transient X-ray diffraction experiment.

For shock pressures greater than about 20 GPa, the diffraction peak from the shocked material was observed to split into two distinct lines a finite time after the signal from the compressed lattice was first seen. The split was significantly different to the fine structure from the α and β lines. This shock structure was tentatively identified as a phase transition, though perhaps to simple hexagonal rather than BCT, as the former is thought to be energetically more favourable. More structure was evident in the traces on release from silicon shocked into a different phase; this may indicate the presence of further phases or of planes temporarily rotated at a different angle to the shock wave. (Fig 4.13.)

The split first appeared ~2.2 ns after the earliest signal from the shocked material in a sample 22 μm thick, with a driving pressure of somewhat over 20 GPa [91].

The 1D Lagrangian hydrocode was used to simulate the shock wave structure, using the 'quasiharmonic 1' equation of state. The simulations demonstrated the formation of an initial shock, a ramp-like structure in the mixed-phase region, and a second shock completing the compression. Each shock was smeared over a few computational cells, because of the artificial viscosity treatment discussed above. The first shock had a slightly lumpy structure, probably caused by the use of linear interpolation over the relatively coarse tabulation of the equation

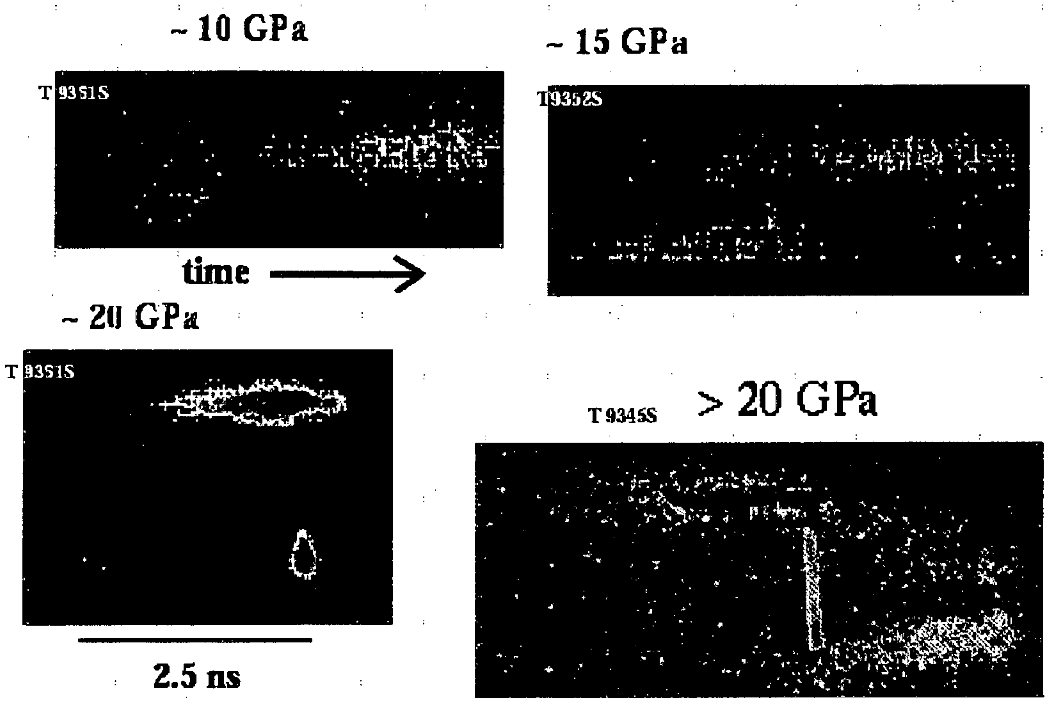

Figure 4.13. Transient X-ray diffraction traces from 'TRIDENT' experiments on silicon.

of state. Other calculations [78] indicate that the BCT cold curve is similar to simple hexagonal in this regime, so the diamond/BCT EOS should provide a reasonable model of the shock structure. (Fig. 4.14.)

According to the hydrocode calculations, a driving pressure of 25 GPa caused a split shock structure where the first shock travelled at ~6.5 km/s and the second at ~4 km/s. The speed of the second shock is quoted with respect to the undisturbed material; its speed with respect to the moving material behind the first shock is lower. These speeds give a time difference ~2 ns after crossing a sample 22 μm thick, clearly consistent with the experimental results. A more precise comparison is not yet possible, as the shock wave in the experiment needs to be characterised more carefully. (The speed of the second shock is relatively sensitive to the pressure.)

There is further scope for developing the experimental technique. If additional diffraction lines are used then it may be possible to identify transient phases with more certainty. The temperature and defect density might be obtained from the angular width of the diffraction lines.

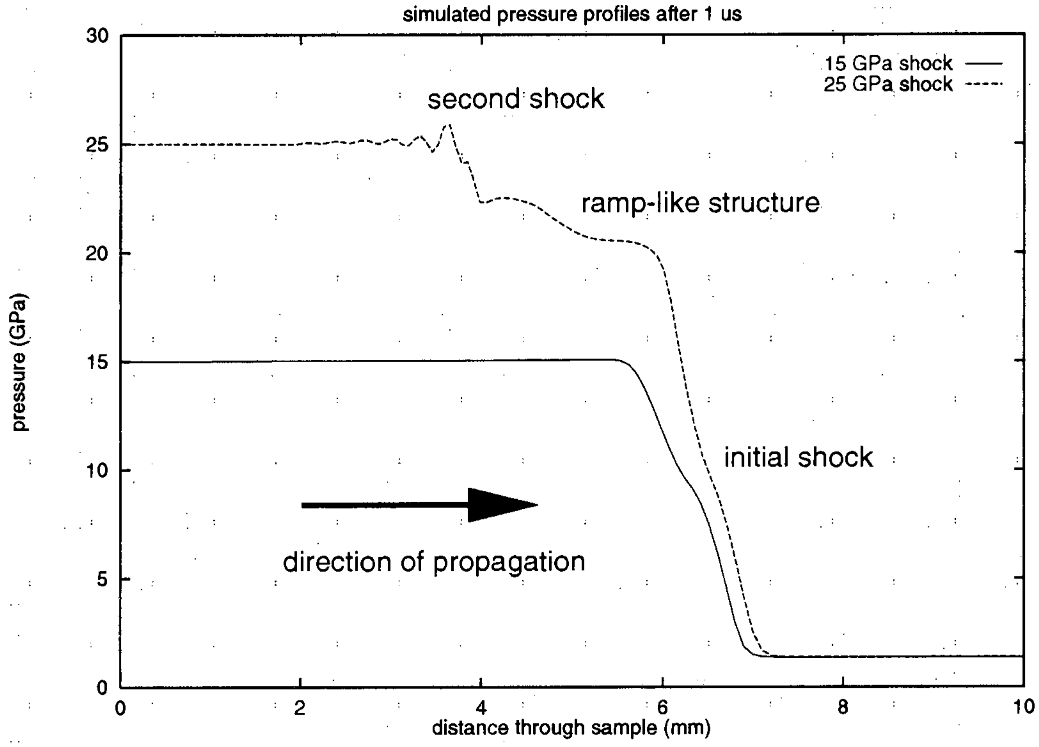

Figure 4.14. Example pressure profiles from hydrocode simulations of shocks in silicon with the diamond/BCT equation of state.
(Driving pressures above and below the threshold for phase transition.)

Chapter 5

Conclusions

Ab initio pseudopotential methods (together with quantum mechanical calculations of the phonon modes and electron-thermal effects based on the band structure) allowed equations of state to be predicted to within a few % in density or a few GPa in pressure, comparing with the density at STP and measurements of the shock Hugoniot. Care was needed that the pseudopotentials were valid in the compression range required, as the accuracy became worse as the cores approached each other.

The method was applied to aluminium, silicon (diamond and body-centred tetragonal structures) and beryllium (hexagonal and body-centred cubic structures). For strong materials such as beryllium, it was necessary to take account of the elastic modulus and yield stress in order to match the shock Hugoniot. This was done in a simple manner.

The *ab initio* phonon calculations rely on force calculations for displaced atoms in the electron ground state code. These calculations are relatively expensive in terms of computational effort, so interpolation schemes were investigated to allow the forces to be calculated for a smaller set of lattice parameters than was used for the cold curve. Considerable care was needed to allow the phonon properties to be interpolated accurately. The more robust scheme was to interpolate the lattice-thermal energy rather than the force constants, frequencies or the density of states.

Interatomic potentials were investigated as a means of representing the effective interatomic interactions in a more compact way. The potentials were obtained by adjusting their parameters to reproduce the energy – volume relations predicted for different structures on the cold curve. Where valid, these potentials allowed phonon modes to be predicted more efficiently, and could be used for Monte-Carlo simulations to investigate anharmonic contributions (phonon-phonon interactions) to the equation of state. Any form of potential seemed capable of fitting the cold curve for a single structure, but multi-body potentials such as the Finnis-Sinclair were necessary to reproduce the energy difference between structures. However, fitting of multi-body potentials to cold curve data, where the stress is isotropic, was found to be an ill-conditioned process, and the lattice bound by the resulting potential was prone to shear instabilities.

The effect of anharmonic contributions was found to be small over the range of states sampled along the shock Hugoniot.

Considerable computational effort was required to obtain *ab initio* phonons in good agreement with the observed density of states. However, because integrations are performed over the density of states in order to calculate the lattice-thermal energy, the equation of state is relatively sensitive to the density of states. It appeared necessary to ensure only that the first few moments of the density of states (mean, variance, skewness) were in agreement. This required much less computational effort.

The accuracy of each equation of state was improved by adding a pressure offset to bring the calculated density at some reference state (e.g. STP) into agreement with the observed value. The resulting *ab fere initio* equation of state matched isothermal and shock compression data well for the same polymorphic phase, up to 50 GPa or more.

The pressure offset was introduced as an energy tilt to the cold curve. Rigorously, the tilt should also manifest itself as a modification to the force constants in the interatomic interaction, and hence to the phonon modes. The effect of this extra correction was evaluated by predicting the lattice-thermal energy with and without the correction by calculating the phonon modes using interatomic potentials fitted to the *ab initio* cold curve or to the adjusted cold curve. The difference was found to be negligible.

An additional attraction of the *ab initio* route compared with empirical equations of state is that predictions of quantities contributing to the equation of state can be verified by non-mechanical measurements. For example, the density of phonon states can be measured by neutron scattering and compared with the predictions used to calculate the lattice-thermal energy.

The basic method developed allows the equation of state to be predicted for a single polymorphic phase. Phase boundaries and transformation pressures can be estimated by repeating the procedure for each phase in turn, and constructing the surface of minimum free energy in density – temperature space. Phase boundaries are extremely sensitive to the precise details of the model used for each phase. Different phases should be calculated in as consistent a way as possible in order to obtain a reasonable prediction of the phase boundary.

A model was developed for predicting the time-dependence of polymorphic phase transitions.

Ab fere initio equations of state for silicon, including the diamond/BCT phase transition, gave a reasonable match to the observed shock Hugoniot. The equation of state also predicted that shocks of between ~20 and 40 GPa should split into a pair of shocks travelling with different speeds. This prediction is in (at least) qualitative agreement with observations make using transient X-ray diffraction.

The overall method seems straightforward to apply to different elements – including structures which are not close-packed or metallic – and should work equally well for stoichiometric compounds (including many alloys). This represents a considerable extension of the more common approach of predicting properties under compression at $T = 0$.

The *ab initio* quantum mechanical method for predicting equations of state could be extended in a variety of ways. There is scope for developing the way in which the electron ground states are predicted, in the techniques used to derive pseudopotentials, all-electron methods and the treatment of multi-electron effects such as exchange and correlation. The use of the band structure in estimating the electron-thermal energy is only valid for temperatures well below the Fermi temperature; it would be possible to predict equations of state at much higher temperatures if excited states of the electrons were considered.

A prototype computer program was written to calculate electron wavefunctions according to the Dirac equation. The program was not made efficient enough for routine equation of state calculations, though it was demonstrated to work accurately on simple trial systems. In the future, this program or a similar one should provide various advantages such as the inclusion of relativistic effects and the use of basis functions that are more efficient for equations of state.

The use of electron ground states in 3D representations of the lattice opens up the possibility of making useful predictions of the elastic constants and perhaps the plastic behaviour of a material. This extension to predictions of tensor material properties is likely to have important technological applications. Elastic constants have been calculated for some elements and compounds, [21, 92, 93, 94, 95, 96], but the approach followed here opens up the possibility of applying a consistent and closely-knit set of techniques to the equation of state, phase diagram, elasticity and plasticity [97]. Other properties such as the creation energy and diffusion rate of point defects [98] could be calculated in a similar way through a mixture of *ab initio* electronic states and interatomic potentials deduced from the electronic states [99, 100].

Appendix A

Units

In generating equations of state from quantum mechanics, it is convenient to work in different sets of units when performing different parts of the calculation.

It is convenient to perform all calculations at the level of the atomic lattice using units based on eV, amu and Å. Similarly, when considering applications to strong shock waves, it is convenient to work in units which reflect typical characteristics of the applications, with times in μs and sizes in mm. In the latter case, it is possible to pick a variant of SI units where quantities are scaled judiciously by factors of 1000.

Table A.1. Units and conversions.

quantity	**SI units**	**atomic level**	**strong shock**
length	m	Å	mm
energy	J	eV	MJ
mass	kg	amu	Mg
time	s	$\text{Å}\sqrt{\text{amu/eV}}$ $\simeq 1.01805466 \times 10^{-14}$ s	μs
speed	m/s		mm/μs = km/s
mass density	kg/m^3	amu/Å^3	Mg/m^3 = g/cm^3
specific energy	J/kg	eV/amu	MJ/kg
pressure	Pa	eV/Å^3	GPa (= 100 kb)

Appendix B

Crystallographic structures

B.1 Introduction

For convenience, the lattice types used in the EOS calculations are listed below.

Cubic structures satisfy 48 rotation matrices. This high degree of symmetry means that a single displacement – for instance, of the atom at $(0,0,0)$ along the x-direction – is all that is required to determine the phonon eigenproblem for each value of the lattice parameter.

B.2 Structures

Face-centred cubic (FCC)

Lattice parameter a; lattice vectors $a,0,0,\ 0,a,0,\ 0,0,a$. Atoms at $(0,0,0)$, $(\frac{1}{2},\frac{1}{2},0)$, $(\frac{1}{2},0,\frac{1}{2})$, $(0,\frac{1}{2},\frac{1}{2})$.

Body-centred cubic (BCC)

Lattice parameter a; lattice vectors $(a,0,0)$, $(0,a,0)$, $(0,0,a)$. Atoms at $(0,0,0)$, $(\frac{1}{2},\frac{1}{2},\frac{1}{2})$.

Diamond cubic (DIA)

Lattice parameter a; lattice vectors $(a,0,0)$, $(0,a,0)$, $(0,0,a)$ – cubic structure. Atoms at $(0,0,0)$, $(\frac{1}{2},\frac{1}{2},0)$, $(\frac{1}{2},0,\frac{1}{2})$, $(0,\frac{1}{2},\frac{1}{2})$, $(\frac{1}{4},\frac{1}{4},\frac{1}{4})$, $(\frac{3}{4},\frac{3}{4},\frac{1}{4})$, $(\frac{3}{4},\frac{1}{4},\frac{3}{4})$, $(\frac{1}{4},\frac{3}{4},\frac{3}{4})$.

Internal translation symmetry: $(\frac{1}{4},\frac{1}{4},\frac{1}{4})$.

Simple hexagonal

Lattice parameters a and c; lattice vectors $(a,0,0)$, $(a/2,\sqrt{3}a/2,0)$, $(0,0,c)$. Atom at $(0,0,0)$.

Hexagonal close-packed (HCP)

Lattice parameters a and c; lattice vectors $(a,0,0)$, $(a/2,\sqrt{3}a/2,0)$, $(0,0,c)$. Atoms at $(0,0,0)$, $(\frac{1}{3},\frac{1}{3},\frac{1}{2})$.

Alternative orthorhombic lattice cell

Lattice vectors $(a,0,0)$, $(0,\sqrt{3}a,0)$, $(0,0,c)$. Atoms at $(0,0,0)$, $(\frac{1}{2},\frac{1}{2},0)$, $(\frac{1}{2},\frac{1}{6},\frac{1}{2})$, $(0,\frac{2}{3},\frac{1}{2})$.

Body-centred tetragonal (BCT) or β-Sn

Lattice parameters a, c; lattice vectors $(a,0,0)$, $(0,a,0)$, $(0,0,c)$. Atoms at $(0,0,0)$, $(\frac{1}{2},\frac{1}{2},\frac{1}{2})$, $(\frac{1}{2},0,\frac{1}{4})$, $(0,\frac{1}{2},\frac{3}{4})$.

Si Imma

Lattice parameters a, b, c; lattice vectors $(a,0,0)$, $(0,b,0)$, $(0,0,c)$. Atom at $(0,\frac{1}{4},\frac{\Delta}{2})$ where $\Delta \simeq 0.386$ at 15 GPa.

To obtain phonon modes, displacements were made of each of the first two atoms in each of the x, y and z directions.

Appendix C

Interatomic potentials

C.1 Introduction

Although realistic quantum mechanical calculations are possible for systems with several atoms, they require a large amount of computer storage and time for collections of atoms large enough to provide meaningful information on thermal behaviour (e.g. the equation of state at finite temperatures) and dynamical questions (e.g. the rate of a phase transition).

Another way to proceed is to represent the total interaction energy of the collection of nuclei and electrons by an effective potential ϕ between individual nuclei at positions $\vec{r}_i$. Such potentials are intended to represent angular or many-body forces and quantum mechanical effects such as exchange and correlation between the electrons. These physical effects give rise to the volume and structure-dependence of the total energy. The structure dependence can in principle predict both polymorphism and local effects such as defect and disclination energies.

Typically, interatomic potentials are derived by fitting the parameters in an empirical form to some property measured experimentally or calculated by a different technique. An infinite range of possible empirical forms could be devised. The popular forms considered here were inverse power, Morse [101] and Finnis-Sinclair [102]. The potentials were fitted to the energies obtained from the ground state calculations, evaluating their ability to represent data for a single lattice structure and their transferability between structures. A phonon density of states at normal density was generated from each potential, and compared against experimental data.

Computer programs were written to calculate the total energy of a regular lattice given the structure and dimensions of the unit cell, for each of the potentials of interest. Programs were also written to fit coefficients in the interatomic potentials to reproduce datum values of specific internal energy at chosen mass densities. Several iterative techniques were used for fitting non-linear coefficients, including Gauss-Newton, steepest descents, conjugate gradients, and Monte-Carlo [51].

C.2 Forms of potential

C.2.1 Inverse power

A common suggestion for trial potentials for the energy between atoms i and j is

$$\phi_{ij} = \sum_i \alpha_i / r^i \tag{C.1}$$

This encompasses the Lennard-Jones potential used for the noble gases.

If the powers of r are known, then the fitting of the parameters α to data is a linear problem. Previous work [103] has shown that it is straightforward to obtain the α_i by the solution of a matrix equation. This can be made more efficient by the use of structure factors, containing the summation of each power of r over all the atoms in each lattice type. The summations need only be performed once. The matrix technique is described in Section C.3.

C.2.2 Morse

Morse potentials of the form

$$\phi_{ij} = D\left[e^{-2\alpha(|\vec{r}_i-\vec{r}_j|-r_0)} - 2e^{-\alpha(|\vec{r}_i-\vec{r}_j|-r_0)}\right] \tag{C.2}$$

for the potential energy between atoms i and j have been used for various purposes including the study of grain boundaries [101]. D, α and r_0 are material-dependent parameters. In some cases, an additional scalar constant ϕ_0 was added to the potential to make it easier to fit some of the quantum mechanical cold curves.

A hybrid fitting scheme was developed for fitting the Morse coefficients to data. The Morse potential is non-linear in α and r_0, but linear in D (and ϕ_0, if used). Given values for α and r_0, D and ϕ_0 can be found quickly and deterministically using the linear fitting method to form a matrix equation (Section C.3). Thus the non-linear fitting scheme can be restricted to use α and r_0 only. As a guide, the difficulty involved in non-linear fitting increases by roughly an order of magnitude for each additional parameter, so the hybrid scheme is valuable. It was found to be considerably faster and more reliable in practice than optimising all the parameters iteratively.

C.2.3 Finnis-Sinclair

An interatomic potential has been proposed [102] which mimics some of the N-body behaviour of the tight-binding approximation in solid state physics. The energy ϕ of atom i can be written [104]

$$\phi_i = \frac{1}{2}\sum_j V_{ij} - f(\rho_i) \tag{C.3}$$

where V_{ij} is a central pair potential between atoms i and j, f an embedding function (taken here to be the square root) and

$$\rho_i = \sum_j \Phi_{ij}. \tag{C.4}$$

The summations extend in principle over all the atoms of the system except atom i, but V and Φ are short-ranged functions only of the interatomic distance r_{ij}.

Previous work [104] has used cubic splines to model the potentials:

$$V(r) = \sum_{k=1}^{6} a_k(r_k - r)\theta(r_k - r) \tag{C.5}$$

$$\phi(r) = \sum_{k=1}^{2} A_k(R_k - r)\theta(R_k - r) \tag{C.6}$$

where θ is the Heaviside step function. The parameters are typically found by fitting to the equilibrium lattice spacing, the cohesive energy, the elastic constants, the vacancy formation energy and the stacking fault energy.

C.3 Interatomic pair potentials from bulk properties

In principle, it is possible to infer an effective interatomic pair potential $V(r)$ for atoms in a solid from the variation of specific potential energy e_c with specific volume v. The energy $e_c(v)$ is assumed to depend on spherically symmetric interatomic potentials through the following relation:

$$e_c(v) = \frac{1}{2}\sum_{i\neq j} V_{ij}(|\vec{r}_i - \vec{r}_j|)/\sum_i m_i \tag{C.7}$$

where i and j run over all atoms, $\vec{r}_i$ and $\vec{r}_j$ are the position vectors of atoms i and j (and hence functions of v), m_i is the mass of atom i and the factor $\frac{1}{2}$ ensures that the contribution from the potential between each pair of atoms is counted only once. In general, the potential between atoms i and j might depend on their type, so the potential function is denoted V_{ij}. For atoms in a crystalline solid, the summation can be performed with respect to a single lattice cell, over the lattice parameters k, l and m and over the s atoms (indexed by n) in each lattice cell.

Consider an atom in lattice cell k, l, m. The origin of the cell is at $k\vec{u} + l\vec{v} + m\vec{w}$, where $\vec{u}$, $\vec{v}$ and $\vec{w}$ are the lattice vectors. If atom n has co-ordinates $(\Delta k_n, \Delta l_n, \Delta m_n)$ with respect to its lattice cell then its position vector is

$$\vec{r}_{klmn} = (k + \Delta k_n)\vec{u} + (l + \Delta l_n)\vec{v} + (m + \Delta m_n)\vec{w}. \tag{C.8}$$

The volume of each lattice cell is vector triple product of the lattice vectors

$$v_c = [\vec{u}, \vec{v}, \vec{w}] = \vec{u}.(\vec{v} \wedge \vec{w}). \tag{C.9}$$

The mass of the cell is

$$m_c = \sum_n m_n, \tag{C.10}$$

where m_n is the mass of atom n. Thus the specific volume is

$$v = v_c/m_c = [\vec{u}, \vec{v}, \vec{w}]/\sum_n m_n. \tag{C.11}$$

The position vectors can be found from the specific volume, so long as the lattice vectors remain at constant relative angles and lengths:

$$\vec{u}(v) = \alpha(v)\vec{u}_0, \vec{v}(v) = \alpha(v)\vec{v}_0, \vec{w}(v) = \alpha(v)\vec{w}_0, \tag{C.12}$$

for some $\alpha(v)$ and constant set of vectors $\vec{u}_0, \vec{v}_0, \vec{w}_0$ giving a reference specific volume v_0. Note that

$$v = \alpha^3[\vec{u}_0, \vec{v}_0, \vec{w}_0] / \sum_n m_n = \alpha^3 v_0. \tag{C.13}$$

Thus

$$\vec{r}_{klmn}(v) = (v/v_0)^{1/3}[(k + \Delta k_n)\vec{u}_0 + (l + \Delta l_n)\vec{v}_0 + (m + \Delta m_n)\vec{w}_0]. \tag{C.14}$$

The specific potential energy is then

$$e_c(v) = \frac{1}{2}\sum_{n'}\sum_{klmn} V_{nn'}([v/v_0]^{1/3}|(k{+}\Delta k_n{-}\Delta k_{n'})\vec{u}_0{+}(l{+}\Delta l_n{-}\Delta l_{n'})\vec{v}_0{+}(m{+}\Delta m_n{-}\Delta m_{n'})\vec{w}_0|) \tag{C.15}$$

where n' runs over the s atoms in the reference cell, and the contribution $n = n'$ when $k = l = m = 0$ is omitted (as it corresponds to $i = j$).

In the discussion below, the equations can be applied easily to the case where a single interatomic potential $V(r)$ is used by ignoring the indices nn', i.e. collapsing this matrix to a single scalar. This is the procedure for elements of any structure.

It is assumed $V_{nn'}(r)$ can be expanded as a linear combination of radial functions $f_{nn'j}(r)$

$$V_{nn'}(r) = \sum_j b_{nn'j} f_{nn'j}(r) \tag{C.16}$$

for some set of parameters $b_{nn'j}$. Given a set of n_f fitting functions $f_{nn'j}$ (and hence n_f unknown values of $b_{nn'j}$) and a value of the specific potential energy e_c at some specific volume v, e_c can be expressed as the inner product of the unknown $b_{nn'j}$ with summations of the $f_{nn'j}$ over the position of each atom:

$$e_c(v) = \frac{1}{2}\sum_{n'j}\left[\sum_{klmn} f_{nn'j}(r_{klmn}(v))\right] b_{nn'j} / \sum_{n'} m_{n'}. \tag{C.17}$$

If the specific energy is known at n_f or more specific volumes v_i, a matrix equation is obtained:

$$\sum_{nn'j} A_{inn'j} b_{nn'j} = E_i \tag{C.18}$$

where

$$A_{inn'j} = \sum_{n'} \sum_{klmn} f_{nn'j}(r_{klmn}(v_i)) \tag{C.19}$$

$$E_i = 2e_c(v_i) \sum_{n'} m_{n'}. \tag{C.20}$$

Note that the matrix of rank 3 spanned by the indices $nn'j$ is considered without loss of generality to form a column vector with the same number of elements.

The equation can only be solved for a primitive lattice cell, because the matrix A is singular for non-primitive cells – columns with different nn' are equal and thus (trivially) linearly dependent. If a non-primitive cell is used, the summations over n' in the energy equation should be performed only for the irreducible set of atoms. In the case of elements where all atoms are in the same charge state, this means that the summation is performed with respect to a single atom, at the origin say. Note that for a radial pair potential, 'primitive' must be taken to refer to the radial pair correlation function, and must not take into account differences in relative angles between neighbouring atoms. If the primitive lattice cell does have more than type of atom or charge state, it is possible in principle to deduce different forms of interatomic pair potential between atoms of different type, though it may be necessary to constrain the basis functions of the different potentials to be unique to prevent A becoming singular.

If e_c is known at more than n_f values of v A least squares approach to determining the b_j can be found by making a trivial change to the scheme. Instead of picking n_f values of v and the corresponding values of e_c to give a square matrix A_{ij}, use as many values of v as are available, n_v say. The energy equation then becomes an overdetermined equation, with n_v values of E_i and a rectangular A_{ij} having n_v rows and n_f columns. The 'Givens rotations' method for solving a system of linear equations works equally well for an overdetermined as an exactly determined system, in which case the solution calculated can be shown to satisfy the 'normal equations' of least-squares fitting with linear coefficients [105].

C.3.1 Structure factors

The process of calculating $A_{inn'j}$ can be streamlined if the $f_{nn'j}$ obey a scaling law in r:

$$f_{nn'j}(\alpha r) = f_{nn'j}(\alpha) f_{nn'j}(r). \tag{C.21}$$

The summations can then be performed with respect to a scale-invariant lattice, which need be done only once for each fitting function on each lattice type, to produce a set of structure factors $F_{nn'j}$:

$$F_{nn'j} = \sum_{klmn} f_{nn'j}(\lambda_{klmnn'}) \tag{C.22}$$

where $\lambda_{klmnn'}$ is the dimensionless distance of atom n in lattice cell k, l, m from atom n' in cell $0, 0, 0$:

$$\lambda_{klmnn'} = |(k+\Delta k_n - \Delta k_{n'})\vec{u}_0 + (l+\Delta l_n - \Delta l_{n'})\vec{v}_0 + (m+\Delta m_n - \Delta m_{n'})\vec{w}_0|. \tag{C.23}$$

The elements of the fitting matrix can then be found for any set of specific volumes v_i using

$$A_{inn'j} = f_{nn'j}(a(v_i)) F_{nn'j}, \tag{C.24}$$

where

$$a(v_i) = (v_i/v_0)^{1/3}. \tag{C.25}$$

Alternative conventions are possible for the definition of the structure factors, as discussed below. The fitting functions $f_{nn'j}$ used so far in this scheme have been the powers of r^{-1}:

$$f_{nn'j}(r) = r^{-(j+3)}, \tag{C.26}$$

which obey the scaling law.

The original proposal was to calculate the structure factors F_j by summing over atoms in a sphere of radius G lattice cells, with a suggested value of $G \sim 15$. The cell indices then satisfy

$$k^2 + l^2 + m^2 \leq G^2 \tag{C.27}$$

The time to calculate each F_j is proportional to G^3, so it is desirable to use the smallest G necessary for a given accuracy.

Values of F_j for the smaller (negative) powers of r converge more slowly, since

Table C.1. Comparison of original and modified structure factor scheme.s

G	F_1 (original)	F_1 (modified)
10	96.442541061199	101.333921648953
15	98.057281784164	101.343940983543
25	99.362817190888	101.349903307008
50	100.353192100431	101.352539439707
100	100.851866806305	101.353047713695
200	101.102215598279	101.353176506389

the corresponding $f_j \to 0$ more slowly as $r \to \infty$. Numerical investigations showed that it takes an impracticably long time for the summation in r^{-4} (i.e. F_1) to converge, as shown in Table 1 for a face-centred cubic (FCC) lattice.

One possible modification is to make a continuum approximation outside the sphere defined by $k^2 + l^2 + m^2 = G^2$:

$$F_j \simeq \sum_{klmn} f_j(\lambda_{klmn}) + 4\pi s \int_G^\infty \lambda^2 f_j(\lambda)\, d\lambda \tag{C.28}$$

where the summation is restricted to points inside the sphere. Care should be taken to avoid counting the contribution from atoms in the region where $k^2 + l^2 + m^2 \sim G^2$ twice or not at all. To make the counting logic simpler, k, l and m can be taken over a region greater than the sphere, and the contribution from each atom used only if $\lambda_{klmn} \leq G$.

For fitting functions of the inverse power form, the structure factors become

$$F_j \simeq \sum_{klmn} \lambda_{klmn}^{-(j+3)} + \frac{4\pi s}{j-3} G^{3-j}. \tag{C.29}$$

The corresponding calculations using this new scheme were found to converge more rapidly and provide a greater number of significant figures for a given value of G (Table C.1).

C.3.2 Using derivatives of the specific energy

Some models can supply derivatives of e_c with respect to v. The derivatives can be used in a similar way to e_c in determining parameters for the radial functions f_j used to fit the interatomic potential.

For compactness, derivatives with respect to a function's single independent variable are denoted by priming (e.g. e_c'), no matter which independent variable is used in practice. The context makes it clear whether this is with respect to v, r etc.

The first two differentials of the energy equation with respect to r are

$$e_c'(v) = \frac{1}{2m_a} \sum_{\text{pairs}} r'(v)V'(r) \tag{C.30}$$

$$e_c''(v) = \frac{1}{2m_a} \sum_{\text{pairs}} [r'(v)]^2 V''(r) + r''(v)V'(r) \tag{C.31}$$

where

$$r_{klmn}'(v) = \frac{1}{3}(sm_a)^{1/3} v^{-2/3} \lambda_{klmn} = \frac{1}{3}\frac{r_{klmn}}{v} \tag{C.32}$$

$$r_{klmn}''(v) = -\frac{2}{9}(sm_a)^{1/3} v^{-5/3} \lambda_{klmn} = -\frac{2}{9}\frac{r_{klmn}}{v^2}. \tag{C.33}$$

Expanding V in terms of functions f_j gives two further equations:

$$e_c'(v) = \frac{1}{2m_a} \sum_j \left[\sum_{klmn} \frac{1}{3} \frac{r_{klmn}(v)}{v} f_j'(r_{klmn}) \right] b_j \tag{C.34}$$

$$e_c''(v) = \frac{1}{2m_a} \sum_j \left[\sum_{klmn} \frac{1}{9} \left(\frac{r_{klmn}(v)}{v} \right)^2 f_j''(r_{klmn}) - \frac{2}{9} \frac{r_{klmn}(v)}{v^2} f_j'(r_{klmn}) \right] b_j. \tag{C.35}$$

Again, these can be regarded as (possibly overdetermined) matrix equations if data are available for several specific volumes v_i:

$$\sum_j A_{ij}^{[1]} b_j = E_j^{[1]} \tag{C.36}$$

$$\sum_j A_{ij}^{[2]} b_j = E_j^{[2]} \tag{C.37}$$

where

$$A_{ij}^{[1]} = \sum_{klmn} \frac{1}{3} \frac{r_{klmn}(v_i)}{v_i} f_j'(r_{klmn}(v_i)) \tag{C.38}$$

$$E_i^{[1]} = 2m_a e_c'(v_i) \tag{C.39}$$

$$A_{ij}^{[2]} = \sum_{klmn} \frac{1}{9} \left(\frac{r_{klmn}(v_i)}{v_i} \right)^2 f_j''(r_{klmn}(v_i)) - \frac{2}{9} \frac{r_{klmn}(v_i)}{v_i^2} f_j'(r_{klmn}(v_i)) \tag{C.40}$$

$$E_i^{[2]} = 2m_a e_c''(v_i). \tag{C.41}$$

If the f_j obey a scaling law as above, the matrix elements can again be found in a more efficient form. Again using

$$\begin{aligned}
r_{klmn} &= a\lambda_{klmn} = (sm_a v)^{1/3}\lambda_{klmn} && \text{(C.42)}\\
r'_{klmn} &= \frac{1}{3}(sm_a)^{1/3}v^{-2/3}\lambda_{klmn} && \text{(C.43)}\\
f'_j(r_{klmn}) &= \frac{f_j(\lambda_{klmn})}{\lambda_{klmn}}f'_j(a) && \text{(C.44)}\\
f''_j(r_{klmn}) &= \frac{f_j(\lambda_{klmn})}{\lambda^2_{klmn}}f''_j(a) && \text{(C.45)}
\end{aligned}$$

we find that

$$\begin{aligned}
A^{[1]}_{ij} &= \frac{1}{3}(sm_a)^{1/3}v^{-2/3}f'_j(a(v_i))F_j && \text{(C.46)}\\
A^{[2]}_{ij} &= \frac{1}{9}(sm_a)^{1/3}v^{-4/3}\left[(sm_a)^{1/3}f''_j(a(v_i)) - 2v^{-1/3}f'_j(a(v_i))\right]F_j. && \text{(C.47)}
\end{aligned}$$

These matrix equations are additional relations to use in determining the b_j. They can be added as extra rows to the fitting equation. The resulting equation can be used to find n_f values of b_j given e_c, de_c/dv and d^2e_c/dv^2 at $n_f/3$ specific volumes, or as an even more overdetermined system (with $3n_v$ rows) for least-squares fitting.

C.3.3 Orthogonalised fitting functions

It is often useful when fitting data to use an orthogonal set of functions. If the fitting coefficients are calculated using a subset of the available functions, the coefficients of these functions should not change when additional functions are used in the fit.

Given a set of functions f_j and a formula for the scalar product $\langle f_i, f_j\rangle$, an orthogonal set of basis functions g_j can be constructed [47] such that $\langle g_i, g_j\rangle = 0$ when $i \neq j$. An important point is that the space over which the scalar product is defined should be the one most relevant to the problem (hence the ancient proverb 'bases for spaces'). Possible choices for a scalar product space here are

$$\langle f_i, f_j\rangle = \int_{r_{\min}}^{\infty} f_i f_j \, dr \qquad \text{(C.48)}$$

and

$$\langle f_i, f_j \rangle = \int_{r_{\text{min}}}^{\infty} f_i r^2 f_j \, dr, \tag{C.49}$$

both of which depend only on r_{min} for functions like r^{-n}. More defensible choices considering where the fitting functions are actually evaluated would seem to be the discrete space

$$\langle f_i, f_j \rangle = \sum_{klmni'} f_i(r_{klmn}(v_{i'})) f_j(r_{klmn}(v_{i'})) \, dr \tag{C.50}$$

or the corresponding mixed space with contributions from the region of continuum approximation. However, when trial calculations were made with fitting functions successfully orthogonalised over the space above, the fitting coefficients changed as successive functions were added, suggesting that even this space is not appropriate.

If the use of orthogonalised functions turns out to be necessary, it is still possible to streamline the calculation of components in the fitting matrix by using structure factors. Suppose we are now looking for fitting coefficients c_j such that

$$V(r) = \sum_j c_j g_j(r) \tag{C.51}$$

where

$$g_j(r) = \sum_k \alpha_{jk} f_k(r) \tag{C.52}$$

and α_{ij} is the transformation matrix between original and orthogonal bases. The (possibly overdetermined) system for the amplitudes of the basis functions is then

$$\sum_j B_{ij} c_j = E_i \tag{C.53}$$

where the matrix B_{ij} can still be calculated using structure factors of the f_j:

$$B_{ij} = \sum_{klmn} g_j(r_{klmn}(v_i)) = \sum_{k'} \alpha_{jk'} f_{k'}(a(v_i)) \sum_{klmn} f_{k'}(\lambda_{klmn}) = \sum_{k'} \alpha_{jk'} f_{k'}(a(v_i)) F_{k'}. \tag{C.54}$$

C.3.4 Conventions for structure factors

If the radial basis functions $f_j(r)$ obey a scaling law as above, structure factors can be calculated once for each function over each lattice type, thus avoiding the need to perform summations over the lattice for each specific volume.

The structure factor defined above is the simplest possible form:

$$F_j^{[a]} = \sum_{klmn} f_j(\lambda_{klmn}). \tag{C.55}$$

This is then related to the fitting matrix A_{ij} by

$$A_{ij} = f_j(a(v_i))F_j^{[a]} \tag{C.56}$$

where the cubic lattice parameter is found from

$$a(v) = (sm_a v)^{1/3}. \tag{C.57}$$

Other forms are possible (and all completely equivalent), depending on how many of the terms it is desired to take into the summation from the ratio of r_{klmn} to λ_{klmn}, i.e. what part of the factor $(sm_a v)^{1/3}$, and any other scaling factors.

Another common choice [40] expresses dimensionless distances as multiples of the nearest-neighbour distance r_{neigh} rather than the lattice spacing a. Defining

$$\gamma \equiv \frac{r_{\text{neigh}}}{a}, \tag{C.58}$$

which is a function only of the lattice type, the structure factors have the form

$$F_j^{[c]} = \sum_{klmn} f_j(\lambda_{klmn}/\gamma), \tag{C.59}$$

giving the fitting matrix

$$A_{ij} = f_j(r_{\text{neigh}}(v_i))F_j^{[c]} \tag{C.60}$$

where

$$r_{\text{neigh}}(v) = \gamma(sm_a v)^{1/3}. \tag{C.61}$$

The following relations are useful in converting between the forms:

$$\begin{aligned} F_j^{[a]} &= f_j(\gamma)F_j^{[c]} & \text{(C.62)} \\ F_j^{[b]} &= f_j(s^{1/3}\gamma)F_j^{[c]} = f_j(s^{1/3})F_j^{[a]}. & \text{(C.63)} \end{aligned}$$

Table C.2. Values of $F_j^{[c]}$ for $f_j(r) = r^{-(j+3)}$.

j	$f_j(r)$	$F_j^{[c]}$ (SC)	$F_j^{[c]}$ (BCC)	$F_j^{[c]}$ (FCC)
1	r^{-4}	16.53	22.64	25.34
2	r^{-5}	10.38	14.76	16.97
3	r^{-6}	8.40	12.25	14.45
4	r^{-7}	7.47	11.05	13.36
5	r^{-8}	6.95	10.36	12.80
6	r^{-9}	6.63	9.89	12.49
7	r^{-10}	6.43	9.56	12.31
8	r^{-11}	6.29	9.31	12.20
9	r^{-12}	6.20	9.11	12.13
10	r^{-13}	6.14	8.95	12.09
11	r^{-14}	6.10	8.82	12.06
12	r^{-15}	6.07	8.70	12.04
13	r^{-16}	6.05	8.61	12.03
≥ 14	$r^{-(j+3)}$	$6 + 12(1/2)^{(j+3)/2}$	$8 + 6(3/4)^{(j+3)/2}$	$12 + 6(1/2)^{(j+3)/2}$

Table C.2, taken from the literature [40] lists values of $F_j^{[c]}$ for simple cubic (SC), body-centred cubic (BCC) and face-centred cubic (FCC) structures when the f_j are powers of r^{-1}. Note that the $F_j^{[c]}$ tend to the coordination number (the number of nearest neighbours) as the power increases.

(The formulægiven for $j \geq 14$ give results to the same accuracy as the rest of the table.)

For comparison, the value of $F_1^{[a]}$ calculated for FCC in Table C.1 (101.353) implies $F_1^{[c]} = 25.338$.

Appendix D

Direct simulation of ensembles of atoms

D.1 Introduction

Direct numerical simulation of the interaction between atoms allows complicated behaviour (such as lattice anharmonicity) to be investigated by starting with simple real-space concepts. Large numbers of calculations are required, making the process slow, and the results must be interpreted statistically, introducing noise, but the accuracy of the simulations can be measured and understood easily.

Consider a set of particles at positions $\vec{r}_i$. The particles interact with one another, such that the total energy is a function of the instantaneous positions of the particles, $H(\{\vec{r}_i\})$.

In the case of atoms, the interaction may be calculated rigorously by finding the wavefunction of the electrons. However, direct numerical simulation tends to be used for very large ensembles of atoms where a proper treatment of the electrons is not practicable. It is more usual to employ an interatomic ('interparticle') potential, $V(\{\vec{r}_i - \vec{r}_j\})$.

For simulations of condensed matter it is usually necessary to model a system many orders of magnitude larger than the number of atoms represented explicitly. Periodic boundary conditions are used to model an infinite system.

In condensed matter, the atoms exhibit co-operative motion such as phonon modes. Even with periodic boundary conditions, the types of co-operative motion which can be simulated are restricted by the number of atoms modelled explicitly.

In the case of phonons, this limitation introduces an artificial cutoff on the wavelength. As the number of atoms in the simulation increases, the states approach those in a macroscopic sample of material. The accuracy of a simulation can be tested in a straightforward way by repeating the simulation with a larger number of atoms N, e.g. a $2 \times 2 \times 2$ supercell of the original simulation. By estimating the variation in an observable O (e.g. the specific internal energy) with N, e.g. $dO/d(1/N)$, for a few different values of N the result can be extrapolated to an infinite system ($1/N \to 0$).

Similarly, observables vary during the course of a simulation, eventually fluctuating about a mean value when the system equilibrates. To obtain the expectation value of an observable O, the system is allowed to reach equilibrium (measured from the history of O), and then the simulation is continued to allow O to be determined with an acceptable statistical uncertainty. This process can break down near phase transitions, because the system may remain in one phase for a long period, then suddenly change phase with an accompanying abrupt change in O.

For the EOS work, typical simulations consisted of taking a fixed volume containing a fixed number of atoms, and simulating the system as it equilibrated at some desired temperature.

If the particles are classical objects, then the total energy can be split simply into potential and kinetic contributions,

$$H = E_p(\{\vec{r}_i\}) + E_k. \tag{D.1}$$

In equilibrium at a temperature T, the classical equipartition theorem [2] allows the mean kinetic energy to be calculated simply as $3k_BT/2$ per atom.

D.2 'Molecular' dynamics

In molecular dynamics (which in the author's opinion would be better called 'particle dynamics'), equations of motion are developed for the particles, and their movement simulated as a function of time.

The force on each particle can be found from the principle of virtual work:

$$\vec{f}_i = -\frac{\partial H(\{\vec{r}_j\})}{\partial \vec{r}_i}. \tag{D.2}$$

Given an interatomic potential or electronic charge distribution, the $\vec{f}_i$ can be obtained more directly for each atom. The equations of motion are then

$$\frac{\partial^2 \vec{r}}{\partial t^2} = \frac{\vec{f}_i}{m_i}, \tag{D.3}$$

where m_i is the mass of particle i.

The equations of motion can be integrated using any solver for coupled partial differential equations. In the present work, a predictor-corrector scheme [84] was used. An overall second order update over a time interval δt is obtained from a first order forward-time prediction over $\delta t/2$ followed by a leapfrog step over the full δt:

$$\begin{aligned} \tilde{F}(t+\delta t/2) &= F(t) + \frac{\delta t}{2}\dot{F}[F(t)] && \text{(D.4)} \\ F(t+\delta t) &= F(t) + \delta t \dot{F}[\tilde{F}(t+\delta t/2)] && \text{(D.5)} \end{aligned}$$

where F is the set of position vectors, $\tilde{F}$ their first order estimates and $\dot{F}$ the derivatives calculated using the position vectors at the appropriate time level.

The time step to use in the integration can be estimated from the highest phonon frequency ω_{max} for a given compression. The time step should be small enough for a period of the vibration to take several steps, N_s say ~10. Thus

$$\delta t_{\text{max}} = \frac{2\pi N_s}{\omega_{\text{max}}}. \tag{D.6}$$

As the simulation proceeds to an equilibrium state, the balance between kinetic and potential energy changes. A perfect numerical scheme would conserve total energy exactly, so the effective temperature, measured by the mean kinetic energy, alters. If the objective is to perform a simulation at some particular temperature, then a correction to the energy is needed to constrain the temperature. Furthermore, the predictor-corrector scheme does not conserve energy exactly. In fact, in a compressed state the potential energy tends to rise with simulation time because of the stiffness of the interatomic potential. Thus it is essential to use a numerical 'thermostat' to hold the temperature to a steady level.

Various thermostats have been devised [12]. In the present work, the velocity vectors $\vec{u}_i$ are rescaled by a global factor ξ every N_r time steps (where N_r ~10 to 100):

$$\xi = \sqrt{\frac{3Nk_BT}{2\bar{E}_k}} \tag{D.7}$$

where $\bar{E}_k$ is the mean kinetic energy of the ensemble of atoms,

$$\bar{E}_k = \frac{1}{N}\sum_i \frac{1}{2}m_i|\vec{u}_i|^2. \tag{D.8}$$

A Gaussian frictional thermostat [106] was tried for comparison.

D.3 Monte-Carlo

Monte-Carlo simulations sample different configurations of the particles $\{\vec{r}_i\}$ with no real time integration of the equations of motion. The kinetic energy is taken to be $3k_BT/2$ per particle. Transitions are made between different random static configurations with a probability chosen so that the potential energy gradually tends to an equilibrium value.

Consider a configuration $\{\vec{r}_i\}$ with potential energy E_p and an alternative configuration $\{\vec{r}_i'\}$ with potential energy E_p'. At a temperature T, the probability that a given configuration (or microstate) will be found is

$$Pr(\{\vec{r}_i\}) \propto e^{-E_p/k_BT}. \tag{D.9}$$

Thus the ratio between the probabilities of the alternative microstates is

$$\frac{Pr(\{\vec{r}_i'\})}{Pr(\{\vec{r}_i\})} = e^{-\Delta E_p/k_BT}, \tag{D.10}$$

where $\Delta E_p = E_p' - E_p$.

A variant of the Metropolis algorithm [12] was used to control the sequence of microstates. Given a configuration $\{\vec{r}_i\}$ and a possible new configuration $\{\vec{r}_i'\}$, the new configuration is always accepted if $E_p' \leq E_p$, and accepted with a probability

$$Pr(\{\vec{r}_i\} \to \{\vec{r}_i'\}) = e^{-\Delta E_p/k_BT} \tag{D.11}$$

if $E_p' > E_p$. To 'accept with a probability Pr,' a random number ν with uniform distribution in the interval $0 \leq \nu < 1$ (from a standard random number generator [51]) and the new configuration accepted if $\nu \leq Pr$.

For each accepted microstate, the next candidate state is generated by adding random perturbations to the position vectors:

$$\vec{r}_i' = \vec{r}_i + \Delta\vec{r}_i. \tag{D.12}$$

Many prescriptions are possible for choosing the $\Delta\vec{r}_i$. In the present work, each component was chosen according to the formula

$$[\Delta\vec{r}_i]_j = 2\left(\nu_j - \frac{1}{2}\right)\delta_j \tag{D.13}$$

where the ν_j are random numbers with uniform distribution in the interval $0 \leq \nu < 1$ and δ_j are scale lengths for each co-ordinate direction (usually equal). The effect of the $2(\nu_j - \frac{1}{2})$ is to produce random numbers with uniform distribution in the interval $-1 \leq \nu < 1$. The random perturbations $\Delta\vec{r}_i$ thus lie in a cuboid of sides $2\delta_j$ centred on the $\vec{r}_i$.

It would be more aesthetically pleasing (and rotationally invariant) to choose perturbations with spherical or ellipsoidal symmetry rather than the cuboids used. Similarly, a different distribution could be used as an alternative to the uniform one chosen, for instance a Gaussian form allowing large perturbations with a lower probability than small ones. However, the cuboidal scheme is computationally more efficient than these alternatives. It is probably more important to sample a large set of microstates than to make minor adjustments to the distribution form of the perturbations.

The perturbation scales δ_j determine the proportion of microstate jumps accepted or rejected. If the perturbations are large, many of them will be rejected (because some atoms will move too close together). If small, a large number of jumps will be taken to sample the same range of energies. The δ_j were adjusted automatically to give reasonable acceptance and rejection rates. If the rejection rate $\mathcal{R}_r$ was less than some minimum $\mathcal{R}_1$, the δ_j were multiplied by a scaling factor σ to give larger jumps. If the rejection rate $\mathcal{R}_r$ exceeded some maximum $\mathcal{R}_2$, the δ_j were divided by σ to give smaller jumps. Typically, $\mathcal{R}_1 \sim 0.3$, $\mathcal{R}_2 \sim 0.7$ and $\sigma \sim 2$.

D.4 Faster calculations

Using interatomic potentials, a simulation with N interacting atoms requires N^m calculations of an m-body term in the potential, thus the computational time scales like N^m, making large calculations expensive. If the potential has finite range r_c so

$$V(r) \simeq 0 \quad : \quad r > r_c, \tag{D.14}$$

then the total computational time can be reduced by splitting the spatial domain of the simulation into cells of size r_c or greater. When calculating the potential energy, the contribution from atom i in cell k involves a sum over only the other atoms in cell k and the atoms in the cells adjacent to k (Fig. D.1). In 3D with cubic cells, only the atoms in 27 cells need to be considered towards the contribution from each atom. The figure of 27 cells remains constant as the total number of cells – and therefore atoms – is changed. Thus the computational effort scales as N rather than N^m.

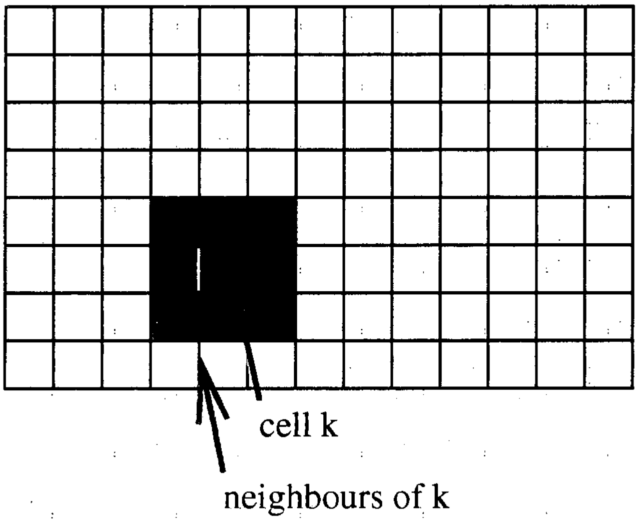

Figure D.1. Decomposition of the region of simulation into cells of the range of the interatomic potential.

Some book-keeping is required to keep track of the list of atoms in each cell, but this is straightforward if the cells have a simple shape. For instance, cuboids are convenient if parallel with the co-ordinate axes because the test if 'in-ness' is just that each component of an atom's position vector lies between the limits corresponding to the extent of a particular cell. The procedure for updating the list of atoms can also be made efficient, by testing cells in the following order:

1. The cell previously containing the atom.
2. The 6 cells sharing a face with the previous cell.
3. The 12 cells sharing an edge with the previous cell.
4. The 8 cells sharing a corner with the previous cell.
5. Cells in the next shell further out.

etc. A simpler alternative method is:

1. The cell previously containing the atom.
2. The 26 cells adjacent to the previous cell.
3. The rest of the cells.

The list is updated every few iterations. If cells are made larger than r_c then updates can be made less frequently. However, the number of atoms included in the calculation of each contribution to the potential energy is correspondingly greater.

Periodic boundary conditions are straightforward to implement using this domain decomposition.

Appendix E

Crystallographic diffraction patterns

E.1 Introduction

One of the theoretical techniques used in this project was the direct simulation of the motion of a collection of atoms, using molecular dynamics or Monte-Carlo methods to predict their evolution as a function of temperature. These simulations produce a large amount of data, which can be difficult to interpret. In particular, it is difficult to discover by inspection of the positions of the atoms whether a phase transformation has occurred, because the new phase may not be aligned conveniently with the co-ordinate axes.

An alternative technique used here is the prediction of the powder diffraction pattern which would be produced by the ensemble of atoms, assuming periodic boundary conditions. In the formation of a powder pattern, information on the orientation of individual crystals in the sample is lost. This can be viewed as a form of data reduction technique, as the resulting patterns can be analysed more simply than the atom positions themselves to identify the presence of common crystal structures.

The theory of X-ray diffraction is usually presented with emphasis on the analysis of an experimentally-measured pattern to obtain an unknown crystal structure. Here we summarise diffraction theory from the point of view of the prediction of a powder diffraction pattern given the position of the atoms. The basic results employed have been published previously by other researchers [107].

E.2 Scattering from a single atom

A single isolated atom scatters photons with an intensity which varies with direction. For X-ray diffraction work, photon wavelengths are of the order of 1 Å. In this regime, various simplifying assumptions can be made. Firstly, it is reasonable to consider purely elastic phonon scattering. Secondly, the scattering effect is dominated by the electrons only, since the classical amplitude scattered from a point particle of mass m and charge q at a distance r is

$$\frac{A_0}{r}\frac{q^2}{mc^2}\sqrt{\frac{1+\cos^2 2\theta}{2}}, \tag{E.1}$$

where A_0 is the incident amplitude and 2θ the angle between the incident beam (forward direction) and the scattered beam, and hence the contribution from the nucleus is much smaller.

Scattering from a 'real' atom includes contributions resulting from the detailed orbital states of all the electrons. However, this can be estimated quite accurately from the charge density distribution, i.e. by integrating over the contribution from each point of space with its associated electronic charge density. Because of phase differences between scattering from different points, this results in a scattering intensity which varies with scattering angle.

It turns out that the ratio between the actual scattering amplitude in a given direction and the value for a classical electron is modelled well by a function of $\sin\theta/\lambda$ (where λ is the wavelength of the X-ray photons). This ratio is called the 'atomic scattering factor', f, which varies from a maximum close to the atomic number Z for $\sin\theta/\lambda = 0$.

It should be emphasised that curves of f vary between elements and also states of a given element, e.g. different states of ionisation.

E.3 Coherent scattering from a crystal

Calculation of the coherent scattering of X-rays by a crystal is made simplest by considering the contribution from each lattice plane. In this instance, 'lattice plane' refers to planes oriented with respect to the Bravais lattice of repeating three-dimensional 'tiles' which make up the crystal. The effect of the lattice basis (the atoms which constitute the lattice cell) is taken into account later.

It is conventional to represent each lattice plane by a set of integers (hkl).

The corresponding plane (passing through a set of lattice nodes) is that which intercepts the lattice vectors at $\vec{a}/h$, $\vec{b}/k$ and $\vec{h}/l$. (This notation is thus independent of the relative orientations of the lattice vectors $\vec{a}$, $\vec{b}$ and $\vec{c}$.) Given the contribution from a given lattice plane (intensity + angle), the contributions can be accumulated over all possible lattice planes, allowing a diffraction pattern to be built up.

Any given set of lattice planes (hkl) will diffract photons if they satisfy the (Laue or) Bragg condition:

$$\lambda = 2d_{hkl}\sin\theta. \tag{E.2}$$

The interplanar spacing d_{hkl} can be found as follows in the case of a general set of lattice vectors, not necessarily orthogonal or of equal length:

$$\begin{aligned}
\cos\alpha &= \frac{\vec{b}.\vec{c}}{|\vec{b}||\vec{c}|} \quad \text{(similarly for } \beta,\ \gamma\text{)} && \text{(E.3)}\\
\cos\alpha^* &= \frac{\cos\beta\cos\gamma - \cos\alpha}{\sin\beta\sin\gamma} \quad \text{(similarly for } \beta^*,\ \gamma^*\text{)} && \text{(E.4)}\\
|\vec{a}^*| &= \frac{1}{\sin\beta\sin\gamma^*} \quad \text{(similarly for } \vec{b}^*,\ \vec{c}^*\text{)} && \text{(E.5)}\\
d^{*2}_{hkl} &= h^2|\vec{a}^*|^2 + 2hl|\vec{b}^*||\vec{c}^*|\cos\alpha^* + \quad \text{corresponding terms} && \text{(E.6)}\\
d_{hkl} &= \frac{1}{d^*_{hkl}}, && \text{(E.7)}
\end{aligned}$$

where α is the angle between lattice vectors $\vec{b}$ and $\vec{c}$, where α^* is the angle between reciprocal lattice vectors $\vec{b}^*$ and $\vec{c}^*$, and d^*_{hkl} is the lattice spacing in reciprocal space.

This analysis is valid for any lattice of repeating cells, whether or not they have a motif of several atoms. In general, the intensity of X-rays scattered by plane (hkl) is given by the sum over the scattering factors of the constituent atoms, taking into account the phase factor between radiation originating at different points within the lattice cell (which leads to systematic absences e.g. for BCC and FCC lattices):

$$I = \sum_{\text{atoms } i} f_i(\theta,\lambda)e^{2\pi i\vec{p}.\vec{r}_i} \tag{E.8}$$

where $\vec{p}$ is the plane index (hkl), $\vec{r}_i$ the position of atom i in the lattice cell, with respect to the general lattice vectors $\vec{a}$, $\vec{b}$ and $\vec{c}$ – i.e. its absolute position with respect to the origin of the lattice cell is

$$\vec{R}_i = [\vec{r}_i]_1\vec{a} + [\vec{r}_i]_2\vec{b} + [\vec{r}_i]_3\vec{c} \tag{E.9}$$

– and f is the atomic scattering factor for atom i at angle θ and photon wavelength λ.

E.4 Software implementation

Software was written to implement all the physical processes described above, except for the classical electron scattering amplitude. For greatest flexibility, the computer program uses atomic scattering factors in tabular form. On calculating the intensity of reflection at any angle, the classical factor

$$r\frac{q^2}{mc^2}\sqrt{\frac{1+\cos^2 2\theta}{2}}, \tag{E.10}$$

is omitted from the explicit formulation. If desired, it can be incorporated in the atomic scattering factor, or the output simply scaled.

Appendix F

Evaluating against experiment

F.1 Introduction

EOS generated by the theoretical methods described here are somewhat uncertain because of the approximations, simplifications and finite-size effects used. Although the EOS could be adjusted to reproduce the results of experiments of arbitrary complexity, it is more useful to compare with data which tests the separate ingredients of the EOS and the way in which they are combined.

F.2 Phonons

The phonon dispersion relations $\omega(\vec{k})$ and the density of phonon states $g(\omega)$ can be measured by inelastic neutron scattering [40].

F.3 Fermi surface

The shape of the Fermi surface can be obtained from de Haas - van Alphen measurements of the magnetisation of materials at high field strengths [40]. This data provides a way of testing predictions of the electron band structure.

F.4 Equilibrium density

The density at STP can be obtained to great accuracy by a variety of techniques including immersion and X-ray crystallography.

For evaluating *ab initio* EOS, it is very useful to have a measurement of the density or lattice constants near $T = 0$ (and $p = 0$ or 1 atm) to compare with the cold curve. This can be obtained by X-ray crystallography.

F.5 Isotherms

Isotherms can be measured conveniently by static press methods, such as the diamond anvil cell and multi-anvil press. The temperature can be controlled most easily by holding the apparatus in thermal equilibrium at the desired temperature. The lattice type and parameters can be measured by X-ray or neutron diffraction.

These methods generally rely on a calibrant to determine the pressure. The calibrant may be a material of known compressibility whose lattice parameters are measured along with those of the sample of interest, or a material with a known relation between pressure and optical properties, such as the wavelengths of the optical resonances in a ruby.

F.6 Shock Hugoniots

A shock wave is a moving interface between regions of different thermodynamic state. The region ahead of the shock has a density ρ_0, specific internal energy e_0 and therefore pressure p_0 and temperature T_0, and moves with a velocity $\vec{u}_0$ which can be taken to be zero by a Galilean velocity transformation. The region behind the shock has a density ρ_1, specific internal energy e_1 and therefore pressure p_1 and temperature T_1, and moves with a velocity $\vec{u}_1$, where

$$\vec{u}_i = \vec{u}_0 + \vec{U}, \tag{F.1}$$

where $\vec{U}$ is parallel to the direction of propagation of the shock wave. The shock wave can be thought of as the interface between undisturbed material and material pushed by a piston moving at a constant speed. (Figs F.1 and F.2.)

Considering a shock wave as a discontinuity between material in different states (p, e, u_p) (where u_p is the material velocity), conservation of mass, momentum and energy across the shock leads to the Rankine-Hugoniot equations linking

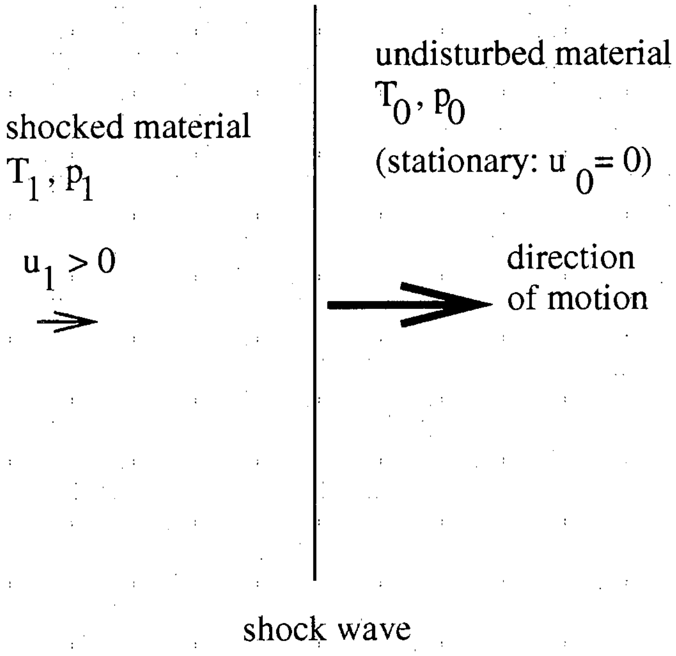

Figure F.1. Idealised shock wave.

the states on either side and the shock velocity u_s: [1]

$$u_s^2 = v_0^2 \frac{p - p_0}{v_0 - v} \tag{F.2}$$

$$u_p = \sqrt{[(p - p_0)(v_0 - v)]} \tag{F.3}$$

$$e = e_0 + \frac{1}{2}(p + p_0)(v_0 - v) \tag{F.4}$$

where subscript '0' denotes material ahead of the shock (with $u_p = 0$). Given the EOS $p(v, e)$ this set of equations can be closed, allowing the Hugoniot (locus of states reached from the initial state by a single shock) to be calculated. Alternatively, the measurement of two quantities in a steady planar shock is sufficient to determine the full mechanical state.

It is possible in principle to measure the density and pressure of the shocked material directly. The density can be determined from radiography, inferring the areal mass over each point in the radiograph from the attenuation of a photon signal. Various types of pressure gauge have been devised, based on the variation of resistance with compression in a wire embedded in the sample, or piezoelectricity. However, the most accurate measurements are usually the shock and

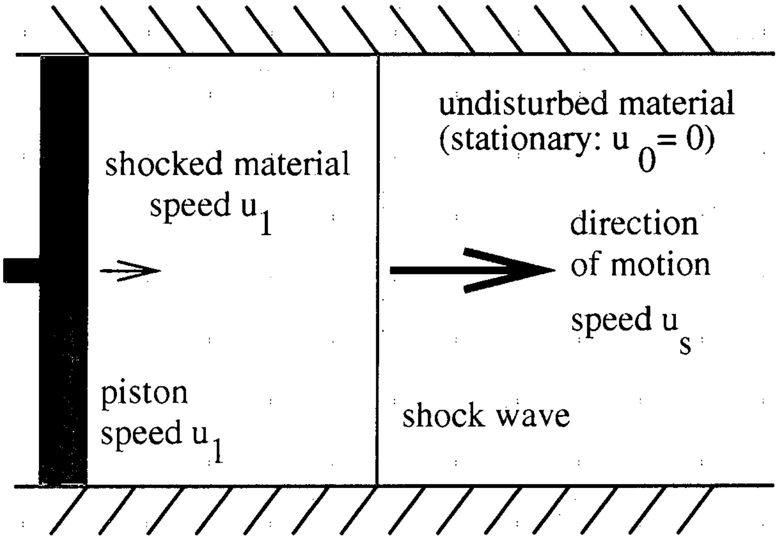

Figure F.2. 1D shock wave produced by a piston.

particle speed. The shock speed can be determined from the rate at which it emerges over an angled surface; the particle speed from the time taken to cross a gap or by the Doppler shift of laser photons reflected from the free surface. The pressure and density can be deduced from the shock and particle speeds, using the Rankine-Hugoniot relations (in effect assuming that the deformation is hydrostatic):

$$\rho = \frac{\rho_0}{1 - u_p/u_s} \tag{F.5}$$

$$p = p_0 + \rho_0 u_p u_s. \tag{F.6}$$

Since shock waves at least commence by a uniaxial deformation of each element of material, the pressure and internal energy in materials with a finite shear modulus will include non-isotropic contributions from the stress tensor: $p = -\mathrm{Tr}\,\sigma/3$ [108]. The strain ϵ caused by a uniaxial compression is of the order of twice the compression $(\rho_0/\rho - 1)$, depending precisely on the model of finite strain taken. The shear modulus of for example aluminium is of the order of 30 GPa [109], but the effect of strength is limited by the onset of plastic flow with a stress of between about 40 and 500 MPa for quasistatic yield [109]. The flow stress depends on the compression and loading rate among other things, so it can be difficult to disentangle the effects of anisotropic stress from hydrodynamic compression. Because static yield points fall in this relatively low stress range, strength effects are usually neglected in analysing Hugoniots [1]. (The validity of this approximation is a possible area for future work.)

Given an EOS of the form $p(\rho, e)$, the shock Hugoniot from the state (ρ_0, e_0) can be constructed as follows:

1. Choose a set of values of ρ at which to calculate the Hugoniot. Note: for 'normal' materials (for which $\partial p/\partial \rho > 0$) ρ must be greater than ρ_0 for a shock wave to exist.

2. For each ρ (in increasing order), increase e until p from the EOS equals p from the Rankine – Hugoniot equations, to some reasonable accuracy. If ρ is spanned in increasing order, then the lower limit for e is the value at the next larger (i.e. the previous) value of ρ. e can thus be found by increasing from the previous value until $p_{\text{EOS}} - p_{\text{RH}}$ changes sign, and then finding the zero of this function by bisection.

Appendix G

Empirical equations of state

G.1 Introduction

Empirical EOS assume a functional form, where the parameters are optimised to reproduce a set of experimental data. The form of an empirical EOS often reflects the data to which it is fitted.

For shock wave work, a convenient source of data is the measured relation between shock speed u_s and particle speed u_p. These parameters are often easiest to measure in an experiment.

G.2 Grüneisen

Grüneisen EOS comprise a functional fit to a reference curve and a model of the mechanical properties away from the reference curve. The Steinberg form of the Grüneisen EOS [70] comprises a polynomial fit to the shock Hugoniot in $u_s - u_p$ space,

$$u_s = c_0 + u_p \sum_{i=0}^{N} s_i \left(\frac{u_p}{u_s}\right)^i, \tag{G.1}$$

and an assumed variation of Grüneisen's γ [40]

$$\gamma = \frac{\gamma_0 + b\mu}{1+\mu} \tag{G.2}$$

where

$$\mu \equiv \frac{\rho}{\rho_0} - 1. \tag{G.3}$$

This leads to a mechanical EOS

$$p(\rho, e) = \begin{cases} \dfrac{\rho_0 c_0^2 \mu \left[1 + \left(1 - \frac{\gamma_0}{2}\right)\mu - \frac{b}{2}\mu^2\right]}{1 - \mu - \sum_{i=0}^{N} s_i \frac{\mu^i}{(\mu+1)^{i-1}}} + (\gamma_0 + b\mu)e & : \quad \mu > 0 \\ \rho_0 c_0^2 \mu + \gamma_0 e & : \quad \mu \leq 0 \end{cases} \tag{G.4}$$

Given an independent method for obtaining the cold curve (such as a theoretical calculation), the Grüneisen parameter $\gamma(v)$ can be deduced using the Dugdale-MacDonald relation [19, 16]

$$\gamma(v) = -\frac{v^2 p_c''/2 + v p_c' + p_c/9}{v p_c' + 2p_c/3} \tag{G.5}$$

where priming denotes differentiation with respect to v.

G.3 Murnaghan

The Murnaghan EOS has the form

$$p = \frac{K_e}{n}\left[\left(\frac{\rho}{\rho_0}\right)^n - 1\right], \tag{G.6}$$

where K_e is the equilibrium bulk modulus. This EOS is a function of ρ only – it does not take account of thermal excitations.

G.4 Rose

The Rose 'universal' EOS form [15][1] has the form

$$p = -3K_e \frac{(v/v_0)^{1/3} - 1}{(v/v_0)^{2/3}}(1 - 0.15a + 0.05a^2)e^{-a}, \tag{G.7}$$

$$a = \frac{r_{WS} - r_{WS}^e}{l}, \tag{G.8}$$

$$\frac{4}{3}\pi r_{WS}^3 = \frac{m_a}{\rho}. \tag{G.9}$$

---

[1] Eq. 1.15 of this reference, defining $p(v)$, appears to be wrong by a factor -1.

Thermal contributions can also be included; these were not needed for the present work.

G.5 Bushman et al

The EOS used by these Russian workers [110] has the form

$$
\begin{aligned}
p &= \begin{cases} \dfrac{v_{0c}}{v_0} \sum a_j \sigma^{j/3} & : \sigma \geq 1 \\ \dfrac{v_{0c}}{v_0} A \left(\sigma^m - \sigma^l\right) + B \left(\sigma^n - \sigma^l\right) & \text{otherwise} \end{cases}, && \text{(G.10)} \\
\sigma &\equiv v_0/v, && \text{(G.11)} \\
\frac{1}{m} &= \frac{1}{A}\left[B\left(\frac{1}{l} - \frac{1}{n}\right) - \frac{e_s}{v_{0c}}\right] + \frac{1}{l}, && \text{(G.12)}
\end{aligned}
$$

where v_{0c} is the equilibrium density at absolute zero and e_s the sublimation energy. Thermal contributions can also be included; these were not needed for the present work.

G.6 Schulte and Holzapfel

The 'H11' form of EOS [111] is

$$
\begin{aligned}
p &= p_0 x^{-5}(1-x)e^{-cx} && \text{(G.13)} \\
x &\equiv (v/v_0)^{1/3} && \text{(G.14)} \\
c &= \ln(p_0/3K_e) && \text{(G.15)} \\
p_0 &= a\left(\frac{Z}{v_{WS}}\right)^{5/3} && \text{(G.16)} \\
a &= (3\pi^2)^{2/3}\frac{\hbar^2}{5m_e} && \text{(G.17)} \\
v_{WS} &= m_a/\rho_0, && \text{(G.18)}
\end{aligned}
$$

where m_a and m_e are the mass of an atom and an electron respectively. This EOS is a function of ρ only – it does not take account of thermal excitations.

Appendix H

Modified Gram – Schmidt Orthogonalisation

H.1 Introduction

Often in mathematics it is desirable to orthogonalise a set of functions. This occurs in data fitting for example, when fitting may not work well for the original set because changes to one function affect another, or when adding another non-orthogonal function requires the whole least-squares process to be repeated. The Gram – Schmidt method [47] is one possible and effective way of creating an orthogonal set of functions (or vectors) from a non-orthogonal set.

For example, the set of polynomials x^i are not orthogonal. If you find the coefficients for best-fitting straight line to some data and then decide to increase the accuracy of the fit by adding a term in x^2, the coefficients of the straight line will change. If on the other hand you used the Chebyshev (or Tschebycheff) polynomials, which *are* orthogonal, adding higher-order terms leaves the old terms unaffected.

The Gram – Schmidt method allows any set of functions to be combined into orthogonal forms. This note describes a modified form of the Gram – Schmidt method which is more efficient and easier to code than the straightforward form.

H.2 Scalar product and orthogonal functions

The scalar product of two functions ψ_1 and ψ_2 is the integral of their product over some space for which both are defined. The space could be for $x \in [0,1] \subset \mathcal{R}$

(a continuous interval), giving a scalar product

$$\langle\psi_1|\psi_2\rangle = \int_0^1 \psi_1(x)\psi_2(x)\,dx; \tag{H.1}$$

or for $x = \{x_i : i \in [1,3] \subset \mathcal{I}\}$ (a set of ordinates), giving a scalar product

$$\langle\psi_1|\psi_2\rangle = \sum_{i=1}^{3} \psi_1(x_i)\psi_2(x_i). \tag{H.2}$$

Just like orthogonal vectors, orthogonal functions are ones with zero scalar product. Given a set of functions ψ_i, they are orthogonal if the matrix of scalar products $\langle\psi_i|\psi_j\rangle = \mathrm{diag}_{ij}\lambda_i$, i.e. all zero apart from the main diagonal of terms of the form $\langle\psi_i|\psi_i\rangle = \lambda_i$. They are *orthonormal* if all the $\lambda_i = 1$.

If the ψ_i are *not* orthogonal, the matrix $\langle\psi_i|\psi_j\rangle$ will have off-diagonal terms.

H.3 Gram – Schmidt method

The essence of the Gram – Schmidt method is to take each function from the non-orthogonal set $\{\psi_i\}$ in turn and subtract off enough of all previous functions to make it orthogonal to them. In this way, an orthogonal set $\{\phi_i\}$ is created.

In practice, the method works as follows:

$$\phi_1 = \psi_1 \tag{H.3}$$

$$\phi_i = \psi_i - \sum_{j=1}^{i-1} \phi_j \frac{\langle\phi_j|\psi_i\rangle}{\langle\phi_j|\phi_j\rangle}. \tag{H.4}$$

By considering $\langle\phi_i|\phi_i\rangle$ and $\langle\phi_i|\phi_{j\neq i}\rangle$, it is easy to show that these functions are orthogonal.

The problem with this method as it stands is that to calculate each ϕ_i for $i > 1$ requires knowledge of all ϕ_j for $j = 1$ to $i - 1$. This is easy enough when working with explicit functions on paper, because each ϕ_j has already been written down in terms of the *original* functions. However, when solving the problem numerically on a computer, this algebra cannot be performed easily. Each ϕ_j is expressed in terms of all previous ϕ_k for $k = 1$ to $j - 1$, each of which is itself expressed in terms of all previous ϕ_l for $l = 1$ to $k - 1$ and so on. Therefore, calculating any reasonably large number of functions involves a huge recursion.

H.4 Modified method

A more compact method of calculating the functions is to make use of the obvious statement that the orthogonal set is a linear combination of the functions in the original set:

$$\phi_i = \alpha_{ij}\psi_j \tag{H.5}$$

for some transformation matrix α_{ij}. For brevity and since it can be calculated from the outset, we define the scalar product matrix of the original functions:

$$S_{ij} = \langle\psi_i|\psi_j\rangle. \tag{H.6}$$

Substituting into the Gram – Schmidt scheme, we find

$$\begin{aligned}
\sum_{j=1}^{i-1} \phi_j \frac{\langle\phi_j|\psi_i\rangle}{\langle\phi_j|\phi_j\rangle} &= \sum_{j,k,l=1}^{i-1} \alpha_{jk}\psi_k \frac{\langle\alpha_{jl}\psi_l|\psi_i\rangle}{\langle\phi_j|\phi_j\rangle} && \text{(H.7)}\\
&= \sum_{k=1}^{i-1}\sum_{j,l=1}^{i-1} \alpha_{jk}\alpha_{jl} \frac{\langle\psi_l|\psi_i\rangle}{\langle\phi_j|\phi_j\rangle}\psi_k && \text{(H.8)}\\
&= \beta_{ik}\psi_k && \text{(H.9)}
\end{aligned}$$

where

$$\beta_{ik} = \begin{cases} \sum_{j,l=1}^{i-1} \alpha_{jk}\alpha_{jl}\frac{S_{li}}{\langle\phi_j|\phi_j\rangle} & : i > k \\ 0 & : i \leq k \end{cases} \tag{H.10}$$

so

$$\phi_i = \alpha_{ij}\psi_j = \psi_i - \beta_{ij}\psi_j = (\delta_{ij} - \beta_{ij})\psi_j \tag{H.11}$$

and

$$\alpha_{ij} = \delta_{ij} - \beta_{ij}. \tag{H.12}$$

In short, the matrix β_{ij} collects all the recursions in terms of the original functions. Note that α_{ij} is lower triangular with ones on the main diagonal.

The norm of the orthogonalised functions is

$$\langle \phi_i | \phi_i \rangle = \sum_{k,l=1}^{i} \alpha_{ik} S_{kl} \alpha_{li}. \tag{H.13}$$

This method works because the original functions are added one at a time. This is why α_{ij} is lower triangular, and means that row i of β depends only on rows 1 to $i-1$ of α. Row i of α and β and also element i of $\langle \phi_i | \phi_i \rangle$ are independent of higher rows and elements, so incomplete matrices can be used to calculate ϕ_i.

The modified method is therefore:

1. calculate S_{ij}
2. set $\phi_1 = \psi_1$ so that $\alpha_{11} = 1$ and $\langle \phi_1 | \phi_1 \rangle = S_{11}$.
3. calculate the next row of β_{ij} and hence α_{ij} from the previous row(s).
4. calculate the value of $\langle \phi_i | \phi_i \rangle$ for use in row $i + 1$ and later

Repeat final two steps for the rest of the original functions.

Note that the only place where the ψ_i come in is in calculating S_{ij}. This method works equally well for orthogonalising a set of vectors $\{\vec{v}_i\}$, with the elements S_{ij} found from

$$S_{ij} = \vec{v}_i . \vec{v}_j . \tag{H.14}$$

A subroutine was written to implement this scheme, calculating the transformation matrix α_{ij} given S_{ij}, from whatever source.

Bibliography

[1] I. C. Skidmore, *An Introduction to Shock Waves in Solids*, Applied Materials Research pp 131 –147 (July 1965).

[2] J. R. Waldram, "The Theory of Thermodynamics", Cambridge (1985).

[3] J A Moriarty, *First-principles equations of state for Al, Cu, Mo and Pb to ultrahigh pressures*, High Pressure Research **13** 6 (1995).

[4] J A Moriarty and M Widom, *First-principles interatomic potentials for transition-metal aluminides: theory and trends across the 3d series*, Phys Rev B (Cond Matt) **56** 13 (1997).

[5] M Widom and J A Moriarty, *First-principles interatomic potentials for transition-metal aluminides. II. Application to Al-Co and Al-Ni phase diagrams*, **58** 14 (1998).

[6] J A Moriarty, *Atomistic simulation of thermodynamic and mechanical properties of materials*, J Computer-Aided Materials Design **5** 2-3 (1998).

[7] P Soderlind and J A Moriarty, *First-principles theory of Ta up to 10 Mbar pressure: Structural and mechanical properties*, Phys Rev B (Cond Matt) **57** 17 (1998).

[8] J A Moriarty, Wei Xu, P Soderlind, J Belak, L H Yang and Jing Zhu, *Atomistic simulations for multiscale modeling in bcc metals*, Trans ASME J of Engineering Materials and Technology **121** 2 (1999).

[9] S J Zhou, D L Preston, P S Lomdahl and D M Beazley, *Large-scale molecular dynamics simulations of dislocation intersection in copper*, Science vol 279 pp 1525 - 1527 (1998).

[10] V Yu Klimenko, *Multiprocess model of detonation (version 4)*, Proc 11th International Symposium on Detonation, held Snowmass CO USA – to appear (1998).

[11] D C Swift, *Shear-banding in Ti-6Al-4V*, Proc Joint AIRAPT/APS Topical Group on Shock Compression of Condensed Matter Conference, AIP Proceedings 309 (1993).

[12] M P Allen and D J Tildesley, "Computer Simulation of Liquids," Oxford (1987).

[13] G I Kerley, *Perturbation theory and the thermodynamic properties of fluids*, J Chem Phys vol 73 no 1 (1980).

[14] M Ross, *High density fluid perturbation theory*, J Chem Phys vol 71 no 4 (1979).

[15] A P Sutton, "Electronic Structure of Materials", Oxford (1993).

[16] P G Parish (Mathematical Physics Associates Limited/AWE Aldermaston), AWE report O-1/95 (1995).

[17] B A Nadykto, *A semiempirical model for calculation of the energies of states of multielectron ions*, Physics - Uspekhi **36** 9 (1993).

[18] B A Nadykto, *Estimation of element and compound electron phase parameters from compressibility data*, presentation to *International Workshop on New Models and Numerical Programs for Computing Shock Waves in Condensed Media*, held Oxford, U.K., 15-19 Sep 1997.

[19] K S Holian (Ed.) "T-4 Handbook of Material Property Data Bases, Vol 1c: Equations of State", Los Alamos National Laboratory report LA-10160-MS (1984).

[20] M C Warren and G J Ackland, *Ab initio structural instabilities in magnesium silicate perovskite*, Physics and Chemistry of Materials **23** 2 (1996).

[21] H-C Hsueh, *Pressure-induced polymorphism in CuCl: an ab initio study*, Phys Rev B **51** 18 (1997).

[22] B B Karki, M C Warren, L Stixrude, G J Ackland and J Crain, *Ab initio studies of high-pressure structural transformation in silica*, Phys Rev B (Cond Matt) **55** 6 (1997).

[23] A A Kelsey, G J Ackland and S J Clark, *Stability and electronic structure of the cinnabar phase in GaAs*, Phys Rev B (Cond Matt) **57** 4 (1998).

[24] U Pinsook and G J Ackland, *Simulation of martensitic microstructural evolution in zirconium*, Phys Rev B **58** 17 (1998).

[25] U Pinsook and G J Ackland, *Calculation of anomalous phonons and the hcp-bcc transition in zirconium*, Phys Rev B **59** 21 (1999).

[26] S G Cochran and J Chan, *Shock initiation and detonation models in one and two dimensions*, Lawrence Livermore National Laboratory report UCID-18024 (1979).

[27] E L Lee and C M Tarver, Phys Fluids vol 23 p 2362 (1980).

[28] P C Souers and L C Haselman Jr, *Detonation equation of state at LLNL, 1993*, Lawrence Livermore National Laboratory report UCRL-ID-116113 (1994).

[29] P Hohenberg and W Kohn, 'Inhomogeneous electron gas,' Phys Rev B vol 136 no 3B, pp 864 - (1964).

[30] W Kohn and L J Sham, *Self-Consistent Equations Including Exchange and Correlation Effects*, Phys Rev **140** 4A (1965).

[31] J Perdew, Phys Rev **B46** 6671 (1992).

[32] J Perdew, Phys Rev **B50** 4954 (1994).

[33] P Soderlind, O Eriksson, J Trygg, B Johansson and J M Wills, *Density-functional calculations for cerium metal*, Phys Rev B (Cond Matt) **51** 7 (1995).

[34] A Delin, L Fast, B Johansson, J M Wills and O Eriksson, *Method for calculating valence stability in lanthanide systems*, Phys Rev Lett **79** 23 (1997).

[35] O Eriksson, J Trygg, H Hjortsam, B Johansson and J M Wills, *Theoretical aspects of the 4f-localization at the surface of* α*-Ce*, Surface Science **382** 1-3 (1997).

[36] A Delin, L Fast, B Johansson, O Eriksson and J M Wills, *Cohesive properties of the lanthanides: Effect of generalised gradient corrections and crystal structure*, Phys Rev B (Cond Matt) **58** 8 (1998).

[37] I Galanakis, M Alouani, J M Wills and H Dreysse, *A real-space full-potential localized LMTO method for non-collinear magnetism*, Phil Mag B **78** 5-6 (1998).

[38] J M Ziman, "Elements of Advanced Quantum Theory," Cambridge (1969).

[39] S Gasiorowicz, "Quantum Physics," Wiley (1974).

[40] N W Ashcroft and N D Mermin, "Solid State Physics," Holt-Saunders (1976).

[41] D A Liberman, 'Self-consistent field model for condensed matter,' Phys Rev B vol 20 no 12 pp 4981 - (1979).

[42] W A Harrison, "Electronic structure and the properties of solids," Freeman (1980).

[43] A J R Aitchison and A J G Hey, "Gauge Theories in Particle Physics," Adam Hilger (1982).

[44] D F Lawden, "An Introduction to Tensor Calculus, Relativity and Cosmology" (3rd Ed), Wiley (1982).

[45] P C W Davies, "Quantum Mechanics," Routledge and Kegan Paul (1984).

[46] F Halzen and A D Martin, "Quarks and leptons," Wiley (1984).

[47] G Arfken, "Mathematical Methods for Physicists" (3rd Ed), Academic (1985).

[48] B I Bennett and D A Liberman, 'INFERNO,' Los Alamos National Laboratory report LA-10309-M (1985).

[49] E R Dobbs, "Electromagnetic Waves," Routledge and Kegan Paul Student Physics Series (1985).

[50] A Sudbery, "Quantum Mechanics and the Particles of Nature," Cambridge (1986).

[51] W H Press, B P Flannery, S A Teukolsky and W T Vetterling, "Numerical Recipes," Cambridge (1989).

[52] M Kaku, "Quantum Field Theory," Oxford (1993).

[53] P Strange, "Relativistic quantum mechanics," Cambridge (1998).

[54] Scottish Examination Board, Science Data Booklet (1982).

[55] M C Payne, M P Teter, D C Allan, T A Arias and J D Joannopoulos, *Iterative minimisation techniques for* ab initio *total-energy calculations: molecular dynamics and conjugate gradients*, Rev Mod Phys **64** 4 (1992).

[56] G Ackland (University of Edinburgh), private communication (1994).

[57] M C Warren (University of Edinburgh), private communication (1994).

[58] S Clark (University of Edinburgh), private communication (1995).

[59] P. Pulay, Mol.Phys. **17** 197 (1969).

[60] G. P. Francis and M. C. Payne, J. Phys. C **2**, 4395 (1990).

[61] H. C. Hsueh, M. C. Warren, H. Vass, G. J. Ackland, S. J. Clark and J. Crain, Phys.Rev. B **53** 14806 (1996).

[62] M-H Lee (University of Cambridge), private communication (1995).

[63] Maradudin, Montroll and Weiss, "Theory of Lattice Dynamics in the Harmonic Approximation", Academic Press (1963).

[64] G J Ackland, M C Warren and S J Clark, *Practical mathods in* ab initio *lattice dynamics*, J. Phys Cond Matt **9** 2 pp 107–118 (1997).

[65] G J Ackland (University of Edinburgh), private communication (1999).

[66] D. A. Liberman and B. I. Bennett, *Atomic vibrations in a self-consistent-field atom-in-jellium model of condensed matter*, Phys Rev B **42** 2475 (1990).

[67] K. Hanson (Los Alamos National Laboratory), private communication (1997).

[68] A D Bruce, N B Wilding and G J Ackland, *Free energy of crystalline solids: a lattice-switch Monte-Carlo method*, Phys Rev Lett **79** 16 (1997).

[69] H M Rosenberg, "The Solid State", Oxford (1988).

[70] D J Steinberg, *Equation of State and Strength Properties of Selected Materials*, Lawrence Livermore National Laboratory report UCRL-MA-106439 change 1 (1996).

[71] R G Greene, H Luo and A L Ruoff, *Al as a Simple Solid: High Pressure Study to 220 GPa (2.2 Mbar)*, Phys Rev Lett **73** 2075 (1994).

[72] S P Marsh (Ed), "LASL Shock Hugoniot Data", University of California (1980).

[73] A C Mitchell and W J Nellis, *Shock compression of aluminium, copper and tantalum*, J Applied Physics **52** 5 (1981).

[74] A Hauer (Los Alamos National Laboratory), private communication (1996).

[75] J Wark (University of Oxford), private communication (1996).

[76] H Kolsky, "Stress Waves in Solids", Dover (1963).

[77] B B Karki, G J Ackland and J Crain, *Elastic instabilities in crystals from ab initio stress-strain relations*, J Phys Cond Matt **9** 41 (1997).

[78] D A Young, "Phase Diagrams of the Elements", University of California (1991).

[79] M. N. Pavlovskii, *Formation of metallic modifications of germanium and silicon under shock loading*, Sov Phys - Solid State **9** 11 (1968).

[80] W J Nellis, J A Moriarty, A C Mitchell and N C Holmes, *Equation of state of beryllium at shock pressures of 0.4 – 1.1 TPa (4 – 11 Mbar)*, J App Phys **82** 5 (1997).

[81] D. G. Pettifor, "Bonding and Structure of Molecules and Solids", Oxford (1995).

[82] R D Richtmeyer and K W Morton, "Finite difference methods for initial value problems," Wiley (1967).

[83] D J Benson, *Computational methods in Lagrangian and Eulerian hydrocodes*, J Computer Methods in Applied Mechanics and Engineering **99** pp 235 - 394 (1992).

[84] E S Oran and J P Boris, "Numerical Simulation of Reactive Flow," Elsevier (1987).

[85] R Courant and D Hilberts, "Methods of mathematical physics," vols 1 and 2, New York (1953,1962).

[86] G A Sod, *A survey of several finite difference methods for systems of nonlinear hyperbolic conservation laws*, J Comput Phys **27** 1 – 31 (1978).

[87] D C Swift, *Ab initio polymorphic equations of state for silicon*, Proc APS Topical Conference on Shock Compression of Condensed Matter, held Snowbird CO USA Jun 1999 - to appear.

[88] J.S. Wark, A. Loveridge, A. Allen, D. Kalantar, B. Remington, R.W. Lee, S. Weber A. Hauer, G. Kyrala, D. Paisley, B. Holian, P. Lomdahl, T. Boehly, M.A. Meyers, D.C. Swift, *Anomalous response of silicon to uniaxial shock compression on nanosecond timescales*, paper in preparation for Phys Rev Lett (1999).

[89] D C Swift, *An evaluation of the 'CERES' X-ray machine for flash crystallography*, Proc International Workshop on New Models and Numerical Codes for Shock Wave Processes in Condensed Media, held Oxford, Sep 97 (to appear) (1997).

[90] D C Swift, G J Ackland, A Hauer, J S Wark, D H Kalantar, G A Kyrala and R Kopp, *Ab initio predictions of material properties compared with time-resolved crystallography of laser-shocked silicon*, 40th APS Plasma Physics Group meeting, held New Orleans, USA, 16 – 20 Nov 1998 (1998).

[91] A Hauer (Los Alamos National Laboratory), private communication (1998).

[92] B B Karki, S J Clark, M C Warren, H-C Hsueh, G J Ackland and J Crain, *Ab initio elasticity and lattice dynamics of $AgGaSe_2$*, J Phys Cond Matt **9** 2 (1997).

[93] R Ahuja, L Fast, O Eriksson, J M Wills and B Johansson, *Elastic and high pressure properties of ZnO*, J App Phys **83** 12 (1998).

[94] D R Allan, S J Clark, M J P Brugmans, G J Ackland and W L Vos, *Structure of crystalline methanol at high pressure*, Phys Rev B (Cond Matt) **58** 18 (1998).

[95] B B Karki and J Crain, *Structure and elasticity of CaO at high pressure*, J Geophys Research **103** B6 (1998).

[96] C R S DaSilva, B B Karki, L Stixrude and R M Wentzcovitch, *Ab initio study of the elastic behaviour of $MgSiO_3$ at high pressure*, Geophys Research Lett **26** 7 (1999).

[97] A Machova and G J Ackland, *Dynamic overshoot in α-iron by atomistic simulations*, Modelling and Simulation in Materials Science and Engineering **6** 5 (1998).

[98] S J Clark and G J Ackland, *Ab initio calculations of the self-interstitial in silicon*, Phys Rev B (Cond Matt) **56** 1 (1997).

[99] G J Ackland, S J Wooding and D J Bacon, *Defect, surface and displacement threshold properties of α-zirconium simulated with a many-body potential*, Phil Mag A **71** 3 (1995).

[100] G J Ackland, D J Bacon, A F Calder and T Harry, *Computer simulation of point defect properties in dilute Fe-Cu alloy using a many-body interatomic potential*, Phil Mag A **75** 3 (1997).

[101] P C Gehlen, J R Beeler (Jr) and R I Jaffee (Eds), "Interatomic Potentials and Simulation of Lattice Defects", Plenum (1972).

[102] M W Finnis and J E Sinclair, *A simple empirical* N*-body potential for transition metals*, Phil Mag A **50** 1 (1984).

[103] D C Swift (AWE Aldermaston), unpublished work (1994).

[104] G J Ackland, G Tichy, V Vitek and M W Finnis, *Simple* N-*body potentials for the noble metals and nickel*, Phil Mag A **56** 6 (1987).

[105] Royal Military College of Science (Shrivenham) notes from COMAG course on "Approximation and Data Fitting" (1989).

[106] J Moriarty and J D Althoff, *First principles temperature-pressure phase diagram of magnesium*, Phys Rev B **51** 9 pp 5609 – 5616 (1995).

[107] D McKie and C McKie, 'Crystalline Solids', Nelson (1974).

[108] D C Wallace, *Thermoelastic-Plastic Flow in Solids*, Los Alamos National Laboratory report LA-10119 (1985).

[109] D J Steinberg, *Equation of State and Strength Properties of Selected Materials*, Lawrence Livermore National Laboratory report UCRL-MA-106439 (1991).

[110] A V Bushman, G I Kanel', A L Ni and V E Fortov, "Intense Dynamic Loading of Condensed Matter", Taylor and Francis (1993).

[111] O. Schulte and W. B. Holzapfel, Phys Rev B **48**, 767 (1993).

[112] *Nuclides and Isotopes – Chart of the Nuclides* (14th Ed.), General Electric Company (1989).

[113] J Moriarty (Lawrence Livermore National Laboratory), private communication (1994).

Interim publications

The following publications were made during the course of this work:

1. D C Swift, '*Ab fere initio* Equations of State for Aluminium', Proc APS Topical Conference on Shock Compression of Condensed Matter, held Seattle WA USA 13 – 18 Aug 1995, AIP (1996).
2. D C Swift, '*Ab fere initio* Equations of State for Solids', Proc APS Topical Conference on Shock Compression of Condensed Matter, held Amherst MA USA 28 Jul – 2 Aug 1997, AIP (1998).

3. D C Swift, '*Ab fere initio* Equations of State for Solids', Proc International Workshop on New Models and Numerical Codes for Shock Wave Processes in Condensed Media, held Oxford, Sep 1997, AWE/Hunting-BRAE (1999).

4. D C Swift, G J Ackland, A Hauer, J S Wark, D H Kalantar, G A Kyrala and R Kopp, *Ab initio predictions of material properties compared with time-resolved crystallography of laser-shocked silicon*, 40th APS Plasma Physics Group meeting, held New Orleans LA USA, 16 – 20 Nov 1998.

5. D C Swift, 'Ab initio polymorphic equations of state for silicon,' Proc APS Topical Conference on Shock Compression of Condensed Matter, held Snowbird CO USA Jun 1999 – to appear.

6. D C Swift, 'Time-dependent polymorphism: prediction and hydrocode models,' Conference on New Models and Numerical Codes for Shock Waves in Solids, held University of Maryland, USA, Jul 1999 – to appear.