INSTITUTE OF PHYSICS PUBLISHING REPORTS ON PROGRESS IN PHYSICS

Rep. Prog. Phys. **69** (2006) 2365–2441 doi:10.1088/0034-4885/69/8/R03

First-principles modelling of Earth and planetary materials at high pressures and temperatures

M J Gillan[1], D Alfè[1,2], J Brodholt[2], L Vočadlo[2] and G D Price[2]

[1] Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, UK
[2] Department of Earth Sciences, University College London, Gower Street, London WC1E 6BT, UK

Received 16 December 2005, in final form 11 April 2006
Published 25 July 2006
Online at stacks.iop.org/RoPP/69/2365

Abstract

Atomic-scale materials modelling based on first-principles quantum mechanics is playing an important role in the science of the Earth and the other planets. We outline the basic theory of this kind of modelling and explain how it can be applied in a variety of different ways to probe the thermodynamics, structure and transport properties of both solids and liquids under extreme conditions. After a summary of the density functional formulation of quantum mechanics and its practical implementation through pseudopotentials, we outline the simplest way of applying first-principles modelling, namely static zero-temperature calculations. We show how calculations of this kind can be compared with static compression experiments to demonstrate the accuracy of first-principles modelling at pressures reached in planetary interiors. Noting that virtually all problems concerning planetary interiors require an understanding of materials at high temperatures as well as high pressures, we then describe how first-principles lattice dynamics gives a powerful way of investigating solids at temperatures not too close to the melting line. We show how such calculations have contributed to important progress, including the recent discovery of the post-perovskite phase of $MgSiO_3$ in the D$''$ layer at the base of the Earth's mantle. A range of applications of first-principles molecular dynamics are then reviewed, including the properties of metallic hydrogen in Jupiter and Saturn, of water, ammonia and methane in Uranus and Neptune, and of oxides and silicates and solid and liquid iron and its alloys in the Earth's deep interior. Recognizing the importance of phase equilibria throughout the planetary sciences, we review recently developed techniques for the first-principles calculation of solid and liquid free energies, melting curves and chemical potentials of alloys. We show how such calculations have contributed to an improved understanding of the temperature distribution and the chemical composition throughout the Earth's interior. The review concludes with a summary of the present state of the field and with some ideas for future developments.

(Some figures in this article are in colour only in the electronic version)

This paper was invited by Professor G Schubert.

Contents

Page

1. Introduction 2368

2. The science of planetary interiors: a brief overview 2371
    - 2.1. The Earth and the terrestrial planets 2371
    - 2.2. The gaseous planets 2374
    - 2.3. Moons 2375

3. First-principles techniques 2376
    - 3.1. Why it is difficult to apply quantum theory to many-electron systems 2376
    - 3.2. Ignoring correlations: Hartree–Fock theory 2377
    - 3.3. Density functional theory 2377
    - 3.4. Approximations to the exchange-correlation energy 2380
    - 3.5. Pseudopotentials 2381
    - 3.6. Basis sets and periodic boundary conditions 2382
    - 3.7. Atomic forces and stress tensor 2383
    - 3.8. Practicalities 2384

4. Energetics of static crystals: athermal calculations 2384
    - 4.1. Crystal structure and equation of state 2385
        - 4.1.1. Magnesium oxide 2385
        - 4.1.2. Magnesium silicate 2386
        - 4.1.3. Silica 2388
        - 4.1.4. Iron 2389
        - 4.1.5. Iron silicide and carbide. 2390
        - 4.1.6. Water and ammonium dihydrate 2391
    - 4.2. Elastic properties 2392
        - 4.2.1. $MgSiO_3$ perovskite 2392
        - 4.2.2. $MgSiO_3$ post-perovskite 2393
        - 4.2.3. Iron 2393
    - 4.3. Crystal defects 2393

5. High-temperature properties from lattice dynamics 2394
    - 5.1. Thermodynamics from quasi-harmonic theory 2395
    - 5.2. Basic theory of lattice vibrations 2396
    - 5.3. First-principles calculation of phonon frequencies 2397
    - 5.4. Phonon dispersion relations at ambient pressure 2398
    - 5.5. Phonons at high pressure 2398
    - 5.6. Thermodynamics in the quasi-harmonic approximation 2400

6. Molecular dynamics 2402
    - 6.1. Summary of first-principles molecular dynamics 2403
    - 6.2. Hydrogen in Jupiter and Saturn 2404
    - 6.3. Uranus and Neptune 2405
    - 6.4. Iron in the Earth's inner core 2407
    - 6.5. Liquid iron at Earth's core conditions 2408
    - 6.6. Materials of the Earth's mantle 2410
    - 6.7. The first-principles determination of crystal structures 2411

7. Free energy 2412
  7.1. Basic concepts 2412
  7.2. Statistical mechanics: thermodynamic integration 2413
  7.3. The practical calculation of first-principles free energies 2415
8. Melting 2417
  8.1. The three approaches to first-principles melting 2418
  8.2. Free-energy approach 2419
  8.3. The model-fitting approach 2422
  8.4. First-principles coexistence approach 2424
9. Chemical potentials 2426
  9.1. First-principles methods 2426
  9.2. Chemical composition and temperature of the core 2428
  9.3. Implications for the core geotherm 2430
10. Summary and prospects 2430
  Acknowledgments 2432
  References 2432

1. Introduction

The evolution of the Earth and the other planets, as well as their present structure and dynamics, depends on processes that take place in their deep interiors. To understand these processes, we need to know at least the chemical composition, temperature and distribution of heat sources throughout their interiors. Yet these interiors are inaccessible to direct observation. For the Earth, fairly precise information about the density and elastic properties as a function of depth has been deduced from seismic data, and there is detailed knowledge of its magnetic field, which also gives information about the properties of the interior. Limited seismic data are also available for the Moon. For the other planets, measurements of their gravitational moments and their magnetic fields provide valuable constraints on models for their interiors. But this information falls far short of what we need to know. Our present understanding of planetary interiors has been patched together from many disciplines, including astronomy, cosmochemistry, nuclear physics, meteoritics, materials science and crystallography, as well as seismology [29, 79, 220, 264]. Laboratory experiments at high pressures and temperatures provide invaluable information about the physical and chemical properties of the materials that are believed to make up the planets. In the past 15 or years or so, a major new tool has begun to play an important role in probing the planetary interiors: materials modelling based on first-principles quantum mechanics. The aim of this review is to explain how this new tool works and to show how it has already yielded insights that would have been hard to obtain in any other way.

At first sight, it might seem surprising that quantum mechanics has anything to offer to the earth and planetary sciences. After all, quantum mechanics is concerned with atomic and sub-atomic processes on the very short length and time scales on which Newtonian physics breaks down, whereas the physics and chemistry of the Earth and the planets involve long length and time scales. The reason why quantum mechanics is relevant is that it gives a way of calculating the properties of materials under extreme conditions, where laboratory experiments may be difficult or even impossible. In order to provide a scientific framework for this review, we shall give an overview of our current understanding of planetary interiors in section 2. It will become clear there that the key to solving some of the outstanding problems is an accurate knowledge of the thermodynamic, structural and transport properties of planetary materials at pressures in the megabar region (1 Mbar = 100 GPa) and temperatures of several thousand Kelvin. This is where first-principles modelling is now playing an indispensable role.

Throughout this review, we shall try to show the close interplay between first-principles modelling and laboratory experiments. Techniques for performing experiments at pressures and temperatures relevant to planetary interiors are described in a number of reviews, for example Mao and Hemley [168]. There are three main methods for achieving these thermodynamic conditions. Multi-anvil high-pressure cells, initially developed in Japan in the 1970s, enable pressures of 25–30 GPa and temperatures of up to 2000 K to be reached, and these are ideal for crystal synthesis and the study of phase equilibria. To obtain higher pressures and temperatures, diamond anvil cells (DACs) are required. The transparency of diamond means that DACs can be coupled with laser heating of the sample to attain pressures of ca 200 GPa and temperatures up to ~ 6000 K. This enables the conditions within the Earth's outer core to be reproduced. Diamonds are also translucent to x-rays, so that laser-heated DACs can be used with synchrotron x-ray radiation to determine the equations of state of planetary materials. The challenge for such studies is the precision with which the pressure and temperature can be measured. Temperature measurement requires fitting the grey-body spectrum of the laser-heated sample, while pressure measurement requires the extrapolation of the ruby fluorescence pressure standard or the use of a calibrated internal standard. A third approach for generating high pressures consists of shock experiments, in which a high-velocity projectile is fired into a target, and the Rankine–Hugoniot relations are used to determine the shock pressure. Shock experiments sample the Hugoniot trajectory through pressure–temperature space, and the temperature of the sample can be estimated from the Mie–Grüneisen equation. This, however, requires a knowledge of the high p–T heat capacity and Grüneisen parameter of the sample, both of which may be poorly known. Although major advances have been made with all of these techniques, the problem of determining p and T at extreme conditions has led in some cases to considerable discrepancies between different experimental approaches. For example, the estimated melting temperature of iron at 300 GPa as inferred by shock experiments and by laser-heated DAC studies differs by some 2000 K [49, 57, 193]. Difficulties such as these provide a strong incentive for developing accurate and reliable modelling methods for the prediction of materials properties at high pressures and temperatures.

The ambition of deriving the macroscopic properties of materials from an atomic-level description of their energetics began to take shape well over 100 years ago, early milestones being the van der Waals theory of the phase diagram of fluids and Einstein's theory of the specific heat of crystals, later improved by Debye. The key idea of van der Waals was that the energetics of matter on the atomic scale can be represented in terms of repulsions and attractions between the molecules. This basic idea was elaborated and refined by many scientists in the early decades of the twentieth century, including Mie, Born, Mayer and Lennard-Jones. Well before the advent of computers, it had been shown that simple models for the atomic interactions, designed using the basic insights given by quantum mechanics, were able to account for quantities such as equilibrium lattice parameters, elastic and dielectric properties, specific heats, lattice expansion and point defect energies.

The rapid growth in the availability of computers in the decades following the 1950s gave an immense impetus to the quantitative atomic-scale modelling of materials. With the development of the molecular dynamics [3, 106] and Monte Carlo [2, 179] simulation techniques, it became possible to calculate from a given interaction model a very wide range of physical properties of solids, liquids and vapours, such as thermodynamic functions, structural and transport properties and dynamical reponse functions. Those years also saw rapid improvements in the sophistication and realism of the atomic interaction models, with successful models being developed for large classes of materials, including ionic and semi-ionic compounds [60, 287], metals [119], molecular and polymeric materials [66], water and aqueous solutions [223, 135]. The adjustable parameters in these models were determined by the fitting of the predicted properties to selected experimental data, sometimes with a judicious admixture of information from quantum chemistry calculations. It was realized that some of these advances could be directly applied to the study of minerals, since many of the oxides and silicates of interest for the terrestrial planets are exactly the semi-ionic materials for which the atomic-scale models were known to be successful. Early successes in the modelling of minerals include the prediction of the elastic constants of $MgSiO_3$-perovskite [68], the calculation of the heat capacity and entropy of the major mantle-forming phases [221], the prediction of high-pressure phase transformations in simple oxides such as Al_2O_3 [171] and the analysis of the diffusion mechanisms in a variety of oxides and silicates [134, 275].

The extensive work on the simulation of materials using empirical interactions, while giving valuable new insights into a wide range of problems, also revealed the strong limitations of this entire approach. There are several reasons why the empirical approach is limited. The models are necessarily based on assumptions about the nature of the interactions, which, though valid under some circumstances, may fail completely under others. One example of this is the modelling of polymorphic phase transformations in which Si changes from four-fold to six-fold coordination. In general, empirical models have grave difficulties with any problem in which chemical bonds are made or broken. The dissociation of molecular materials such as hydrogen or methane under the high pressures found in the gaseous planets is an example where one would not expect empirical models to be useful. Even when the form of the empirical model is valid, the determination of the adjustable parameters may pose problems, which get worse as the sophistication of the model and the number of its parameters increase. Sometimes, the only option available is intelligent guesswork.

A major paradigm shift in the atomic-scale modelling of materials occurred with the development of first-principles methods [58, 123, 152, 170]. Instead of treating a material as an assembly of atoms, this approach regards it as consisting of nuclei and electrons. The description is based from the start on the fundamental and essentially exact equations of quantum mechanics. This gives the enormous advantage of liberating the modelling of materials from adjustable parameters. In its purest form, the aim of the first-principles approach is to provide a universal method, which can be applied without change to broad classes of materials under wide ranges of conditions. The basic problem to be overcome in this approach is that the equations of quantum mechanics, though essentially exact themselves, cannot be solved exactly for complex systems containing many nuclei and electrons. An outstandingly successful (partial) solution to the problem is given by density functional theory (DFT) [123, 152, 170], a reformulation of the quantum mechanics of many-electron systems, within which it has proved possible to find universal approximations which deliver very good accuracy for the physical properties of wide ranges of materials, including the majority of those of interest in the earth and planetary sciences. Remarkably, one and the same approximation within DFT gives a similarly reliable account of problems that look completely unrelated, for example the phase diagram of $MgSiO_3$ at the conditions of the Earth's lower mantle and the vibrational density of states of iron under Earth's core conditions, as we shall see later.

In the early days of first-principles modelling, computer power was enough to allow the calculation only of rather simple, though still important, zero-temperature properties, such as the relative enthalpies of different crystal structures as a function of pressure. However, the last 10 years have seen the development of robust first-principles techniques for the calculation of high-temperature properties of solid and liquid materials, using the harmonic approximation in the case of crystals, or the much more general, but also more expensive, technique of first-principles molecular dynamics simulation. Our aim in this review is to describe in some detail how all these first-principles methods work, how their reliability and accuracy have been demonstrated by comparison with high-pressure and sometimes high-temperature experiments and how the calculations are placing entirely new constraints on the physical and chemical conditions that exist inside the Earth and the other planets.

Since we want to convey an understanding of the methodology, we place a strong emphasis on the techniques themselves, starting with the fundamentals of density functional theory, and then explaining how it is being used in the earth and planetary sciences, first for the simplest kinds of calculation, and then for more complex kinds. However, to provide background to the scientific work, we give in section 2 a brief summary of our present understanding of planetary interiors, indicating some of the areas where first-principles modelling can contribute. Section 3 then outlines the foundations on which first-principles modelling rests. We particularly stress the nature of the basic approximation that is made in all first-principles calculations. Those who wish to pass over these theoretical matters by omitting section 3 should still find that the rest of the review makes sense. In section 4, we then describe the use of DFT for the simplest kinds of calculations: the energetics of crystals at zero temperature. One of the important purposes of this section is to show by comparisons with experiment the level of accuracy that can be expected of DFT for the simplest properties. We then progress in section 5 to the calculation of vibrational frequencies of crystals and hence the calculation

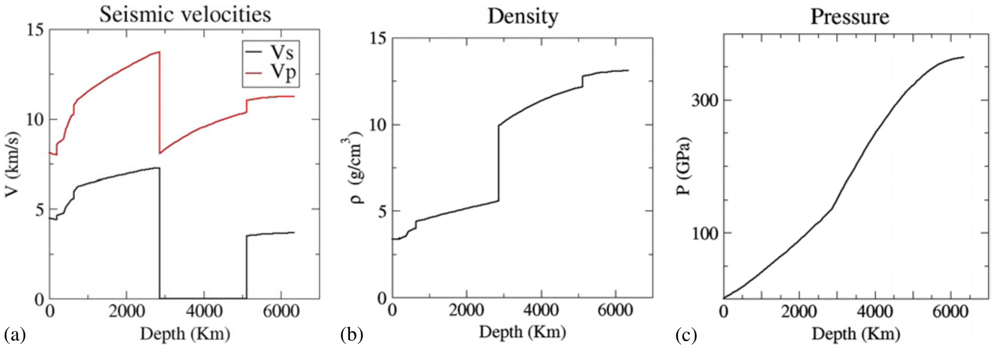

Figure 1. Structure of the Earth from seismic data: (*a*) velocities v_P and v_S of longitudinal (pressure) and transverse (shear) seismic waves, (*b*) density and (*c*) pressure as function of depth.

of thermodynamic and other properties of crystals in the quasiharmonic approximation. This gives further opportunities for comparisons with experiment as well as for predictions that go beyond experiment. The technique of first-principles molecular dynamics simulation is described in section 6, where we also illustrate how this technique can be used to study solids and liquids at the extreme conditions relevant to planetary interiors. Because phase equilibria are so important in this field, the first-principles calculation of free energies of solids and liquids is described separately in section 7, and the calculation of high-pressure melting properties, both by free-energy methods and by other approaches, is outlined in section 8. We return to phase equilibria in section 9, describing how it has recently become possible to calculate chemical potentials in solid and liquid mixtures using first-principles simulation. A summary of the present state of the field and prospects for future progress conclude the review.

2. The science of planetary interiors: a brief overview

2.1. The Earth and the terrestrial planets

The overall structure of the Earth as revealed by seismic measurements is summarized in figure 1, showing the velocities of longitudinal (pressure) and transverse (shear) seismic waves, the density and the pressure as a function of depth from the surface. The most striking feature is a large discontinuity at a depth of 2880 km, which is about half-way to the centre. The region above the discontinuity is the mantle—here we will not be concerned with the thin layer of rock at the Earth's surface known as the crust. The mantle itself consists of the upper mantle, extending down to a depth of 400 km, and the lower mantle, which is the region below a depth of 650 km; the two are separated by a 250 km thick 'transition zone'. The variation of properties through the lower mantle is relatively smooth, but detailed analysis reveals an anomalous region of thickness ~ 200 km at the base of the mantle, known as the D$''$ zone. This zone is seismically anisotropic—seismic velocities vary significantly with the direction of propagation [211]. The region below the mantle is the Earth's core. The data show that the density in the core is much greater than that of the mantle. They also show that the core consists of an outer part, in which transverse waves do not propagate, and an inner part having a radius of 1220 km. The absence of transverse waves means that the outer part is liquid, but the inner core can support a low frequency shear stress and is solid. At the boundary between the inner and outer core, there is a clear discontinuity in the density; the solid core at this boundary is

Table 1. Abundances of the major elements (number of atoms per atom of Si) in the Universe, the Sun and the whole Earth.

	Universe[a]	Sun	Whole Earth
H	2.72×10^4	2.3×10^4	4×10^{-2} [b]
He	2.18×10^3	1.95×10^3	—
C	12.1	6.49	1×10^{-2} ? [b]
N	2.48	2.78	3×10^{-4} ? [b]
O	20.1	22.9	3.73[a]
Mg	1.08	1.09	1.06[a]
Si	1.00	1.00	1.00
S	0.52	0.46	3×10^{-2} ? [b]
Fe	0.90	1.27	0.90[a]
Ni	0.05	0.05	0.05 ? [b]

Note: References: [a] [29]; [b] [176].

6.5 ± 1.4% more dense than the liquid [174]. We shall see later that this density discontinuity plays a key role in constraining the chemical composition of the core. Not shown in figure 1 is the fact that the inner core is seismically anisotropic: compressional waves travel some 3% faster along the polar axis than in the equatorial plane [242, 260].

The chemistry of the Earth's crust can be sampled directly, and information about the upper mantle is available from the study of volcanic products and inclusions that are brought to the surface by eruptions. However, our present understanding of the composition of the deep interior is based mainly on meteoritics and cosmochemistry. A widely accepted first approximation to the overall chemical composition of the Earth is the so-called 'chondritic' model, which assumes that it is essentially the same as that of chondritic meteorites [25, 121]. This indicates that the composition is close to that of the universe as a whole, but with strong depletion of volatile elements, such as hydrogen and helium (table 1). It follows from this that the mineralogy of the mantle must be dominated by $MgSiO_3$. High-pressure crystallography, supported by first-principles calculations, shows that $MgSiO_3$ in the lower mantle has the perovskite crystal structure. However, more detailed analysis shows that the average atomic number in the mantle (21.3) is significantly above that of $MgSiO_3$ (20.1), and the difference is accepted to be due to ∼10% substitution of Mg by Fe. In the lower mantle, there is a also a minor mineral content of magnesiowüstite (MgO with ∼ 30% of Mg substituted by Fe; this is also known as 'ferropericlase') and Ca-perovskite. The nature of the D'' zone at the base of the mantle was mysterious until very recently. The discovery that it is almost certainly a new crystal structure of $MgSiO_3$, in which first-principles calculations played an important role, will be recounted later. Our understanding of the mineralogy of the mantle is summarized in figure 2.

The Earth's core has the density of a heavy metal. The metallic nature of the core is also beyond doubt because the only credible mechanism for the generation of the Earth's magnetic field is based on convection-generated electric currents in the liquid outer core. Cosmochemistry (table 1) and meteoritics indicate that the core must consist mainly of Fe, though a Ni content of somewhat less than ∼10% is also implied. However, the density in the core is ∼7% too low for a Fe/Ni alloy, and light impurities must be present. The leading candidates are S, Si and O, though C and H have occasionally been suggested [25].

The picture we have sketched provides a good starting point for understanding the Earth, but much more is needed. A crucial piece missing from the picture is the temperature distribution throughout the Earth (the so-called geotherm), without which it is impossible to understand convection and heat transport in the core and the mantle. In addition, to be able

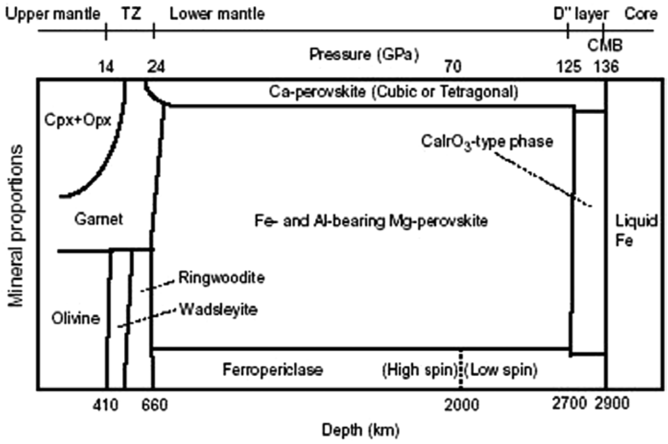

Figure 2. Mineralogy of Earth's mantle: proportions of different minerals as function of depth. Reproduced with permission from [210].

to constrain geochemical and dynamical evolution models of the Earth, we need to know more about the chemical composition of both the core and the mantle. It is clear that further progress depends on a more accurate knowledge of the materials making up the Earth.

A long-established strategy for estimating the geotherm is based on the fact that both the lower mantle and the outer core are in a state of turbulent convection, which implies that the dependence of temperature T on pressure p follows an adiabat. This leads to the following equivalent expressions for the dependence of temperature T on radius r:

$$\mathrm{d}T/\mathrm{d}r = -(\partial T/\partial p)_S \mathrm{d}p/\mathrm{d}r = -\rho g(\partial T/\partial p)_S = -\rho g T\gamma/K_S, \qquad (1)$$

where p is pressure; the second equation follows from the first, since the variation of pressure with radius is $\mathrm{d}p/\mathrm{d}r = -\rho g$, with ρ the mass density and g the acceleration due to gravity at radius r; the third equation expresses the adiabatic variation of T with p in terms of the Grüneisen parameter γ and the adiabatic bulk modulus K_S. Since K_S and g can be obtained from seismic measurements, the only unknown is γ, which can be calculated at any thermodynamic state by first-principles methods, as we shall see. In the mantle, there are widely accepted estimates for the temperature at the base of the transition zone [130], from which equation (1) can be integrated to obtain $T(r)$ throughout the lower mantle. The same procedure can be followed to estimate the geotherm in the outer core, provided we can obtain the temperature at the boundary between the inner and outer core. We shall see that first-principles modelling provides a strategy for doing this. It should be noted that the temperatures thus obtained for the mantle-side and core-side of the core-mantle boundary are not expected to be the same. The reason for this is that the lowest 150 km of the mantle forms a thermal boundary layer, in which conduction, rather than convection, is the mechanism of heat transfer. The temperature difference across this boundary can be estimated from a knowledge of the thermal conductivity in this region. The entire geotherm strategy thus depends on an accurate knowledge of the properties of materials at high temperatures and pressures.

The outline we have given of the structure and chemical composition of the Earth also applies with some changes to the terrestrial planets Mercury, Venus and Mars. However, we are not aware of first-principles work on problems specific to these planets, so they will not figure in this review.

2.2. *The gaseous planets*

The giant planets Jupiter and Saturn consist mainly of H and He, though both planets presumably have small rocky cores, about which nothing is known. Because they are almost entirely gaseous, they do not have sharply defined surfaces, and they are not believed to have internal surfaces either, except for the presumed core boundary. The measured mass fractions of helium in their atmospheres are significantly lower than the initial solar abundance of 0.272 ± 0.005 [248]. For Jupiter, this mass fraction is 0.236 ± 0.004, and for Saturn it may be even lower. This may be due to unmixing in the H–He mixture, because of He precipitation. One reason for thinking this is that Saturn in particular has an unexpectedly high luminosity, which would be explained by gravitational energy release due to the sinking of He. Both planets have large magnetic fields, which can only be explained by electric currents generated in their interiors by a convection-driven dynamo mechanism. The spatial distribution of these fields revealed by the Pioneer 11 and Voyager 1 space probes indicates that the planets are electrically conducting over most of their interiors. Models of their interiors based on estimated equations of state of hydrogen–helium mixtures and the asssumption of adiabaticity predict that the temperature and pressure in the central regions of Jupiter and Saturn are in the region of 20 000 K, 4 TPa and 10 000 K, 1 TPa, respectively [62, 248].

Clearly, an understanding of Jupiter and Saturn depends heavily on an understanding of the properties of H and H/He mixtures at pressures and temperatures up to those in the centres of the planets. Since no laboratory experiments are currently able to reproduce the conditions in their deep interiors, theory must play a key role. The prediction that H becomes a metal under high compression was first made in a celebrated paper by Wigner and Huntington [284], in which they estimated that the pressure of the transition to the metallic state would be at least 25 GPa. Subsequently, it has become clear that the true transition pressure is far above this. Static compression experiments on low-temperature solid hydrogen have been performed at pressures up to over 300 GPa [165, 166, 188]. The evidence from these experiments is not entirely consistent, with some measurements indicating that hydrogen is still transparent at 342 GPa [188], while others appear to show that it becomes opaque at 320 GPa [166]. The experimental evidence has been interpreted to suggest that solid hydrogen may become fully metallic at around 450 GPa [166]. Shock experiments [73, 125, 147, 189, 280] explore an entirely different part of the phase diagram, since temperature increases rapidly along the Hugoniot. Experiments performed with the 'reverberating shock' technique [280], which allows the exploration of pressures and temperatures not too far from those in the outer regions of Jupiter and Saturn, indicate that hydrogen becomes a poor metal at 140 GPa and 3000 K, corresponding to a compression factor of about 9. The static compression and shock results can only be reconciled by postulating that the metallization is strongly promoted by increased temperature.

We note some of the key issues [35, 248] which first-principles methods may be able to help resolve. First, we want to know in more detail the insulator–metal transition line. A second issue that may or may not be related to this is the transition from molecular to atomic hydrogen and, in particular, the question whether this transition is intimately related to the insulator–metal transition. A third issue concerns the question of whether either or both of these transitions may be associated with a first-order phase transition in the fluid phase. The important consequences of such a phase transition for both the nature of convection and the temperature distribution inside the giant planets have been emphasized by Stevenson [248]. In addition to these questions, an understanding of the temperature distribution also depends on an improved knowledge of the entropy and Grüneisen parameter as functions of p and T, while an understanding of the magnetic fields requires knowledge of the electrical conductivity. Finally, the important issue of whether and how much helium precipitation contributes to energy release demands a much improved understanding of the thermodynamics of H/He mixtures under planetary conditions.

Uranus and Neptune are believed to have similar structures, consisting of a small rock core, an outer gaseous part and, between the two, a thick ‘ice layer’, consisting of water, methane and ammonia in solar proportions, the approximate mole fractions being 56% H_2O, 36% CH_4 and 8% NH_3 [61, 128]. (The commonly used phrase ‘ice layer’ is perhaps misleading, since the layer is believed to be fluid.) The pressure and temperature in the ice layer are believed to follow adiabats, spanning the range 20 GPa and 2000 K to 600 GPa and 7000 K in both planets. In trying to understand the physical and chemical states of the material along this adiabat, an important fact is the existence of the very anisotropic magnetic fields of the two planets measured by the Voyager 2 spacecraft. The fields presumably must originate from electric currents within the ice layer, which could be either ionic or electronic in origin or perhaps both. In fact, there have been speculations for many years about high-pressure forms of ice and about possible superionic conduction in high-pressure solid forms of ice (see, e.g. [83]).

There is now convincing experimental evidence of the pressure-induced transition in water from a solid consisting of well-defined molecules to one consisting of hydrogen and oxygen ions [34, 110, 255]. This transition is accompanied by so-called ‘hydrogen-bond symmetrization’, in which the protons go from asymmetrical positions associated with particular oxygens to symmetrical positions mid-way between two oxygens. Measurements of the high-pressure melting curve of water [109, 235] agree in showing a strong break in the melting slope in the region of $p \sim 45\,\text{GPa}$, indicating that the solid in coexistence with the liquid at higher pressures has a high entropy, perhaps connected with superionic conduction. There have been a number of dynamic compression experiments on water, using both single-shock and reverberating shock techniques [65, 124]. Measurements of electrical conductivity in these experiments show that the conductivity increases exponentially up to pressures of $\sim 30\,\text{GPa}$, but above that has rather constant values of $10\,\Omega^{-1}\,\text{cm}^{-1}$, typical of an ionically conducting liquid. There are also shock data [227] and static-compression data [44] on methane, which have been interpreted as showing dissociation to carbon and hydrogen at pressures and temperatures in the region 10–50 GPa and 2000–3000 K, respectively.

Issues that might be clarified with the aid of first-principles simulation include the following. First, one needs to know the $p(T)$ regions in which H_2O, NH_3 and CH_4 go from the molecular to the dissociated atomic state. Second, the ionic conductivity (if any) of the dissociated materials is needed as a function of state. Third, for the purpose of understanding convection, it is clearly vital to know whether the material is a fluid throughout the ice layer or whether there is a possibility that it might be partially solid. Finally, one hopes that first-principles simulation will be able to show whether the material might become an electronic conductor in the lower regions of the ice layer, and if so what values of electrical conductivity might be attained.

2.3. Moons

There are more than 120 moons in the solar system, and the vast majority contain a significant amount of ice and other volatile solids, such as methane, ammonia and ammonia dihydrate; typically, the objects are roughly half ‘ice’ and half ‘rock’. Yet for many of the planet-forming ices, basic properties, such as thermal expansion and equation of state, are either poorly known or not known at all. There has been a considerable amount of first-principles work on water ice, some of it relevant to understanding the planetary moons. Recent work on this and on ammonia dihydrate will be reviewed in section 4.1.6.

3. First-principles techniques

Since the significance of the advances described in this review rests entirely on the validity of the first-principles modelling techniques, we present here a summary of these techniques. We shall not attempt to go into depth, because this would not be possible within the scope of a brief review such as this. The reader who wants to achieve a detailed understanding of density functional theory (DFT) will turn to papers and books [149, 170, 213]. The recent book by Martin [170] is particularly thorough. A simplified (almost) non-mathematical explanation of DFT has been given elsewhere by one of the present authors [107]. The summary presented here is no more than a quick guided tour through the leading ideas. We have tried to make it accessible both to those who plan to study the methods in greater depth and to those whose interests may be mainly experimental. After a reminder of the quantum mechanics of many-electron systems, we state the two fundamental theorems of DFT, outline the schemes that are used to apply the theory, with a strong emphasis on pseudopotential methods, and give some practical information about computer codes and the computational resources needed.

3.1. Why it is difficult to apply quantum theory to many-electron systems

The central aim of DFT is to calculate the ground-state total energy $U(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N)$ for a system of interacting electrons in the electrostatic field of N nuclei at the positions $\boldsymbol{r}_i$ $(i = 1, 2, \ldots, N)$. From this total energy function, one can calculate a vast range of other quantities, including all thermodynamic quantities. If the equations of quantum mechanics could be implemented without any approximations, the function $U(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N)$ could be calculated exactly, and it would be possible to compute the properties of materials more accurately than they could be measured in any experiment. Unfortunately, there is a huge gulf between ‘in principle’ and ‘in practice’. The only systems of nuclei and electrons that can be treated exactly in practice are those containing a single electron, examples of which include the hydrogen atom and the H_2^+ molecular ion. (The helium ion He^+ and other single-electron ions such as Li^{2+} are also examples.) To find the allowed energies E_n of the hydrogen atom and the corresponding wavefunctions $\psi_n(\boldsymbol{r})$ of the electron, one has to solve the Schrödinger equation, according to which

$$-\frac{\hbar^2}{2m}\nabla^2\psi_n(\boldsymbol{r}) + V(\boldsymbol{r})\psi_n(\boldsymbol{r}) = E_n\psi_n(\boldsymbol{r}). \tag{2}$$

Here, $\hbar$ is Planck’s constant h divided by 2π and $V(\boldsymbol{r})$ represents the potential of the electron in the electrostatic field of the proton:

$$V(\boldsymbol{r}) = -e^2/4\pi\epsilon_0 r, \tag{3}$$

with r the distance between the electron and the proton (we use SI units). This problem can be solved with pencil and paper, and there are well-known formulae for the ground-state energy and the other energy eigenvalues E_n and associated eigenfunctions $\psi_n(\boldsymbol{r})$.

As soon as one considers atoms with more than a single electron, it becomes impossible to solve the equations of quantum mechanics exactly, and the problem is even worse for systems containing many atoms. The helium atom is already non-trivial. Here, we have two electrons, acted on by the electrostatic attraction of the nucleus and their own electrostatic repulsion. It is this repulsion that causes all the trouble. If there were no repulsion, each electron would behave as if the other were not there, and we would be back to a single-electron problem such as equation (2), which is easy to solve. Without repulsion, the two electrons have their own wavefunctions, and the wavefunction $\psi(\boldsymbol{r}_1, \boldsymbol{r}_2)$ of the two-electron system can be expressed in terms of the single-electron wavefunctions. In doing this, we have to respect the fact that electrons are fermions, which means that the total wavefunction must change its sign when the two electrons are swapped. To express this requirement of 'exchange symmetry', we have to bring in the spin of the electron. For example, if both electrons have spin up then $\psi(\boldsymbol{r}_1, \boldsymbol{r}_2)$ for the non-interacting electrons is given by

$$\psi(\boldsymbol{r}_1, \boldsymbol{r}_2) = \psi_a(\boldsymbol{r}_1)\psi_b(\boldsymbol{r}_2) - \psi_a(\boldsymbol{r}_2)\psi_b(\boldsymbol{r}_1), \tag{4}$$

where ψ_a and ψ_b are the two spatial wavefunctions. This exchange symmetry is the origin of the Pauli exclusion principle, which says that two electrons cannot be in the same quantum state. It also implies that two electrons with the same spin cannot be in the same place: if $\boldsymbol{r}_1 = \boldsymbol{r}_2$ in equation (4), then $\psi(\boldsymbol{r}_1, \boldsymbol{r}_2)$ vanishes.

The repulsion between electrons completely spoils this simplicity. The true wavefunction, instead of being expressible in the product form of equation (4), is some very complicated function of the electron positions, which we can never hope to find exactly. This reflects the fact that the electrons do not move independently: their motion is correlated. This is the fundamental problem in doing quantum calculations on systems containing more than a single electron.

3.2. *Ignoring correlations: Hartree–Fock theory*

We cannot simply ignore the repulsion between electrons. When the distance between two electrons is 1 Å, their electrostatic interaction energy is roughly 14 eV, and if we ignored an energy as large as this we would get completely wrong results. However, instead of ignoring their interaction, as a first step we can ignore their correlation. Ignoring correlation means assuming (incorrectly) that the overall wavefunction can be represented in an antisymmetrized product form, such as that for a two-electron system is given in equation (4). We do the best job we can using this product form. This idea leads to Hartree–Fock theory. One way of thinking of this is in terms of the variational principle of quantum mechanics. This states that if the total energy of the system is calculated using an approximate wavefunction, then the result is guaranteed to be above the true ground-state energy. The equations of Hartree–Fock theory yield the nearest approximation to the true ground-state energy that can be obtained within the approximation that the wavefunction has the product form. These equations are very like the Schrödinger equation for independently moving single electrons, but the potential $V(\boldsymbol{r})$ that appears in the normal single-electron Schödinger equation (2) now depends on the quantum states occupied by the electrons.

Hartree–Fock theory works quite well for many-electron atoms and can also be used quite successfully for some kinds of systems containing many atoms. But for most materials the complete neglect of electron correlations is too drastic an approximation. For quantitative quantum calculations on systems containing many atoms, more accurate approximations are needed, and DFT provides the theoretical framework within which such approximations can be formulated.

3.3. *Density functional theory*

DFT theory addresses the general problem of finding the ground-state energy E_g of a system of interacting electrons acted on by a potential $V(\boldsymbol{r})$. The interaction between the electrons is, of course, the usual electrostatic potential $e^2/4\pi\epsilon_0 r$, with r the distance between electrons. Initially, $V(\boldsymbol{r})$ is allowed to be completely arbitrary, though naturally when DFT is applied in practice $V(\boldsymbol{r})$ represents the field due to the atomic nuclei. The distinctive feature of the DFT formulation of quantum mechanics is that it focuses on the density distribution $n(\boldsymbol{r})$ of electrons in the ground state, which depends on $V(\boldsymbol{r})$. The meaning of $n(\boldsymbol{r})$ is that at any position $\boldsymbol{r}$ in space, the probability of finding an electron in an infinitesimal volume $\mathrm{d}v$ at that point is $n(\boldsymbol{r})\,\mathrm{d}v$.

The whole of DFT rests on two exact fundamental theorems [123, 152]. The first theorem concerns the relationship between the ground-state density $n(\boldsymbol{r})$ and the potential $V(\boldsymbol{r})$. Assuming that the ground-state is non-degenerate, then clearly $n(\boldsymbol{r})$ is uniquely determined by $V(\boldsymbol{r})$. It is straightforward to show that the converse is also true, so that we have the following theorem.

- Theorem 1. It is impossible that two potentials $V_1(\boldsymbol{r})$ and $V_2(\boldsymbol{r})$ that differ by more than a constant give rise to the same ground-state density $n(\boldsymbol{r})$.

This implies that the ground-state energy E_g, which is obviously completely specified by $V(\boldsymbol{r})$, can also be thought of as completely speficied by $n(\boldsymbol{r})$, since there is only one possible ground state in which the density is $n(\boldsymbol{r})$. We can say that E_g is a *functional* of $n(\boldsymbol{r})$, a fact that is usually expressed by writing it as $E_g[n(\boldsymbol{r})]$.

It is convenient to separate the ground-state energy into two parts: the first is the interaction energy of the electrons with the potential $V(\boldsymbol{r})$, which is the integral of $V(\boldsymbol{r})n(\boldsymbol{r})$ over the volume of the system; the second, called $F[n(\boldsymbol{r})]$, is the sum of the kinetic energy of the electrons and their mutual interaction energy:

$$E_g[n(\boldsymbol{r})] = \int V(\boldsymbol{r})n(\boldsymbol{r})\,\mathrm{d}v + F[n(\boldsymbol{r})]. \tag{5}$$

Since E_g is uniquely determined by $n(\boldsymbol{r})$, then so is $F[n(\boldsymbol{r})]$. The second theorem of DFT then states the following.

- Theorem 2. The true ground-state energy of the system acted on by the given potential $V(\boldsymbol{r})$ is obtained by minimizing the right-hand side of equation (5) with respect to $n(\boldsymbol{r})$, subject to the condition that the total number of electrons is held fixed; the density $n(\boldsymbol{r})$ that yields this minimum is the ground-state density associated with the given $V(\boldsymbol{r})$.

To explain how these two seemingly innocuous and abstract theorems have revolutionized the modelling of materials, we need to separate the energy $F[n(\boldsymbol{r})]$ of equation (5) into two parts. The first part is the kinetic energy of a system of non-interacting electrons acted on by an external potential which is adjusted so that the ground-state density of this non-interacting system is $n(\boldsymbol{r})$; we denote this by $T[n(\boldsymbol{r})]$. The second part is simply the part of $F[n(\boldsymbol{r})]$ not accounted for by $T[n(\boldsymbol{r})]$, and is denoted by $G[n(\boldsymbol{r})]$. We note that if there were no interactions between the electrons, then by definition $G[n(\boldsymbol{r})]$ would be zero, so in that sense $G[n(\boldsymbol{r})]$ is the part of $F[n(\boldsymbol{r})]$ due to the mutual interaction of the electrons. With this definition, the quantity that has to be minimized to obtain the ground state, according to theorem 2, is

$$E_{\mathrm{tot}}[n(\boldsymbol{r})] = \int V(\boldsymbol{r})n(\boldsymbol{r})\,\mathrm{d}v + T[n(\boldsymbol{r})] + G[n(\boldsymbol{r})]. \tag{6}$$

In the ground state, E_{tot} is stationary with respect to small variations $\delta n(\boldsymbol{r})$ of $n(\boldsymbol{r})$, so that

$$\int V(\boldsymbol{r})\delta n(\boldsymbol{r})\,\mathrm{d}v + \int \frac{\delta T}{\delta n(\boldsymbol{r})}\delta n(\boldsymbol{r})\,\mathrm{d}v + \int \frac{\delta G}{\delta n(\boldsymbol{r})}\delta n(\boldsymbol{r})\,\mathrm{d}v = 0. \tag{7}$$

However, we only allow variations $\delta n(\boldsymbol{r})$ for which the total number of electrons is held fixed, so we have to impose the constraint $\int \delta n(\boldsymbol{r})\,\mathrm{d}v = 0$. This constraint is conveniently incorporated using a Lagrange undetermined multiplier, and we then obtain

$$\frac{\delta T}{\delta n(\boldsymbol{r})} + V(\boldsymbol{r}) + \frac{\delta G}{\delta n(\boldsymbol{r})} = \mu, \tag{8}$$

where the constant μ plays the role of the chemical potential of the electrons.

Now here is the key point. For a system of non-interacting electrons, $G = 0$, so that the third term on the left-hand side of equation (8) vanishes and the ground state is determined by

$$\frac{\delta T}{\delta n(\boldsymbol{r})} + V(\boldsymbol{r}) = \mu. \tag{9}$$

But equation (8) has exactly the same form, provided we define an effective potential $V_{\text{eff}}(\boldsymbol{r}) = V(\boldsymbol{r}) + \delta G/\delta n(\boldsymbol{r})$. The ground state of the system of interacting electrons is therefore given by

$$\frac{\delta T}{\delta n(\boldsymbol{r})} + V_{\text{eff}}(\boldsymbol{r}) = \mu. \tag{10}$$

This means that the ground-state density $n(\boldsymbol{r})$ of a system of interacting electrons acted on by potential $V(\boldsymbol{r})$ is exactly the same as that of a system of non-interacting electrons acted by the effective potential $V_{\text{eff}}(\boldsymbol{r}) = V(\boldsymbol{r}) + \delta G/\delta n(\boldsymbol{r})$.

Now we know how to calculate the ground-state density $n(\boldsymbol{r})$ of a system of non-interacting electrons acted on by a potential $V_{\text{eff}}(\boldsymbol{r})$. To do this, we solve Schrödinger's equation:

$$-\frac{\hbar^2}{2m}\nabla^2\psi_n + V_{\text{eff}}\psi_n = \epsilon_n\psi_n \tag{11}$$

to find the eigenfunctions $\psi_n(\boldsymbol{r})$ and the eigenvalues ϵ_n. The ground-state density is then given by

$$n(\boldsymbol{r}) = 2\sum_n f_n\,|\psi_n(\boldsymbol{r})|^2\,, \tag{12}$$

where the occupation numbers f_n are equal to 1 for states $\epsilon_n < \mu$ below the Fermi energy and 0 for $\epsilon_n > \mu$; the factor of 2 is for spin. From what we have said, the $n(\boldsymbol{r})$ calculated this way is exactly the same as the ground-state density of the interacting electron system. The total ground-state energy of the interacting system can then be obtained by substituting $n(\boldsymbol{r})$ back into equation (6).

The remarkable achievement of DFT is thus to map the system of interacting, correlated electrons acted on by potential $V(\boldsymbol{r})$ onto a system of independent non-interacting electrons acted on by the 'effective' potential $V_{\text{eff}}(\boldsymbol{r})$. In this context, the basic equation (11) having the form of a Schroedinger equation for non-interacting electrons is known as the Kohn–Sham equation, and the potential $V_{\text{eff}}(\boldsymbol{r})$ is known as the Kohn–Sham potential. Note that the Kohn–Sham equation has to be solved in a self-consistent manner. The Kohn–Sham potential $V_{\text{eff}}(\boldsymbol{r}) \equiv V(\boldsymbol{r}) + \delta G/\delta n(\boldsymbol{r})$ depends on the electron density $n(\boldsymbol{r})$, which itself depends on the solutions of the Kohn–Sham equation *via* equation (12). The ground-state density is therefore the $n(\boldsymbol{r})$ such that the $V_{\text{eff}}(\boldsymbol{r})$ constructed from it yields solutions to the Kohn–Sham equation which, when substituted into equation (12), reproduce the same $n(\boldsymbol{r})$.

There is an important generalization of DFT to finite temperatures, in which the electron system is in thermal equilibrium at a specified temperature [178]. The quantity of interest is then the free energy of the system, rather than the ground-state energy. The two basic theorems remain valid, and the form of the theory is also essentially unchanged, except that the occupation numbers f_n are now given by the Fermi–Dirac formula. This generalization is relevant to the present review, because of the very high temperatures in planetary interiors.

To turn this formal scheme into a practical calculation tool, we need an expression for $G[n(\boldsymbol{r})]$. We recognize in advance that an exact expression for $G[n(\boldsymbol{r})]$ is unlikely to be accessible, since if we had this it would be tantamount to solving exactly the quantum mechanics of many-electron systems. So in practice, we are talking about good approximate expressions for $G[n(\boldsymbol{r})]$. We address this question next.

3.4. Approximations to the exchange-correlation energy

We already pointed out that $G[n(\boldsymbol{r})]$ represents in some sense the mutual interaction energy of a system of electrons whose density distribution is $n(\boldsymbol{r})$. Now one large and important part of this interaction energy is the electrostatic Coulomb energy of the charge distribution associated with the electron density $n(\boldsymbol{r})$. Since each electron has charge $-e$, the charge density $\rho(\boldsymbol{r})$ is equal to $-en(\boldsymbol{r})$. So a key part of $G[n(\boldsymbol{r})]$ is the Coulomb energy of the charge distribution $\rho(\boldsymbol{r})$. This energy, often called Hartree energy E_{Har}, is given by

$$E_{\mathrm{Har}} = \frac{1}{2}\int \frac{\rho(\boldsymbol{r})\rho(\boldsymbol{r}')}{4\pi\epsilon_0|\boldsymbol{r}-\boldsymbol{r}'|}\,\mathrm{d}\boldsymbol{r}\,\mathrm{d}\boldsymbol{r}'. \tag{13}$$

The energy $G[n(\boldsymbol{r})]$ is not the same as E_{Har}, for two reasons: first, because of quantum exchange energy (see section 3.1) and, second, because the electrons are correlated. The difference between $G[n(\boldsymbol{r})]$ and E_{Har} is called exchange-correlation energy, denoted by E_{xc}:

$$G[n(\boldsymbol{r})] = E_{\mathrm{Har}}[n(\boldsymbol{r})] + E_{\mathrm{xc}}[n(\boldsymbol{r})]. \tag{14}$$

This exchange-correlation energy E_{xc} is the heart of DFT because it has been found that rather simple approximations to E_{xc} yield approximations to the ground-state density and total energy which are accurate for large classes of materials.

The simplest approximation to E_{xc}, and the one that first demonstrated the remarkable capabilities of the DFT philosophy, is the so-called local density approximation (LDA) [152]. This says that at any position in space the density of exchange-correlation energy depends on the electron density at that point in the same way as in a system of interacting electrons of uniform density, otherwise known as 'jellium'. The exchange-correlation energy per electron $\epsilon_{\mathrm{xc}}(n)$ in the uniform electron gas of density n is accurately known. The LDA then says that E_{xc} in a system of electrons acted on by a potential $V(\boldsymbol{r})$ is given in terms of the jellium $\epsilon_{\mathrm{xc}}(n)$ by

$$E_{\mathrm{xc}} \simeq \int n(\boldsymbol{r})\epsilon_{\mathrm{xc}}(n(\boldsymbol{r}))\,\mathrm{d}v. \tag{15}$$

Practical experience accumulated by hundreds of research groups worldwide over the past thirty years and more has demonstrated that this simple approximation already yields a fairly accurate decription of the energetics of electrons in a wide variety of materials [170], including many of geological interest.

Approximations to E_{xc} that are considerably better than LDA for many purposes have been discovered in the past 15–20 years. Perhaps the most widely used of these are the so-called generalized gradient approximations (GGA) [215–217]. These recognize that the LDA must be inadequate because it does not allow for the dependence of E_{xc} on the rapidity of spatial variations of $n(\boldsymbol{r})$. To remedy this, GGAs include an explicit dependence on the gradient $\nabla n(\boldsymbol{r})$. The physical arguments used to support proposed forms of this dependence are described in the original papers.

A variety of examples of the accuracy of LDA and GGA for geological materials will be given in the later sections of this review. However, before this, we need to indicate more explicitly how it is possible to go from a given approximation for exchange-correlation energy to practical computations on real materials. One of the powerful techniques that has proved crucial here is the pseudopotential method, which we outline next.

3.5. Pseudopotentials

Although DFT tells us in principle how to calculate the electron orbitals and eigenvalues, the practical implementation of this general scheme poses many conceptual and technical problems, which have been addressed by a variety of different approaches, some of which predate DFT itself. One feature that must be addressed is that there are radically different kinds of orbitals in real materials. There are core orbitals, which are strongly localized around the nuclei and have very deep binding energies. At the other extreme, there are valence and conduction orbitals (for brevity, we shall simply say valence orbitals in the following), which by comparison are rather weakly bound and which extend throughout the material. Remarkably, in some materials, particularly the s–p bonded metals and semi-conductors, the energies of the valence bands strongly resemble those of free electrons [120].

Pseudopotential theory [67, 122, 170, 263] is a reformulation of DFT which exploits the features that we have just mentioned to achieve major simplifications without significant loss of accuracy. It recognizes first that the core electrons play little or no role in the energetics of the material, which depends almost entirely on the valence electrons. It therefore sets out to eliminate the explicit description of the core orbitals and to rewrite the Kohn–Sham equations so that their lowest energy levels are the valence states. The way it does this is linked to the similarity of the valence energies to free electron energies in s–p bonded materials. The theory demonstrates that this can be understood in two complementary ways. The first way recognizes that the manner in which valence electrons are scattered by the atomic cores can be accurately mimicked by weak valence-core potentials—these are the 'pseudopotentials' created by the theory. The second way arrives at essentially the same concept of a weak effective valence–core pseudopotential by showing the strong potential attracting the valence electrons to the cores is largely cancelled by the effects of orthogonality between the valence and the core states, which have the effect of a repulsive potential.

The net result is that the total energy of a material can be calculated by solving the Kohn–Sham equations for the valence electrons alone, which interact with the atomic cores *via* pseudopotentials. The modern versions of the theory provide procedures for constructing these pseudopotentials by DFT calculations on free atoms.

In some versions of the pseudopotential method, the pseudopotentials have the property of 'norm conservation', which is an expression of the 'transferability' of the pseudopotential, i.e. the the maintenance of good accuracy when the atom is transferred from one environment to another. However, over the past 10 years, there has been a steadily increasing tendency to use so-called 'ultra-soft' pseudopotentials [263], which are considerably more efficient and in practice are at least as transferable, if not more so. Closely related to the ultra-soft pseudopotential technique is an all-electron technique known as the 'projector augmented wave (PAW)' method [48, 156].

The pseudopotential implementation of DFT is the general approach used in the vast majority of the work we shall review. However, it is important to appreciate that there are a number of important approaches in which the core orbitals are explicitly treated. A common feature of these approaches is that the wavefunctions in the regions near the nuclei are treated by 'atomic-like' methods, while in the regions between the atoms they are treated by 'free-electron-like' methods. An example is the augmented plane-wave (APW) technique, in which the eigenfunctions of the Kohn–Sham equation are represented within spheres surrounding the nuclei by spherical harmonics multiplied by radial functions, and the region between the spheres by plane waves. Somewhat similar in spirit is the Korringa–Kohn–Rostocker (KKR) method, which regards the electronic structure as arising from the multiple scattering of plane waves by the potentials within atom-centred spheres. These and other methods of implementing DFT in condensed matter, as well as a thorough exposition of pseudopotential theory, are provided in the book by Martin [170].

3.6. Basis sets and periodic boundary conditions

A major issue in any quantum mechanical calculation is the representation of the wavefunctions, which in the case of DFT means the Kohn–Sham single-electron orbitals $\psi_n(\boldsymbol{r})$ (see section 3.3). There are two schools of thought about how to do this. One school says that the orbitals of electrons in an assembly of atoms are like the wavefunctions $\phi_\alpha(\boldsymbol{r})$ of the isolated atoms, so that the ψ_n can be represented as superposition of the ϕ_α:

$$\psi_n(\boldsymbol{r}) = \sum_\alpha c_{i\alpha}\phi_\alpha(\boldsymbol{r}). \tag{16}$$

The determination of the ground state then means the determination of the coefficients $c_{i\alpha}$ that yield the lowest DFT total energy (recall theorem 2, section 3.3). The other school of thought says that the valence or conduction electrons in condensed matter run rather freely, so that they are like free particles. Now the wavefunction of a free electron is $\exp(\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r})$, where $\boldsymbol{k}$ is the momentum of the electron divided by $\hbar$ (in other words, $\boldsymbol{k}$ is the wavevector of the de Broglie wave). So the idea is to use these plane waves $\exp(\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r})$ as basis functions:

$$\psi_n(\boldsymbol{r}) = \sum_{\boldsymbol{k}} c_{n\boldsymbol{k}} \exp(\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}), \tag{17}$$

and the coefficients that have to be varied in searching for the ground state are the $c_{n\boldsymbol{k}}$.

The majority of DFT codes designed for the kind of modelling of interest in this review are based on the plane-wave approach, but there is an important minority based on atomic basis sets (see, e.g. [241]), and these play a distinctive role. It should be stressed that it would be completely impossible to use plane-wave basis sets were it not for pseudopotentials. The point is that the core orbitals vary rapidly on an extremely short length scale, so that a prohibitive number of plane waves would be needed to represent them. The representation of valence and conduction orbitals would be equally impossible because orthogonality requires these too to vary rapidly in the region of the atomic cores. Pseudopotentials solve both these problems because the explicit representation of core orbitals becomes unnecessary and because the smoothness of the pseudopotential in the core regions makes the pseudo-wavefunctions of valence and conduction orbitals in these regions manageably smooth.

The use of plane-wave basis sets is closely linked to the question of boundary conditions. In condensed matter modelling, we usually want to study macroscopic assemblies of atoms, but of course the explicit calculations become extremely challenging for more than about a thousand atoms. The standard technique for solving this problem is periodic boundary conditions (pbc). The assembly of atoms on which the explicit calculations are performed is contained in a cell having the shape of a parallelepiped. This cell is periodically repeated, so that it is surrounded by exact images of itself. Another way of saying this is that the material, which may not have any periodicity at all (amorphous or liquid systems are examples), is treated as though it were a crystal having a large, complex unit cell. For this reason, the pbc method is sometimes called the 'supercell' method. Naturally, the imposition of an artificial periodicity causes errors in the calculated quantities, but for most quantities of interest these errors fall off quite rapidly with the cell size—usually at least as rapidly as $1/N_{\mathrm{cell}}$, where N_{cell} is the number of atoms in the repeating cell. This kind of error always needs to be tested for, and efforts should be made to ensure that N_{cell} is large enough to make the size errors negligible. Examples will be referred to in the later sections.

Plane-wave basis sets fit very nicely with pbc, because these boundary conditions make it natural to restrict the wavevectors $\boldsymbol{k}$ in the representation of the orbitals (equation (17)) to those that are commensurate with the imposed periodicity. If this is done, so that the $\boldsymbol{k}$ are reciprocal lattice vectors of the superlattice, then an error is made, which decreases with the size of the repeated cell. The suppression of this error by so-called k-point sampling is a technical matter which is discussed in depth elsewhere [170, 180].

Before leaving the discussion of basis sets, we mention that the use of atomic-like basis functions $\phi_\alpha(\boldsymbol{r})$, sometimes called pseudo-atomic orbitals (PAOs), has great advantages for some purposes, the main advantage being that it is capable of giving much more rapid calculations, though at the cost of some loss of accuracy. The design of accurate, flexible and transferable PAOs has made important progress in recent years [241].

3.7. Atomic forces and stress tensor

A major advantage of the pseudopotential implementation of DFT is that the calculation of the forces on the atoms is straightforward. To be more precise, we should say the forces on the ionic cores, since pseudopotential theory regards the system as composed of a collection of ionic cores and valence electrons.

There are two ways of thinking about the forces on the ions. The first way is to say that the physical force on an ion is the sum of the electrostatic forces due to all the other ionic cores plus the force exerted by the valence electrons. The calculation of the force due to the other ions is straightforward, since it is just a matter of classical electrostatics. The force due to the valence electrons comes from the interaction of the latter with the ionic pseudopotentials and can be calculated from the occupied valence orbitals $\psi_n(\boldsymbol{r})$. The second way of considering the force on an ionic core is to say that it must be (minus) the gradient of the ground-state energy with respect to the position of that ion. The equivalence of the two ways of obtaining the forces is embodied in a statement known as the Hellmann–Feynman theorem [92].

The computation of ionic forces using the Hellmann–Feynman theorem is particularly simple with plane-wave basis sets and adds very little cost to the computer effort required in any case to calculate the ground-state energy. With atomic-like basis sets, forces are somewhat more complicated to compute because the movement of the basis functions with the atoms creates an additional contribution known as the Pulay correction [222]. Provided this is done correctly, the resulting forces are still the exact derivatives of the ground-state energy with respect to the ionic positions.

In the geological context, we often need to model materials under high pressures, so that the pressure, or more generally the stress tensor, is an important quantity. Just as the ionic forces are the derivative of the total energy with respect to the ionic positions, so the stress tensor is the derivative of the total energy with respect to the strain. If we have a system in which the ions have positions $\boldsymbol{r}_i$ then the effect of a uniform strain is to shift the ions to new positions $\boldsymbol{r}'_i$, which are related to the old positions by the transformation

$$r'_{i\alpha} = r_{i\alpha} + \sum_\beta \epsilon_{\alpha\beta} r_{i\beta}, \tag{18}$$

where α denotes the Cartesian components of the vectors, and the 3×3 matrix $\epsilon_{\alpha\beta}$ is the strain tensor. The tensor $\epsilon_{\alpha\beta}$ can be assumed to be symmetric, since any antisymmetric part corresponds to overall rotations of the system, which are of no interest. With these definitions, the stress tensor $\sigma_{\alpha\beta}$ is defined as

$$\sigma_{\alpha\beta} = \frac{1}{\Omega} \frac{\partial U_{\text{tot}}}{\partial \epsilon_{\alpha\beta}}, \tag{19}$$

where Ω is the volume of the system whose total energy is U_{tot}.

In the pseudopotential formulation of DFT, an analytic expression can be derived for $\sigma_{\alpha\beta}$ [170, 194, 195], and it is a standard operation to calculate the entire stress tensor at the same time as the total energy and the forces, though this does add appreciably to the cost of the whole calculation.

3.8. Practicalities

The pseudopotential/plane-wave implementation of DFT is by far the most widely used method used for the quantum-based modelling of complex condensed-matter systems. (The term 'pseudopotential/plane-wave' is not quite accurate here, since we mean to also include the closely related projector augmented wave method—see above.) These techniques have become very firmly established over the past 20 years and are in routine use by hundreds of research groups. Some of the leading groups have developed comprehensively tested codes that implement the techniques. The codes generally come with either a standard library of first-principles pseudopotentials or separate code for generating the pseudopotentials or both. In some cases, the codes can be freely downloaded from web sites or else obtained from the authors at nominal cost. As with any complex technique, it is unwise to try to use the codes without a good understanding of the underlying theory. A powerful aid to this understanding is to learn the use of the chosen code in close collaboration with experienced practitioners.

In the present review, we survey the work of a number of different research groups, and a variety of codes were used. In our own work at UCL, the main workhorse has been the VASP code [153, 154], which is particularly efficient for large-scale DFT molecular dynamics simulation, and has the advantage of being very robust and reliable for metallic systems. Our implementation of a scheme for the extrapolation of the charge density also increased the efficiency of the molecular dynamics simulations by up to a factor of two [5]. However, several other pseudopotential/plane-wave codes have a comparable performance, among which are the ABINIT [112], the CASTEP [236] and the ESPRESSO [234] packages.

The computer requirements for performing these calculations depend to some extent on the material being studied and depend enormously on the number of atoms that need to be handled. Static relaxation calculations are in general much less demanding than DFT molecular dynamics. Static calculations on crystals containing no more than a few tens of atoms in the unit cell can easily be performed on PCs or modest workstations. At the time of writing, extensive calculations on phonon dispersion relations or static relaxation on materials containing defects, surfaces or interfaces require somewhat more powerful facilities, which usually means parallel machines having a modest number of processors. DFT molecular dynamics simulation on systems of up to ~100 atoms typically requires parallel machines having of order 64 processors. Finally, the most ambitious DFT molecular dynamics simulations mentioned in this review, performed on systems containing of order 1000 atoms, can at present only be done on large national facilities. Moore's law will ensure that the foregoing statements become rapidly out of date.

4. Energetics of static crystals: athermal calculations

The simplest possible kind of first-principles calculation consists of determining the total energy of an assembly of atoms in given fixed positions. For very simple crystals, in which the positions of the atoms in the unit cell are completely fixed by symmetry, this elementary kind of calculation, repeated at different unit cell volumes, is already useful because it allows the computation of the equilibrium volume and the zero-temperature equation of state (pressure as a function of volume). The geologically important MgO (the mineral periclase) is an

Table 2. First-principles predictions of properties of MgO crystal at zero temperature and pressure, compared with experiment: lattice parameter a_0, volume of cubic unit cell $V_0 \equiv a_0^3$, bulk modulus K_0, pressure derivative $K_0' \equiv \partial K_0/\partial p$, second pressure derivative $K_0'' \equiv \partial K_0'/\partial p$.

	LDA[a]	LDA[b]	LDA[c]	LDA[d]	GGA[e]	expt[f]
a_0 (Å)	4.240	4.222	4.25	4.167	4.253	4.212
V_0 (Å^3)	76.2	75.2	76.8	72.4	77.0	74.7
K_0 (GPa)	172.6	159	159.7	172	150.6	160.2
K_0'	4.00	4.30	4.26	4.09	4.10	3.99
K_0''	−0.025	−0.030	−0.026	−0.023	−0.027	−0.024

Note: References: [a] [203]; [b] [142]; [c] [137]; [d] [177]; [e] [202]; [f] [243].

example. For less symmetrical crystals, the equilibrium positions of the atoms in the unit cell must be determined by using the first-principles forces to relax the system. Examples include $MgSiO_3$ and water ice. With calculations such as these, it is also possible to compare the stability of different crystal structures as a function of pressure and hence to determine low-temperature phase boundaries. A slight generalization of these ideas leads to the first-principles calculation of the stress in a crystal as function of strain, and hence the elastic properties. Static relaxation performed on defective crystals allows one to compute quantities such as the formation and migration energies of point defects (vacancies, impurities, etc) or the energetics of dislocations, grain boundaries and surfaces, all of which are extremely important in the physics and chemistry of the Earth and the planets. Caution is often needed in using the results of zero-temperature static calculations, particularly for problems involving high temperatures. Nevertheless, these simple and routine calculations form the essential baseline for the more sophisticated high-temperature calculations described later.

In reviewing work on low-temperature energetics, we shall start by describing first-principles calculations on the crystal structure and equation of state of some of the materials that are crucial in the earth and planetary sciences. Illustrations of work on elastic properties will then follow. The section ends with some examples of work on crystal defects of various kinds.

4.1. *Crystal structure and equation of state*

4.1.1. Magnesium oxide. Because of its simplicity and its geological importance, MgO has received a large amount of experimental and theoretical attention. (We recall from section 2.1 that, in the form of magnesiowüstite, with approximate composition $(Mg_{0.8}Fe_{0.2})O$, it has ~20% abundance in the Earth's lower mantle.) Because of this, there has been strong interest in determining its phase diagram. Diffraction measurements performed in the diamond anvil cell [88] show that, over the whole pressure range from 0 to 227 GPa, MgO has the rock-salt structure, though it is generally agreed that the CsCl structure becomes more stable at much higher pressures. The high-pressure melting curve of MgO is also of great interest.

DFT calculations on MgO go back to the early 1980s [64,86,132,137,177,202,203]. First-principles and experimental values of the lattice parameter a_0, the bulk modulus K and its first two pressure derivatives $K' \equiv \mathrm{d}K/\mathrm{d}p$ and $K'' \equiv \mathrm{d}K'/\mathrm{d}p$ at $T = 0$ are compared in table 2. The agreement of a_0 to within $\sim 1\%$, and of K within ~5% is rather typical for this kind of ionic material. The low-temperature equation of state $p(V)$ is also good (figure 3), though we note significantly better agreement with GGA than with LDA. However, the inclusion of room temperature effects would shift both curves upwards, with the experimental results falling roughly in the middle of the LDA and the GGA curves.

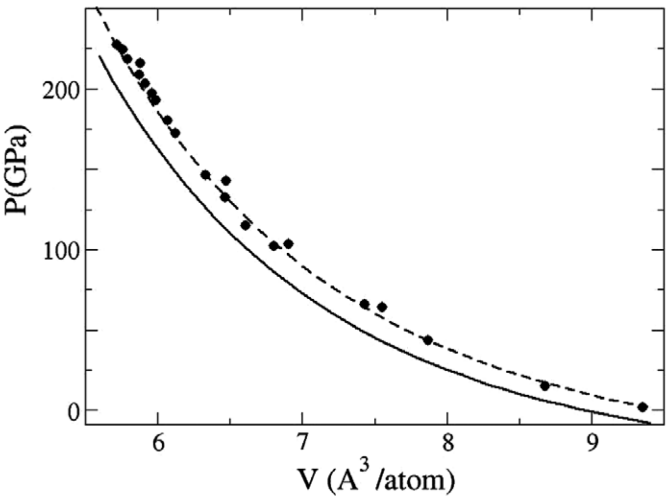

Figure 3. Pressure as function of atomic volume for rock-salt MgO calculated with DFT-LDA (——) and DFT-GGA (- - - -) compared with experiments ($*$) [88]. Calculations do not include zero-point or thermal effects. Reprinted with permission from [7]. Copyright 2005 American Physical Society.

Even though MgO is such a simple material, the first-principles calculation of the rock-salt/CsCl structural transition raises tricky technical issues that are important for all phase diagrams. The curves of enthalpy as a function of pressure for the two structures cross at a very shallow angle, so that small errors that do not cancel between the two phases can lead to large spurious shifts of the transition pressure. However, the consensus from the most recent calculations is that the transition occurs in the region of 500 GPa (see [202] and references therein). This is far above the pressure at the base of the Earth's mantle (135 GPa), and the implication is that magnesiowüstite in the mantle must be in the rock-salt structure. To be sure of this, we need to know the temperature variation of the transition pressure up to $\sim$ 4000 K, and we return to this in section 5.6.

4.1.2. Magnesium silicate. Since $MgSiO_3$ perovskite is the major component in the Earth's mantle (section 2.1), it is outstandingly important. Although this crystal structure is only metastable at ambient pressure, it can be experimentally studied under these conditions. However, the challenges in obtaining accurate experimental data under lower mantle conditions mean that first-principles calculations have been invaluable in obtaining the properties of $MgSiO_3$-perovskites in the Earth. For example, in 1993 three papers were published by different authors using different first-principles methods: LDA and pseudopotentials [283], LDA and LAPW [250], and Hartree–Fock theory [76]. All the calculations predicted that the orthorhombic Pbnm phase of $MgSiO_3$ is more stable than the higher symmetry cubic and tetragonal perovskite phases throughout the pressure range of the Earth's mantle; this important conclusion still stands today.

In table 3, we show first-principles predictions for the three lattice parameters a, b and c of $MgSiO_3$ perovskite at zero temperature and pressure obtained from DFT calculations, compared with experimental values. The table also shows the fractional coordinates x, y and z specifying the relaxed positions of the Mg ion and the two inequivalent O ions (the Si position and the z-coordinates of Mg and O_1 are fixed by symmetry). If the 298 K temperature difference

Table 3. Structure of $MgSiO_3$ perovskite crystal at ambient conditions: comparison of theoretical and experimental values of orthorhombic cell parameters a, b and c (Å units), and components of Mg and O positions not constrained by symmetry (fractions of cell parameters). Theoretical values are from Hartree-Fock (HF) approximation and from LDA and GGA approximations within DFT; GGA+P refers to pressure-corrected GGA values (see text).

	HF[a]	LDA[b]	LDA[c]	LDA[d]	GGA[e]	GGA+P[f]	Expt[g]
a	4.786	4.711	4.789	4.727	4.834	4.777	4.775
b	4.913	4.880	4.922	4.889	4.983	4.933	4.932
c	6.896	6.851	6.893	6.831	6.977	6.892	6.899
Mg x	0.5112	0.5174	0.5129	0.5160	0.5148	0.5154	0.5143
y	0.5476	0.5614	0.5521	0.5576	0.5552	0.5570	0.5556
O_1 x	0.0999	0.1128	0.0980	0.1044	0.1018	0.1030	0.1037
y	0.4643	0.4608	0.4686	0.4654	0.4664	0.4666	0.4655
O_2 x	0.1980	0.1928	0.1983	0.1969	0.1983	0.1972	0.1974
y	0.2029	0.1995	0.2033	0.2015	0.2031	0.2019	0.2011
z	0.5522	0.5582	0.5507	0.5539	0.5527	0.5530	0.5538

Note: References: [a] [76, 75]; [b] [282]; [c] [141]; [d] [198]; [e] [200]; [f] [200]; [g] [228].

between experiment and theory is taken into account, the LDA results are in good agreement with the experimental data. In order to correct for the GGA error, Oganov *et al* [201] showed that adding a simple pressure correction of about −8.5 GPa to the GGA pressure resulted in an extremely good agreement between calculated and experimental lattice parameters, bulk modulus and elastic constants of $MgSiO_3$ perovskite. For instance, the GGA bulk modulus corrected for pressure was 267 GPa, compared with the experimental value of 264 GPa found by Yeganeh–Haeri [289]. More recent experimental data, however, show that the bulk modulus may in fact be 253 GPa [239], so that the simple pressure correction of Oganov *et al* may need to be reevaluated. We note from table 3 that in all the cases the theoretical values of the relaxed ionic positions are in excellent agreement with experiment. Comparison of the calculated $p(V)$ curve with experiment (figure 4) shows the same excellent agreement over the whole range of mantle pressures that we saw for MgO (figure 3).

Perovskite in the lower mantle is not pure $MgSiO_3$ and contains up to approximately 10% iron (as both ferric and ferrous) and 5% Al_2O_3. The effect of Al_2O_3 became a topic of interest when Zhang and Weidner [293] predicted that small amounts of Al_2O_3 in mantle perovskite may have an unexpectedly large effect on its compressibility. GGA calculations by Brodholt [54], Yamamoto *et al* [288] and Walter *et al* [277] showed that this could only be achieved if Al^{3+} was substituted for Si^{4+}, with the charge balanced by oxygen vacancies. First-principles work on the effect of iron on the properties of perovskite has been limited, partly because standard DFT methods fail to predict the correct electronic structure of FeO, predicting it to be metallic under ambient conditions, whereas in fact it has a sizeable band-gap [70]. Kiefer *et al* [143] have used LDA and GGA to predict the effect of Fe^{2+} on the elasticity of magnesium-silicate perovskite containing 25% FeO, and there are works in progress calculating the spin state of Fe^{2+} and Fe^{3+} in mantle perovskites at very high pressures [162].

We noted in section 2 the existence of the mysterious D″ zone at the base of the mantle, characterized by marked seismic anisotropy, and having a very variable thickness. One suggested explanation was that $MgSiO_3$ transforms from Pbnm perovskite to another structure in this region, but until very recently there had been no evidence for this. The situation changed dramatically in 2004, when Murakami *et al* [187] and Oganov and Ono [207] reported high-pressure diffraction experiments showing that a new crystal structure having the CmCm $CaIrO_3$ structure appears at the pressure of the D″ zone. This new

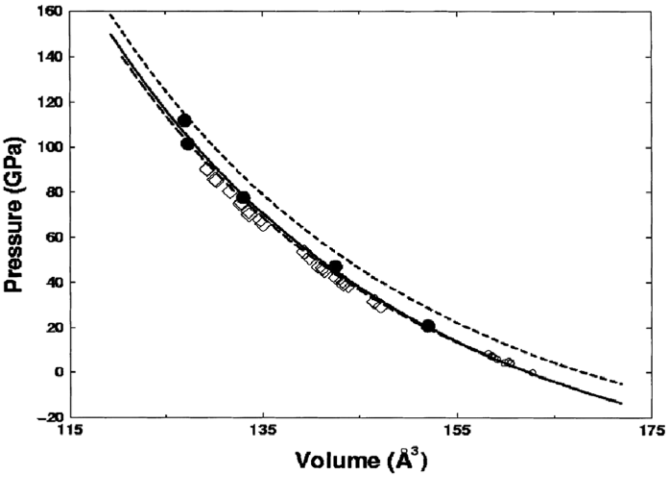

Figure 4. Pressure as function of atomic volume for $MgSiO_3$ perovskite calculated with DFT-LDA (lower dashed line) and DFT-GGA (upper dashed line). Solid line shows DFT-GGA results corrected to include thermal pressure at room temperature. Small and large circles and diamonds show experimental data. Reprinted from [200] with permission. Copyright 2001 Elsevier.

structure (figure 5), christened 'post-perovskite', was immediately investigated by first-principles calculations [129, 207, 261], which confirmed that at low temperature its enthalpy falls below that of orthorhombic perovskite at a pressure of ∼100 GPa (the value depends somewhat on the exchange-correlation functional). However, the major contribution of DFT calculations to this story has been to demonstrate that the elastic properties of the new phase do account for the seismic anistropy and to relate the variable thickness of the D″ zone to the Clapeyron slope of the boundary between the two phases, as will be described in sections 4.2.2 and 5.6.

4.1.3. Silica. Although an end-member silica (SiO_2) phase is not thought to be a large component of the Earth's interior (section 2.1), it has been postulated as being a minor phase depending on the exact composition of the Earth's mantle or as being a major phase in some restricted parts of the mantle. For example, Dobson and Brodholt [84] suggested that ultra-low velocity zones (regions of very low seismic velocity found in the bottom 200 km or so of the mantle) may be the result of the subduction of banded iron formations (BIFs). These BIFs are composed of layers of iron oxide and layers of silica. If, as suggested by Dobson and Brodholt, they are subducted to deep in the Earth's interior, the SiO_2 component will transform into its higher pressure polymorphs. At low pressures, silica exists as quartz (trigonal, 4-fold coordinated Si), but above 2 GPa it transforms to coesite (monoclinic) and above 8 GPa to stishovite (tetragonal, rutile structure, 6-fold coordinated Si). Above these pressures, Kingma *et al* [144] reported both experimental measurements and DFT calculations to confirm that stishovite transforms at about 45 GPa to a $CaCl_2$ structure, a pressure corresponding roughly to a depth of about 1200 km, at which there is a weak seismic discontinuity. Further first-principles calculations showed that the $CaCl_2$ structure breaks down to a PbO_2 structured phase which is stable up to about 200 GPa, where it then transforms to a pyrite-type phase (cubic, 6 + 2-fold coordinated Si) [138, 204, 259]. This last transformation has very recently been confirmed experimentally [158], albeit at a slightly higher pressure.

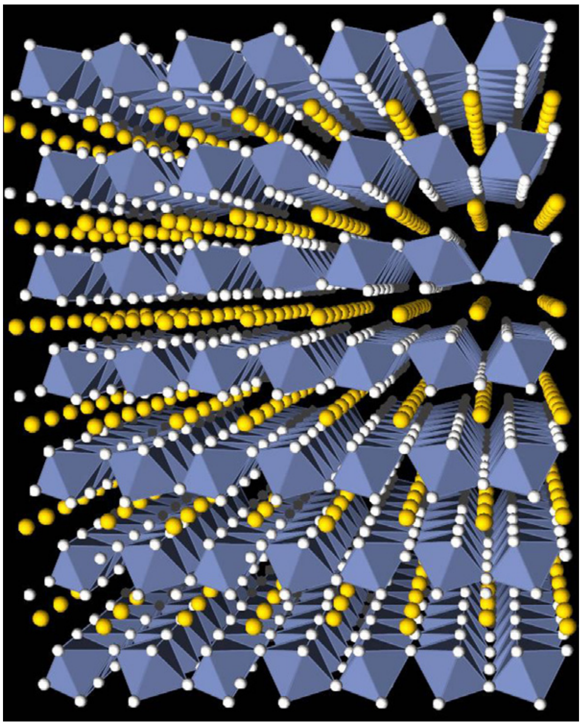

Figure 5. Crystal structure of post-perovskite $MgSiO_3$, showing SiO_6 octahedra with oxygen ions (small spheres) at their vertices and Mg ions (large spheres).

4.1.4. Iron. Since the solid inner core of the Earth is composed mainly of Fe (section 2.1), an understanding of the crystal phase diagram of Fe is extremely important. It has long been known that at ambient pressure the ferromagnetic α (bcc) structure is stable below 1180 K and the non-magnetic γ (fcc) phase is stable above that, with the non-magnetic δ (bcc) phase appearing in a narrow temperature range just below the melting temperature at 1810 K (see phase diagram, figure 6). At low temperatures, there is a transition from the bcc structure to the ϵ (hcp) phase at a pressure of 10–15 GPa, and ϵ appears to remain stable at all pressures of interest above that. At high p and T, there have been reports of new crystal structures [33,232], so far unconfirmed. In the absence of firm evidence to the contrary, there has been a general presumption that hcp is the stable phase at inner-core conditions ($p \sim 330$ GPa, $T \sim 5500$ K). As a baseline for high-temperature studies, the comparison of $T = 0$ DFT calculations with experimental data is clearly crucial.

Extensive static calculations on different crystal structures of Fe using the full-potential LAPW [253] and LMTO [240] implementations of DFT were reported about 10 years ago. It was already known then that the choice of exchange-correlation functional is important. In particular, at zero pressure LDA incorrectly predicts that the non-magnetic fcc structure is more stable than ferromagnetic bcc [278]. However, the use of GGA brings a large improvement, with ferromagnetic bcc now correctly predicted as the most stable structure at zero pressure and the experimental bcc → hcp transition pressure in the region of 12 GPa also correctly given [240, 253]. GGA also gives good agreement with experiment for the pressure–volume relation at $T = 0$ K, the agreement improving steadily with increasing pressure up to the maximum experimental pressure of ~300 GPa. Essentially identical results

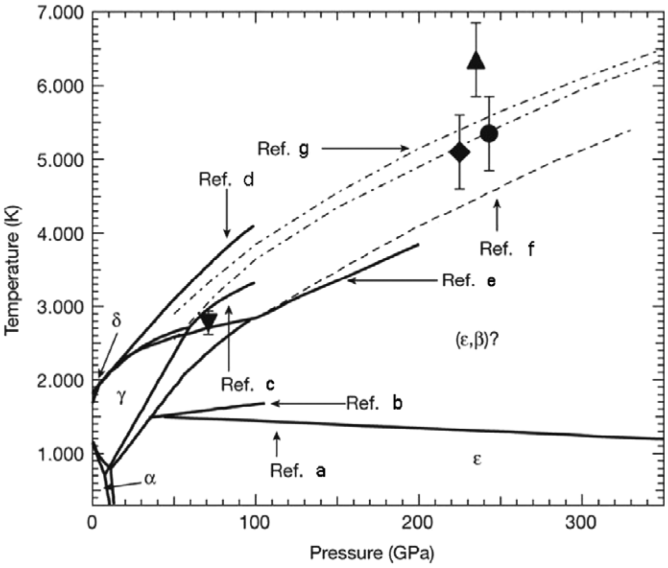

Figure 6. Phase diagram of iron. Solid lines represent phase boundaries and melt lines from DAC experiments, symbols with error bars are points on melting curve from shock experiments, and broken lines are melting curves from first-principles calculations. DAC data: a: [231], b: [32], c: [238], d: [285], e: [49]. Shock data: triangle: [290]; circle: [57]; reverse triangle: [1]; diamond: [193]. First-principles data: f (dashed curve): [159], g (pair of chain curves): [23]. Reproduced with permission from [193]. Copyright 2004 Macmillan Publishers Ltd.

were later reported with the pseudopotential and projector-augmented-wave implementations of DFT [18].

Detailed studies have also been reported on the relative enthalpies of the bcc and hcp structures, as well as of all the other candidate structures (fcc, body-centred tetragonal (b.c.t.), double hcp and orthorhombically distorted hcp), with the magnetic effects included through spin-polarized calculations [240, 246, 253, 272]. These studies confirm that hcp remains the most stable structure up to Earth's core pressures at $T = 0$; weak anti-ferromagnetism stabilizes the hcp structure at lower pressures, the inclusion of this effect improving the agreement with the experimental pressure–volume relation. It has also turned out that bcc, bct and orthorhombically distorted hcp become mechanically unstable at pressures well below those of the Earth's inner core [251, 252, 267, 272]. These results, while useful, do not settle the question of the crystal structure in the inner core because thermal effects still need to be included (see sections 5.6 and 6.4).

4.1.5. Iron silicide and carbide. Silicon is likely to be one of the light elements alloyed with iron in the Earth's core. In a first-principles study of the relative stability of the FeSi structure and of some hypothetical polymorphs of FeSi, Vocadlo *et al* [274] predicted that a new high-pressure phase having the CsCl structure should be the most stable; subsequently, Dobson *et al* [85] performed high-pressure multi-anvil experiments in which this proposed phase was synthesized for the first time, confirming the results of the calculations. The implications of this are extremely significant because the high-pressure phase is topologically equivalent to the bcc phase of iron; moreover, recent experiments [163] have found that the presence of silicon stabilizes bcc Fe under pressure with respect to the hcp phase. This suggests the bcc structure of iron containing some substitutional Si as a candidate for the stable phase in the Earth's core.

In a critical examination of the role of carbon in the inner core, Wood [286] proposed that iron carbide (Fe_3C) might be the major phase crystallizing out of the liquid outer core. Amongst other evidence, he based this suggestion on an equation of state which led to a density at inner-core pressure (338 GPa) in excellent agreement with the probable range of density determined from seismology. However, the equation of state he used was based on estimated values of the bulk modulus, K_0, and of its pressure derivative K' at ambient pressure and temperature. Diffraction experiments performed to determine the thermal expansion of Fe_3C showed a kink at about 480 K, the point at which the permanent magnetism of Fe_3C is destroyed by heat. These results immediately cast doubt on the applicability to the study of the Earth's core of any measurements of the incompressibility of Fe_3C made at room temperature, since under core conditions ($330 < p < 360$ GPa and ~6000 K), this material will no longer be magnetic. An investigation into the effect of pressure on the equation of state of Fe_3C using first-principles calculations was performed for both 'magnetic' and 'non-magnetic' forms of Fe_3C. As expected, a transition to a non-magnetic state was found at a pressure of about 60 GPa. By applying an approximate temperature correction to this calculated equation of state, incompressibility-density plots for Fe_3C under core conditions could be constructed, which clearly show that Fe_3C cannot be a major core-forming phase.

4.1.6. Water and ammonium dihydrate. We mentioned in section 2.3 the need for data on water ice and related solids such as ammonium dihydrate (ADH) in order to interpret the properties of planetary moons. First-principles calculations by Fortes [96] on the low-pressure proton-ordered phases of ice, ice XI (the proton-ordered form of normal water ice, phase Ih) and ice II (occurring at 0.2–0.5 GPa below ~250 K) [97], yield equations of state that are rather poor when compared with experiment, with large overestimates in V_0 and K, particularly in the case of ice XI. This is thought to result from the overbinding of hydrogen-bonded solids in the DFT calculations—an effect which is greatly reduced as the pressure increases. However, when compared at the same volume, the structural properties of these ice phases are in excellent agreement with experiment [97, 96].

Calculations on the high pressure phases of water ice show excellent agreement with experiment [96]. For example, the transition from ice VIII to ice X is predicted to occur at ~ 70 GPa compared with the experimental value of 60–65 GPa. Furthermore, these simulations predict higher pressure transitions to dhcp ice (500–600 GPa) and antifluorite ice (2.5 TPa).

ADH is very likely to be the most significant mineral after water ice in the mantles of ice-rich moons orbiting Saturn, Uranus and Neptune, but the physical properties and phase relations of ADH are poorly constrained. There are currently two known phases (I and II) at low pressures. Neutron diffraction experiments at 150 K [192–191] suggest that ADH I is cubic; the water molecules are orientationally disordered with partially occupied hydrogen sites. First-principles calculations [98] show that there is a slight monoclinic distortion from the proposed cubic structure at zero Kelvin; the suggestion is that ADH becomes ordered at some temperature below ~150 K, the temperature at which experimental observations suggest discontinuities in thermal conductivity and dielectric properties [164]. This low-temperature ordered structure has yet to be found experimentally. Neutron diffraction experiments by Fortes *et al* [99] have revealed a wealth of previously unknown phases at pressures up to 9 GPa. There are at least four new polymorphs of ADH at pressures of 0.5–2 GPa, the pressure regime of particular interest in the study of icy moons.

4.2. Elastic properties

Since most of our understanding of the Earth's interior derives from seismic measurements, the elastic properties of Earth-forming materials are outstandingly important. These properties are usually described by the elastic constant tensor, whose components give the coefficients of proportionality between infinitesimal changes in strain and the resulting changes in stress. The theory of elasticity is expounded in many standard texts and reviews, including Nye [197], Wallace [276] and Born and Huang [51]. Elasticity theory for solids under a pre-existing stress is somewhat more complex than for unstressed solids, and a paper that disentangles the complexities in a clear way is by Barron and Klein [39].

Briefly, if a crystal suffers an infinitesimal strain, described by strain tensor $d\epsilon_{\alpha\beta}$, then the resulting change of stress $d\sigma_{\alpha\beta}$ is given by

$$d\sigma_{\alpha\beta} = \sum_{\gamma\delta} c_{\alpha\beta\gamma\delta} d\epsilon_{\gamma\delta}, \tag{20}$$

where $c_{\alpha\beta\gamma\delta}$ are the components of the rank-4 tensor of elastic constants. Since the stress and strain tensors can both be assumed symmetric, equation (20) is usually rewritten in the Voigt notation, in which the 3×3 tensors $d\epsilon_{\alpha\beta}$ and $d\sigma_{\alpha\beta}$ are regarded as 6-component vectors, with the index pairs 11, 22, 33, 23, 31 and 12 represented by the six symbols 1, 2, 3, 4, 5 and 6, respectively. In this notation, the stress–strain relation appears as

$$d\sigma_i = \sum_j C_{ij} d\epsilon_j, \tag{21}$$

with i and j going from 1 to 6. Elastic constants are usually given as the coefficients C_{ij} in this notation. The matrix C_{ij} is symmetric, so that the maximum number of independent elastic constants of a crystal is 21. Because of crystal symmetries, the number of independent constants is usually much smaller. For example, in cubic crystals there are only three elastic constants, in hcp Fe there are five and in orthorhombic $MgSiO_3$ perovskite there are nine.

4.2.1. $MgSiO_3$ perovskite. The low-temperature elastic tensor $C_{\alpha\beta}$ for $MgSiO_3$ perovskite was first calculated using first-principles methods by Wentzcovitch *et al* [282] using LDA and Troullier-Martins pseudopotentials. Seven of the nine elastic constants of the orthorhombic phase agreed with the experimental values [289] to within 10%, though C_{66} differed by ca 15% and C_{44} by ca 25%. Later LDA calculations by Karki *et al* [136] obtained a set of elastic constants in much better agreement with the data, with the worst discrepancy being only 11% and most discrepancies being less than 5%. The difference in the two calculations must be due to the different pseudopotentials used, so the better agreement of the results of Karki *et al* may be fortuitous. Oganov *et al* [200, 201] used GGA to calculate the elastic constants of $MgSiO_3$ perovskite, but in order to account for the general tendency of GGA to overestimate volumes and underestimate bulk moduli, they implemented a small pressure correction. They also took into account the temperature difference between the athermal calculations and the ambient temperature of the experiments. Using this scheme, they produced a set of elastic constants of which eight out of nine agreed with experimental data to within 4%, and the other agreed to within 10%. The very recent experimental data of Sinogeikin *et al* [239] are in slightly worse agreement with the calculated values of Oganov *et al* , but nevertheless all but one of the elastic constants agree to within 6%. The predicted shear modulus of Oganov *et al* is 179 GPa, which agrees within error with both experimental data sets. The predicted bulk modulus, however, is slightly higher than the latest experimental data (267 GPa versus 253 GPa, respectively).

At higher pressures, experimental data on the bulk modulus are obtained from the volumes measured as a function of pressure in a diamond anvil cell or a multi-anvil device. The high pressure volumes are then fitted to an equation of state (such as that due to Birch and Murnaghan) where the pressure effect on the bulk modulus K is reported by a single parameter as $K' = \mathrm{d}K/\mathrm{d}p$. The quantities K and K' are, however, highly correlated with each other and with the zero-pressure volume V_0, so that K' obtained from different experiments or theory cannot be compared independently of K and V_0. When fitted to such an equation of state, the GGA calculations of Oganov *et al* [200, 201] produce $V_0 = 162.4$ Å^3, $K = 267$ GPa and $K' = 4.10$. This compares well with experimental data of Fiquet *et al* [95] of $V_0 = 162.3$ Å^3, $K = 259$ GPa and $K' = 3.7$. The main difference is that the experiments find that K' is less than the commonly assumed value of 4.0.

4.2.2. $MgSiO_3$ post-perovskite. In section 4.1.2, we outlined the recent discovery that a new 'post-perovskite' crystal structure of $MgSiO_3$ becomes more stable than perovskite at a pressure corresponding to that of the D$''$ zone at the base of the mantle. In addition to showing that the new phase is thermodynamically more stable than perovskite, the athermal elastic constants calculated using GGA or LDA by three groups [129, 207, 262] showed that the post-perovskite phase is elastically very anisotropic. This high anisotropy means that shear-induced lattice-preferred orientation of post-perovskite can explain the observed seismic anisotropy of D$''$. This anisotropy of D'' has always been difficult to explain with perovskite itself, since the experimentally observed slip systems in perovskite do not lead to the correct type of seismic anisotropy. Subsequently, another DFT work has examined the effect of temperature and the effect of Al_2O_3 and ferrous Fe on the stability and elasticity of post-perovskite. It is fair to say that the DFT calculations have provided the crucial link between the experimental discovery of post-perovskite and the seismic properties of D$''$.

4.2.3. Iron. The problems of interpreting seismic data on the Earth's inner core, including its observed seismic anisotropy, were mentioned earlier (section 2.1). These problems stimulated early efforts to calculate the elastic properties of perfect-crystal hcp Fe and other hcp transition metals [246, 252]. The evidence from this work suggests that GGA is capable of giving fairly good results for the elastic constants, though attempts to compare them with existing experimental data for Fe up to ∼210 GPa are hampered by uncertainties about the assumptions used to interpret the data. The DFT results indicated that at inner-core pressures the difference in compressional sound velocities parallel to and perpendicular to the hexagonal axis would be enough to account for the observed seismic anisotropy, if there was a high degree of alignment of the crystallites in the inner core. However, it has since been appreciated that the complete absence of thermal effects in the calculations casts doubt on the usefulness of this conclusion [247].

4.3. Crystal defects

First-principles calculations on defects and impurities on Earth materials at high pressures are rather rare and have mostly been restricted to olivines. Brodholt [53] used DFT methods to calculate the relative energetics of different point defects in Mg_2SiO_4 olivine (forsterite), which is generally thought to be the major phase in the Earth's mantle down to about 410 km. Forsterite has three inequivalent oxygen sites and two inequivalent magnesium sites, and the calculations showed that there are large energetic differences between the vacancies on the different sites. For instance, oxygen vacancies on the O3 site are between 1 and 3 eV lower in energy than those on the O2 and O1 sites. Similarly, magnesium vacancies on the M1 sites are 0.8 eV lower in energy than those of the M2 site. Assuming that diffusion of magnesium occurs by a vacancy-hopping mechanism, the observed anisotropy in diffusion rates with respect to crystallographic orientation could be understood by the fact that jumps must only be between M1 to M1 sites. A later TEM study [133] confirmed that magnesium vacancies were indeed restricted to the M1 site.

The role of hydroxyl defects in nominally anhydrous mantle minerals (i.e. those that contain water as a defect and are normally water-free) has been the subject of much interest since it was empirically observed that mantle minerals such as olivine, pyroxenes and garnet can contain significant amounts of water [40] and that small amounts of water can strongly affect the transport properties of these minerals (i.e. hydrolytic weakening). The solubility of water in quartz and the effect on hydrolytic weakening were studied using DFT methods by McConnell *et al* [175]. The calculations showed that hydrogen was incorporated into quartz via a hydrogarnet type substitution ($4H^+$ substituting for one Si^{4+}). They proposed that the subsequent conversion of these defects to molecular water involved the nucleation of prismatic dislocation loops and that the development of these loops led to dislocation-induced plasticity. Brodholt and Refson [55] proposed a very different mechanism for hydrolytic weakening in forsterite. They noted from their DFT calculations that the energy involved in the hydrogarnet substitution was comparable to substituting $2H^+$ for a Mg^{2+} ion. In dry forsterite, the energetic difference between Si^{4+} and Mg^{2+} vacancies is huge and the concentration of Si^{4+} vacancies is vanishingly small, but in wet forsterite, the energy differences are relatively small. They postulated, therefore, that the role of water was simply to greatly increase the concentration of Si^{4+} vacancies (albeit filled with protons) and that this would have the effect of increasing the mobility of Si^{4+} ions. Since Si^{4+} is the slowest diffusing species in olivine and, therefore, rate limiting in any diffusion-controlled creep (i.e. diffusion creep or when rate limited by dislocation climb), this could provide a mechanism for hydrolytic weakening of the olivine. Later work by Braithwaite *et al* [52] on protons in forsterite used an embedded cluster model which calculated the total energy of the central core of atoms constraining the defect using Hartree–Fock theory and the outer atoms with a classic shell potential. In addition to confirming the general findings of Brodholt and Refson [55], they calculated the infrared frequencies of the different hydrous defects. By comparing with experimental data, they suggested that the most common hydrous defect in olivine is three protons subsituting for Si^{4+}, rather than the hydrogarnet substitution.

First-principle simulations on defects in Earth-forming magnesium silicate perovskites have been restricted to understanding the mechanism for the incorporation of Al^{3+}. This has implications in understanding the role of Al^{3+} on the elastic properties of mantle perovksites. The two models considered were, first, Al^{3+} substituting for Si^{4+}, with the charge balanced by oxygen vacancies; and second, Al^{3+} substituting for Mg^{2+} and Si^{4+}, which does not require the formation of oxygen vacancies. The first of these may have important implications for the Earth's mantle since perovskites with a large concentration of oxygen vacancies should have very different transport properties from those that are relatively defect poor. Both the DFT results of Brodholt [54] and Yamamoto *et al* [288], however, suggest that at high-pressures Al^{3+} will substitute for both Mg^{2+} and Si^{4+} and, therefore, mantle perovskites will not have an unusually high concentration of oxygen vacancies.

5. High-temperature properties from lattice dynamics

For many, perhaps most, problems in the Earth and planetary sciences, it is not enough to consider the properties of materials at zero temperature. This is self-evident if we are talking of liquids, but is also true of high-temperature solids, since the elastic and other properties often vary significantly with temperature. This means including entropy effects, so that we need to calculate not just energies but free energies. By far the simplest way of doing this is to use quasi-harmonic theory to treat lattice vibrations. This is, of course, an approximation, since it means ignoring anharmonic effects. However, it is a good approximation for most materials up to about half the melting temperature.

There are also other reasons for calculating lattice vibration frequencies, apart from the desire to predict thermodynamics. One very strong reason is that phonon frequencies give a powerful way of checking the validity of the DFT methods against experiment. Phonon dispersion relations have been measured by inelastic neutron scattering for many materials, including some of geological interest. The large quantity of information in these dispersion relations allows rather detailed tests of DFT exchange-correlation functionals. Until recently, the experimental phonon data were available almost exclusively at ambient pressure, but modern synchrotron techniques are now providing phonon data at pressures well into the megabar regime, and this gives an important new way of validating DFT at high pressures. Because of the close link between phonon frequencies and thermodynamics, such comparisons also give an indirect control of the thermodynamic predictions of DFT.

5.1. *Thermodynamics from quasi-harmonic theory*

The general ideas behind the harmonic calculation of thermodynamic properties are easy to understand, and we summarize them here. In a harmonically vibrating crystal in thermal equilibrium, the total Helmholtz free energy $F(V, T)$ as a function of volume and temperature can be separated into the energy $E_{\mathrm{perf}}(V)$ of the static perfect crystal, which we discussed at length in section 4, and the free energy $F_{\mathrm{harm}}(V, T)$ associated with the harmonic lattice vibrations:

$$F(V, T) = E_{\mathrm{perf}}(V) + F_{\mathrm{harm}}(V, T). \tag{22}$$

(If the temperature is high enough for thermal excitation of electrons to be significant, the finite-temperature version of DFT may be needed, and E_{perf} should then be interpreted as the *free* energy of the static perfect lattice.) Now it is shown in all the standard text books that the vibrational free energy of a single harmonic oscillator of frequency ω is $k_{\mathrm{B}}T \ln(2\sinh(\hbar\omega/2k_{\mathrm{B}}T))$. Since F_{harm} is the sum of the vibrational free energies of all the modes of vibration of the crystal, it is given by

$$F_{\mathrm{harm}} = k_{\mathrm{B}}T \sum_{n} \ln(2\sinh(\hbar\omega_n/2k_{\mathrm{B}}T)), \tag{23}$$

where ω_n is the frequency of the nth vibrational mode.

In fact, each mode of vibration of a perfect crystal is associated with a wavevector $\boldsymbol{k}$, and for each such $\boldsymbol{k}$ there is a number of modes equal to three times the number of atoms in the primitive unit cell. If the frequency of the sth mode at wavevector $\boldsymbol{k}$ is denoted by $\omega_{\boldsymbol{k}s}$ then the vibrational free energy is

$$F_{\mathrm{harm}} = k_{\mathrm{B}}T \sum_{\boldsymbol{k}s} \ln(2\sinh(\hbar\omega_{\boldsymbol{k}s}/2k_{\mathrm{B}}T)). \tag{24}$$

If we can obtain the vibrational frequencies $\omega_{\boldsymbol{k}s}$ from first-principles calculations, this provides a way of calculating the Helmholtz free energy as a function of volume and temperature, and by appropriate differentiation of the result we can obtain all other thermodynamic quantities.

A particularly important quantity that can be obtained like this is pressure p. Since $p = -(\partial F/\partial V)_T$, equation (22) implies that the pressure has two parts:

$$p = p_{\mathrm{perf}} + p_{\mathrm{harm}}, \tag{25}$$

where $p_{\mathrm{perf}} = -(\partial E_{\mathrm{perf}}/\partial V)$ is the contribution from the static lattice and $p_{\mathrm{harm}} = -(\partial F_{\mathrm{harm}}/\partial V)_T$ is the contribution due to lattice vibrations. The part due to lattice vibrations, sometimes called thermal pressure, exists because the harmonic frequencies $\omega_{\boldsymbol{k}s}$ depend on the volume of the crystal. In most materials, compression causes the vibrational frequencies to increase, so that $V\,\partial\omega_{\boldsymbol{k}s}/\partial V$ is negative. This means that $p_{\rm harm}$ is usually positive, which is why materials expand on heating. In fact, $p_{\rm harm}$ is usually positive even at zero temperature, where the harmonic free energy becomes $F_{\rm harm} \to \sum_{\boldsymbol{k}s}(\frac{1}{2})\hbar\omega_{\boldsymbol{k}s}$, i.e. the sum of zero-point energies of the individual modes; the resulting zero-point pressure gives a slight increase in the zero-temperature volume. It is because thermal expansivity can be obtained from the calculation of volume-dependent harmonic frequencies that the term ‘quasi-harmonic’, rather than ‘harmonic’ is generally used. The point here is that in a perfectly harmonic system, the frequencies would not depend on volume, and the thermal expansivity would be zero.

The quasi-harmonic contribution to the free energy is particularly easy to analyse at high temperature, where $\hbar\omega_{\boldsymbol{k}s} \ll k_{\rm B}T$ for all $\boldsymbol{k}$ and s. In this situation, where the temperature is well above the Debye temperature, $F_{\rm harm}$ becomes equal to

$$F_{\rm harm} = k_{\rm B}T\left[\sum_{\boldsymbol{k}s}\left(\ln(\hbar\omega_{\boldsymbol{k}s}/k_{\rm B}T) + \frac{1}{24}\,(\hbar\omega_{\boldsymbol{k}s}/k_{\rm B}T)^2 + \cdots\right)\right]. \tag{26}$$

For $\hbar\omega_{\boldsymbol{k}s} \ll k_{\rm B}T$, which is the case for most planetary science problems, the higher terms are negligible, and we have the simple form

$$F_{\rm harm} = N_{\boldsymbol{k}s}k_{\rm B}T\,\ln(\hbar\bar{\omega}/k_{\rm B}T), \tag{27}$$

where $N_{\boldsymbol{k}s}$ is the total number of vibrational modes in the crystal (three times the number of atoms in the primitive cell times the number of primitive cells) and $\bar{\omega}$ is the geometric mean vibrational frequency, defined by

$$\frac{1}{N_{\boldsymbol{k}s}}\sum_{\boldsymbol{k}s}\ln(\omega_{\boldsymbol{k}s}/\bar{\omega}) = 1. \tag{28}$$

We now turn to the theory needed to calculate the phonon frequencies $\omega_{\boldsymbol{k}s}$ of a perfect crystal by first-principles methods, starting with a brief summary of the theory of lattice vibrations.

5.2. Basic theory of lattice vibrations

The quasiharmonic theory of lattice vibrations (see, e.g. [51]) is based on the assumption that the vibrations are very small, so that the total energy function is accurately represented by its expansion to second order in these displacements. Let us suppose that there are ν atoms in the primitive unit cell of the perfect crystal, and let $\boldsymbol{R}_{ls}$ be the equilibrium position of the sth atom ($s = 1, 2, \ldots \nu$) in the lth primitive cell. If $\boldsymbol{u}_{ls}$ denotes the displacement away from its equilibrium position of the sth atom in the lth unit cell then the total energy U can be expanded thus:

$$U = E_{\rm perf} + \frac{1}{2}\sum_{ls\alpha,l't\beta}\Phi_{ls\alpha,l't\beta}u_{ls\alpha}u_{l't\beta} + \cdots, \tag{29}$$

where $E_{\rm perf}$, as before, is the total energy of the static perfect crystal and α, β indicate Cartesian components. (As usual, U and $E_{\rm perf}$ represent *free* energies if the finite-temperature version of DFT is used.) The matrix whose elements are $\Phi_{ls\alpha,l't\beta}$ is called the force-constant matrix and is given by the double derivative $\partial^2U/\partial u_{ls\alpha}\partial u_{l't\beta}$ evaluated with all atoms at their equilibrium positions. This force-constant matrix gives the relation between the forces $\boldsymbol{F}_{ls}$ and the displacements $u_{l't}$, as can be seen by differentiating equation (29) and ignoring the higher-order anharmonic terms:

$$F_{ls\alpha} = -\partial U/\partial u_{ls\alpha} = -\sum_{l't\beta}\Phi_{ls\alpha,l't\beta}u_{l't\beta}. \tag{30}$$

The equation of motion describing the harmonic lattice vibrations is then

$$m_s \partial^2 u_{ls\alpha}/\partial t^2 = F_{ls\alpha} = -\sum_{l't\beta} \Phi_{ls\alpha,l't\beta} u_{l't\beta}, \tag{31}$$

with m_s the mass of the sth atom in each cell. The normal mode solutions to this equation have the form:

$$u_{ls\alpha}(t) = u^0_{ls\alpha} \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{R}_{ls}} \mathrm{e}^{\mathrm{i}\omega t}, \tag{32}$$

where $\boldsymbol{k}$ is the wavevector of the mode, ω is its frequency and the time-independent constants $u^0_{ls\alpha}$ specify the pattern of displacements characterizing the vibrational mode. Substituting this form into equation (31), we have

$$-m_s \omega^2 u^0_{s\alpha} \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{R}_{ls}} = -\sum_{l't\beta} \Phi_{ls\alpha,l't\beta} u^0_{t\beta} \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{R}_{l't}}. \tag{33}$$

This equation can be simplified by multiplying through by $\exp(-\boldsymbol{k}\cdot\boldsymbol{R}_{ls})$ and defining the so-called dynamical matrix $D_{s\alpha,t\beta}(\boldsymbol{k})$:

$$D_{s\alpha,t\beta}(\boldsymbol{k}) = (m_s m_t)^{-1/2} \sum_{l'} \Phi_{ls\alpha,l't\beta} \mathrm{e}^{-\boldsymbol{k}\cdot(\boldsymbol{R}_{ls}-\boldsymbol{R}_{l't})}. \tag{34}$$

In terms of this, we have

$$\sum_{t\beta} D_{s\alpha,t\beta}(\boldsymbol{k}) v^0_{t\beta} = \omega^2 v^0_{s\alpha}, \tag{35}$$

where $v^0_{s\alpha} = m_s^{1/2} u^0_{s\alpha}$. This shows that every vibrational mode is associated with a wavevector $\boldsymbol{k}$, and that the squared frequencies ω^2 at this wavevector are the eigenvalues of the dynamical matrix $D_{s\alpha,t\beta}(\boldsymbol{k})$. The dimension of this matrix is three times the number of atoms in the primitive cell, and this is the number of distinct vibrational modes at each wavevector. The graphs of these frequencies $\omega_s(\boldsymbol{k})$ against the wavevector are often referred to as phonon dispersion relations.

Examples of phonon dispersion relations for geological materials will be given later in this section. But before this, we want to indicate how the phonon frequencies can be obtained from first-principles DFT calculations.

5.3. First-principles calculation of phonon frequencies

In fact, there are two rather different first-principles strategies for calculating phonon frequencies. The method that is easier to understand starts from the fact that the force-constant matrix expresses the proportionality between displacements and forces, when the displacements are small enough for this relationship to be linear. Since the forces on the atoms are calculated in any case in first-principles codes, all that has to be done in principle is to displace a single atom t in cell l' in the Cartesian direction β, all other atoms being held fixed at their equilibrium positions; the forces $F_{ls\alpha}$ on all the atoms then give directly the elements of the force-constant matrix $\Phi_{ls\alpha,l't\beta}$ for the given $(l't\beta)$. If this procedure is repeated for all other $(l't\beta)$, all the elements of the force-constant matrix can be obtained. Note that the number of separate calculations required to do this is at most three times the number of atoms in the primitive cell, but for most materials symmetry relations can be used to reduce this number substantially. This strategy, sometimes called the 'small displacement method' [155], is implemented in the PHON code written by one of the present authors [4]. Although the small displacement method is widely used, and can be very accurate, a word of caution is in order. Since DFT calculations on condensed matter always use periodic boundary conditions, the repeating cell must be large enough so that the elements $\Phi_{ls\alpha,l't\beta}$ have all fallen off to negligible values at the boundary of the repeating cell. This is readily achieved for some materials, particularly metals.

In ionic materials, the force-constant elements fall off only as r^{-3}, and convergence can be slow. Moreover, in polar materials Coulomb forces produce a macroscopic electric field in the limit of zero wavevector. This electric field is responsible for a splitting in the frequencies of the vibrational modes parallel and perpendicular to the electric field (the so-called LO–TO splitting). This effect can be taken into account by adding a nonanalytic contribution to the dynamical matrix at wavevector $\boldsymbol{k}$ which has the form [105]

$$D^{\mathrm{na}}_{s\alpha,t\beta} = (m_s m_t)^{-1/2} \frac{4\pi e^2}{\Omega} \frac{(\boldsymbol{k}\cdot\boldsymbol{Z}^*_s)_\alpha(\boldsymbol{k}\cdot\boldsymbol{Z}^*_t)_\beta}{\boldsymbol{k}\cdot\epsilon^\infty\cdot k}, \tag{36}$$

where $\boldsymbol{Z}^*_s$ is the Born effective charge tensor for atom s and ϵ^∞ the high frequency static dielectric tensor. These two quantities can be calculated in the framework of density functional perturbation theory (DFPT) [37, 38, 105], which also provides a second elegant strategy for the calculation of phonons in crystals. The main idea in DFPT, pioneered by Baroni *et al* [38] (see also [111, 113]) is to exploit the Hellmann–Feynman theorem to show that a linear order variation in the electron density upon application of a perturbation to the crystal is responsible for a variation in the energy up to second (in fact, third [114]) order of the perturbation. Using standard perturbation theory, this linear order variation of the electronic charge density can be calculated using only unperturbed wavefunctions, which therefore only requires calculations on the ground state crystal. If the perturbation is a phonon wave with wavevector $\boldsymbol{k}$, calculation of the density change to linear order in the perturbation can be used to determine the force-constant matrix at wavevector $\boldsymbol{k}$. This can be done for any arbitrary wavevector, without the need for the construction of a supercell. The implementation of the method is by no means straightforward, and for further details the reader should consult the original papers [38, 105].

5.4. Phonon dispersion relations at ambient pressure

It was shown many years ago that the LDA gives excellent results for the phonon dispersion relations of MgO [80]. We illustrate this from the recent work of Oganov *et al* on the phase diagram of MgO [203] (figure 7). The agreement to within ∼3% for all phonon branches across the whole Brillouin zone implies that thermodynamic predictions from quasi-harmonic theory are likely to be reliable. As a rather different illustration, we show in figure 8 the phonon curves of ferromagnetic Fe from GGA calculations [18] compared with experimental results. Again, the agreement is very satisfactory.

5.5. Phonons at high pressure

Measurements of the phonon density of states (DOS) of hcp iron up to Earth's core pressures have recently been reported by Mao *et al* [169]. They used the technique of nuclear resonant inelastic x-ray scattering (NRIXS) [237,256], which exploits the presence of a nuclear excitation in ^{57}Fe at 14.4136 keV. The large cross section for incident photons of this nuclear mode and its convenient lifetime (141 fs) make it suitable for the study of the vibrational properties of this iron isotope. The measured DOSs were also compared with those obtained from first-principles calculations, using the small displacement method described above [169] with forces obtained from DFT-GGA calculations. The good agreement between theory and experiment (figure 9) provides valuable confirmation that first-principles thermodynamic predictions are likely to be accurate at high pressures.

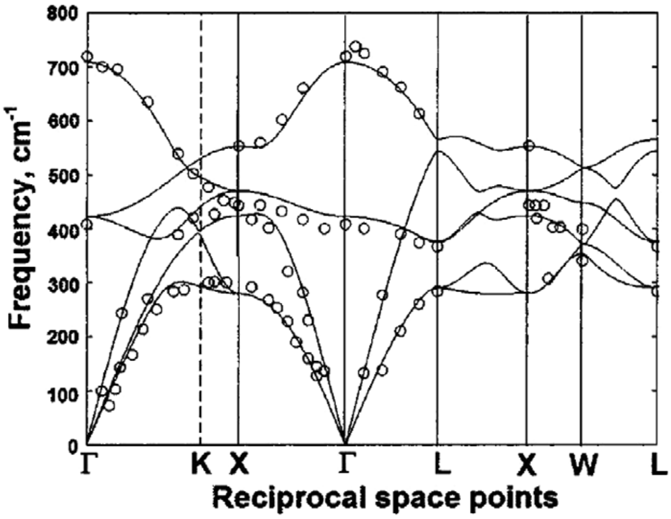

Figure 7. Phonon dispersion relations of MgO in the rock-salt structure at ambient pressure. Lines and open circles show first-principles theory and experiment, respectively. Reproduced with permission from [202]. Copyright 2003 American Institute of Physics.

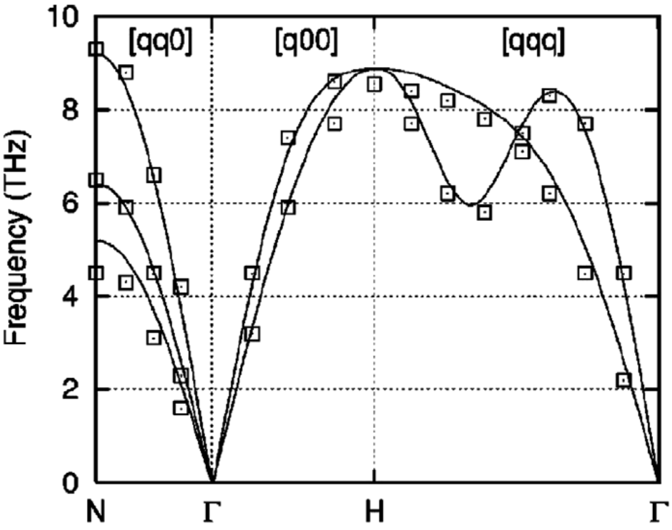

Figure 8. Phonon dispersion relations of ferromagnetic bcc Fe. Lines and open squares show first-principles theory and experiment, respectively. Reproduced with permission from [18]. Copyright 2000 American Physical Society.

The NRIXS technique has also been recently used to measure the partial density of states of FeS as a function of pressure [148]. Measurements were taken at pressures 1.5, 4.0 and 9.5 GPa, in the troilite, MnP-type and monoclinic crystal structures of FeS, respectively, and were compared with first-principles calculations based on DFT-GGA. The agreement beteween the calculations and the experiments was reasonably good, although the FeS in the troilite structure was found to be unstable. The calculations were also used to provide the total density of states, which provided thermodynamic quantities such as the entropy and the specific heat (see below).

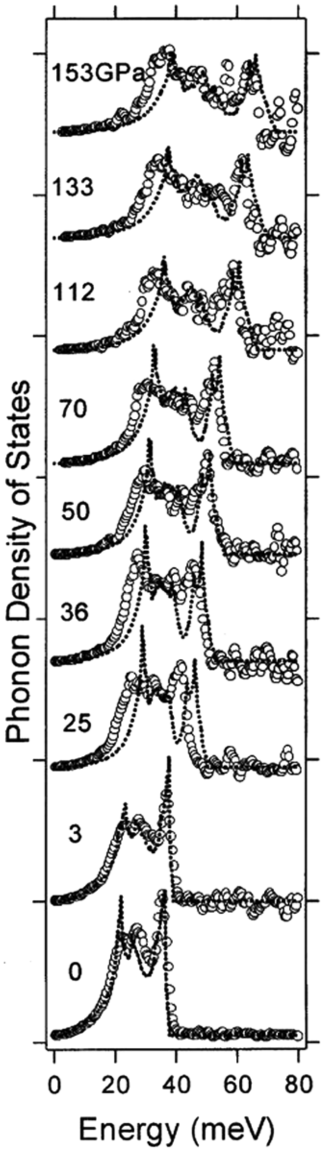

Figure 9. Phonon density of states of bcc Fe (pressure $p = 0$ and 3 GPa) and hcp Fe (p from 25 to 153 GPa). Dotted curves and open circles show first-principles theory and experiment, respectively. Reproduced with permission from [169]. Copyright 2001 American Association for the Advancement of Science.

5.6. Thermodynamics in the quasi-harmonic approximation

An important practical use of first-principles phonon frequencies is the calculation of quasi-harmonic free energies and hence the temperature-dependent phase boundaries between coexisting crystal structures. We noted in section 4.1.1 that the transition in MgO from the rock-salt structure to the CsCl structure occurs at $\sim$ 500 GPa at $T = 0$. To confirm that MgO cannot exist in the CsCl structure in the Earth's lower mantle ($p < 135$ GPa), quasi-harmonic calculations of $p(T)$ on the phase boundary have been reported. Initially [86], it

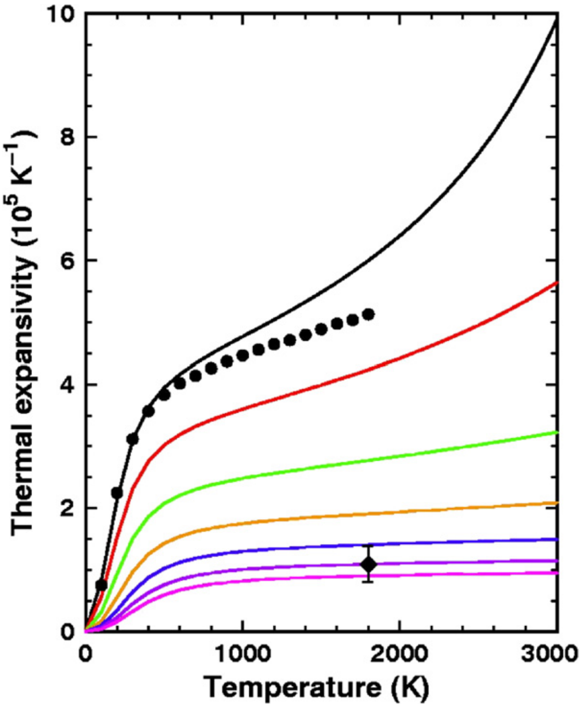

Figure 10. Temperature dependence of thermal expansivity α of MgO on isobars at 0, 10, 30, 60, 100, 150 and 200 GPa (curves from top to bottom). Experimental data at zero pressure are indicated by filled circles. The average value of α over temperature range 300–3300 K and pressure range 169–196 GPa derived from shock experiments is indicated by the diamond. Reproduced with permission from [140]. Copyright 1999 American Association for the Advancement of Science.

seemed that $\mathrm{d}p/\mathrm{d}T$ might have the very large value of -50 GPa per 1000 K, but more recent calculations [203], in which the linear response is used to calculate the phonon frequencies, indicate that, even by 5000 K, $p(T)$ has fallen below its $T = 0$ value by no more than ~50 GPa. First-principles quasi-harmonic calculations have also been used to study the high-T/high-p phase diagrams of SiO_2 [204] and Al_2O_3 [208].

As an illustration of thermodynamic calculations which demonstrate the accuracy of the first-principles quasi-harmonic approach, we show in figure 10 the temperature dependence of the thermal expansivity α of MgO at pressures up to 200 GPa, compared with experimental results. At ambient pressure, the first-principles $\alpha(T)$ agrees very closely with experiment up to ~1000 K (about one-third of the melting temperature), as would be expected from the good phonon dispersion relations (see figure 7). Good agreement is also found with the isolated shock datum at $p \sim 180$ GPa. (The increasingly poor agreement at higher T was shown by Oganov and Dorogokupets [202] to be due to failure of the quasiharmonic approximation; it was shown there that inclusion of anharmonic corrections restores agreement with the experiment.)

We described in section 4.1.2 the important role of first-principles calculations in the discovery of the post-perovskite structure of $MgSiO_3$, which appears to solve the mystery about the D$''$ zone at the base of the Earth's mantle. At low temperatures, the pressure of the transition from perovskite to post-perovskite is in the region of 90 GPa, which is well below the pressure of 127 GPa at the top of the D$''$ zone. However, quasi-harmonic first-principles calculations [207, 261] show that the transition pressure increases with temperature, the Clapeyron slope $\mathrm{d}p/\mathrm{d}T$ being between 7 and 10 MPa K^{-1}. This brings the predicted transition pressure close to that at the top of D″. These values of the Clapeyron slope also explain semi-quantitatively the variation in thickness of the D″ zone.

We mentioned earlier the controversies surrounding the phase diagram of solid Fe at conditions relevant to the Earth's inner core. Harmonic free-energy calculations on the candidate crystal structures of Fe were reported several years ago [267]. Since the bcc structure becomes mechanically unstable above $\sim$ 180 GPa, and some of the other structures also become unstable before inner-core pressures are reached, the only surviving candidate structures for quasi-harmonic calculations are hcp, fcc and double-hcp. The free energies of these three structures lead to the conclusion that hcp is the most stable structure, as has generally been assumed. However, we caution that this is not conclusive evidence of Fe in the inner core having the hcp structure, first because anharmonic effects are neglected and second because the presence of light elements such as S and Si might tip the balance. Harmonic DFT calculations have also played a role in helping to resolve a controversy over the temperature dependence of c/a of hcp Fe under inner-core conditions. There have been suggestions [247] that the c/a ratio increases rapidly with temperature from its $T = 0$ value of 1.59, attaining a value of nearly 1.70 in the inner core. This, if true, would have important consequences for the elastic properties of Fe in the inner core. However, a re-examination of the question [104, 103], using methods including quasi-harmonic DFT theory, indicates that the variation of c/a is actually very weak, in agreement with recent experimental results [167].

There has been a considerable amount of work on the high-temperature elastic constants of Earth's mantle materials using first-principles quasi-harmonic theory. Examples include work by Karki *et al* on MgO [140] and by Karki and Wentzcovitch [139] on $\mathrm{MgSiO_3}$ ilmenite. Of particular interest is the work of Wentzcovitch *et al* [281] on the high-T elastic properties of $\mathrm{MgSiO_3}$ perovskite, which is relevant to the question of whether the chemical composition of the lower mantle varies with depth. Their predictions for the relationship between the longitudinal and transverse velocities and the density led them to conclude that the seismic data is best understood by postulating that composition does depend on depth.

6. Molecular dynamics

The methods of quasi-harmonic lattice dynamics discussed in the previous section are powerful within their domain of applicability, but for high-temperature solids near the melting temperature and for liquids, we must turn to molecular dynamics (m.d.) simulation. The m.d. technique has been in use for over 40 years [26, 101, 224]. Before the advent of first-principles m.d. in 1985 [58], all m.d. simulations were based on empirical models for the total energy function, and these models remain extremely important today [94]. Since this review is about first-principles modelling, we shall focus here on m.d. simulations based on DFT, which were pioneered by Car and Parrinello [58]. Roughly speaking, first-principles m.d. is the same as empirical m.d., but with the difference that, as the system evolves in time, the energy and the atomic forces for the atomic positions at any instant are calculated using the DFT techniques outlined in section 3. Because of this, our summary of the m.d. techniques themselves (section 6.1) will be very brief, concentrating mainly on the DFT aspects. We then review a range of applications to problems relevant to the Earth and planetary sciences, including the properties of hydrogen and hydrogen–helium mixtures under the conditions that exist in the planets Jupiter and Saturn, the properties of water, ammonia and methane under the conditions in Uranus and Neptune, the thermodynamic and transport properties of iron in the Earth's core and the properties of materials in the Earth's mantle, particularly magnesium silicate $\mathrm{MgSiO_3}$.

At the end of the section, we briefly review first-principles statistical–mechanical methods for determining crystal structures.

6.1. Summary of first-principles molecular dynamics

The two basic assumptions of first-principles m.d. are that the positions and momenta of the nuclei evolve in time according to Newton's equations of motion and that the forces on the nuclei at any instant of time are those produced by the electron sub-system in its ground state, calculated as though the nuclei were completely static at their current instantaneous positions. The first assumption is accurate provided the temperature of the system is well above the Debye temperature (see section 5.1), which is usually the case for problems in the Earth and planetary sciences. The assumption about the forces is the well-known Born–Oppenheimer approximation, which says that the electrons follow the motions of the nuclei adiabatically. This assumption is also usually accurate and is justified by the fact that the frequencies associated with the nuclear motions (typically $\sim 10^{13}$ Hz) are very much lower than those of electronic excitations (typically $\sim 10^{15}$ Hz), so that the nuclear motions do not significantly excite the electrons from their ground state.

With these basic assumptions, the simplest type of m.d. simulation that can be performed is that in which the total energy is conserved, with 'total energy' here meaning the sum of the kinetic energy of the nuclei and the ground-state DFT energy $U_{\mathrm{tot}}(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N)$ at each instant of time. In this sense, the energy U_{tot} plays the role of the 'potential energy' in an empirical m.d. simulation. This type of constant energy m.d., with the volume of the system also held fixed, usually called (N, V, E) m.d., generates phase-space configurations with a probability distribution characteristic of the microcanonical ensemble, provided the time evolution is ergodic. This fact implies that the time averages of dynamical variables are equal to their thermal averages in the microcanonical ensemble.

The usual way of generating trajectories in (N, V, E) m.d. is to divide time into time steps Δt and to approximate the equation of motion $m_i \partial \boldsymbol{v}_i / \partial t = \boldsymbol{F}_i$ (m_i is the mass of atom i, $\boldsymbol{v}_i$ is its velocity, and $\boldsymbol{F}_i$ is the force acting on it) by a finite-difference equation. The simplest and most effective of such algorithms is the Verlet equation [26, 101, 265], which can be expressed in a number of equivalent forms, one of which is

$$\begin{aligned} \boldsymbol{r}_i(t + \Delta t) &= \boldsymbol{r}_i(t) + \boldsymbol{v}_i(t)\Delta t + \frac{1}{2m_i}\boldsymbol{F}_i(t)\Delta t^2, \\ \boldsymbol{v}_i(t + \Delta t) &= \boldsymbol{v}_i(t) + \frac{1}{2m_i}(\boldsymbol{F}_i(t) + \boldsymbol{F}_i(t + \Delta t))\Delta t. \end{aligned} \tag{37}$$

With an appropriately chosen time step, of order 1/20 of typical vibration periods, this algorithm conserves energy rather well, over both short and long time scales. The algorithm also has the highly desirable properties of preserving the time reversibility of the equations of motion and of reproducing the incompressibility of flow in phase space required by Liouville's theorem. (See, e.g. [101] for a detailed explanation of these important statements.)

A key requirement in first-principles m.d. is that the electrons remain in the self-consistent ground state, or at least very close to it, as the nuclear positions evolve. There are two ways of achieving this. The most direct way is to relax the electrons to the ground state, to within a specified tolerance, at each time step [153]. The other way, introduced by Car and Parrinello [58], is to treat the wavefunction coefficients (in practice, the plane-wave coefficients $c_{n\boldsymbol{k}}$ of equation (17)) as fictitious dynamical variables. If these variables are given small enough (fictitious) masses then they follow the nuclear motions adiabatically, at least for non-metallic systems.

The primary thermodynamic quantities that can be computed in an (N, V, E) m.d. simulation, apart from the energy itself, are the temperature T and the pressure p (more generally the stress tensor $\sigma_{\alpha\beta}$). The temperature is deduced from the average kinetic energy of the nuclei, using the fact that the thermal average kinetic energy per atom is $(\frac{3}{2})k_{\mathrm{B}}T$. The pressure or stress is obtained from the time average of the DFT stress tensor (see section 3.7), also including a contribution from the kinetic energy of the nuclei. Also normally calculated in the course of an m.d. simulation are the radial distribution functions $g_{\alpha\beta}(r)$ giving the probability of finding an atom of type β at distance r from an atom of type α; these provide the simplest way of characterizing the average structure of a liquid and are also rather directly related to quantities measured in diffraction experiments. Many other thermodynamic, structural and dynamical quantities can also be calculated by taking time averages of appropriate variables, and a number of these quantities will be mentioned when we present illustrative results later in this section.

Like the m.d. simulation based on empirical interaction models, first-principles m.d. can also be performed at constant temperature and/or constant pressure (more generally, constant stress). The preferred methods for doing this are those that produce phase-space sampling corresponding to well defined statistical–mechanical ensembles. For constant-T simulation, the thermostats having this property include those due to Andersen [28] and Nosé and Hoover [127, 196]. When combined with the Parrinello–Rahman constant-stress technique [214], this also allows simulations to be performed at constant T and $\sigma_{\alpha\beta}$ (see e.g. [283]).

6.2. Hydrogen in Jupiter and Saturn

Our presentation of recent first-principles simulation work on hydrogen will focus mainly on studies that help with the planetary issues outlined in section 2. Work not reviewed here includes the many interesting studies on the crystal structures of solid hydrogen [146, 150, 151]. A considerable effort has been devoted to the question of whether first-principles simulations correctly reproduce the thermodynamic properties and the electrical conductivity of hydrogen in the region explored by single- and multiple-shock experiments. It is now clear that simulations by different groups [74, 102, 126, 161] agree closely with each other and also reproduce reasonably well the experimental pressure–density relation on the single-shock Hugoniot [125, 147, 189], if one discounts the results of laser-shock experiments [73], which appear to be out of line with other data at pressures above $\sim$ 30 GPa. The first-principles calculations of electrical conductivity have been done using the so-called Kubo–Greenwood approximation [116, 157], which ignores electron–electron correlations, and can be expected to give only rather qualitative results, except for nearly free electron metals. Nevertheless, the simulations [102] agree with experiment [280] in predicting that on the single-shock Hugoniot H becomes a poor metal, with a conductivity in the region of 5000 Ω^{-1} cm^{-1} when the compression factor and temperature exceed 5 and 10 000 K, respectively.

A particularly comprehensive study of the structure, band gap and electrical conductivity has been reported recently by Hood and Galli [126]. Recognizing that the band gap calculated by DFT may not be trustworthy, these authors have used the much more reliable quantum Monte Carlo technique [100] to calculate the band gap in the simulated system. The simulations were performed at 3000 K and compression factors of 4, 6, 9 and 14. (The region probed by reverberating shock experiments [280] passes near the state $T \simeq$ 3000 K and a compression factor of 9.) The simulations show that for compressions greater than $\sim$4, the lifetime of molecules starts to approach the very short value of 10 fs, which is the normal vibrational period of the molecule, and the radial distribution function $g(r)$ starts to lose the sharp peak associated with the intra-molecular bond length. At 9-fold compression, where the reverberating shock experiments show a poor metal, the calculated band gap goes to zero, and the system is ∼40% dissociated. At this state, the conductivity given by the Kubo–Greenwood approximation is about four times the value deduced experimentally, which for present purposes counts as reasonable agreement. An important conclusion from this simulation work is that in this part of the phase diagram metallization and molecular dissociation are intimately linked.

The simulations just outlined give no evidence that the transition from the molecular insulating state to the dissociated metallic state occurs by a first-order phase transition. Nevertheless, theoretical arguments supported by first-principles data suggest that such a first-order transition should exist [145]. At lower temperatures, near the melting curve, first-principles simulations convincingly indicate the presence of such a transition. The evidence comes from work by Scandolo [233], in which the properties of fluid H were studied at 1500 K in the pressure range 75–175 GPa. Along this isotherm, a discontinuity in the volume of ∼ 6% was found at ∼125 GPa. The almost complete disappearance of the intra-molecular peak in $g(r)$ across this transition indicates that it is associated with major molecular dissociation. The rapid loss of the band gap in the electronic density of states as p goes from 112 to 137 GPa suggests that metallization occurs across the transition. More recent first-principles work on the melting curve of H up to pressures of 200 GPa [50], further reviewed in section 8.4 below, indicates that the liquid in contact with the solid remains molecular up to this high pressure but that there may well be a triple point between solid, insulating liquid and metallic liquid at a pressure of ∼300–400 GPa. These two simulation studies confirm that the insulator–metal transition pressure is indeed strongly temperature dependent. The transition line presumably ends in a critical point, but the location of this point remains unknown.

To conclude this summary of first-principles simulations on hydrogen, we note briefly the work of Pfaffenzeller *et al* on H/He mixtures [218]. This work addresses the question of the anomalous heat release from Saturn, and why this should be so much higher than that of Jupiter. It has been speculated that this could be because He is immiscible with H for conditions in some regions inside Saturn, so that the heat release comes from sinking of precipitated He, whereas at the higher temperatures inside Jupiter, the H/He mixture is always stable. Their simulations focus first on the mixture at low temperatures, where they show that there is a substantial positive enthalpy of mixing, so that phase separation would always occur. Their estimates of the free energy of mixing at high temperatures allow them to estimate the phase separation boundary, which indeed confirms the proposed reason for heat release in Saturn, and also supports the idea that the H/He mixture will lie at temperatures in the miscible region above the boundary everywhere inside Jupiter.

6.3. *Uranus and Neptune*

Our survey of scientific issues concerning Uranus and Neptune in section 2 showed the need to understand the properties of water, ammonia and methane at high pressures and temperatures. A first-principles study with strong links to experiment, which gives important insight into pressure-induced changes in the structure of ice is that of Bernasconi *et al* [47]. This work focuses on comparisons with experimental infrared spectra and the interpretation of the observed spectral changes in terms of structural changes. The simulations indicate that at pressures in the region of 40 GPa the protons have an asymmetric distribution, each proton being bound for long periods to one oxygen; at 80 GPa, rapid hopping of protons is observed, while at 125 GPa the spatial distribution of protons consists of peaks on the O–O mid-points, in the way predicted for the ice-X structure. Several other first-principles m.d. studies by the same group [45, 46] provide further information.

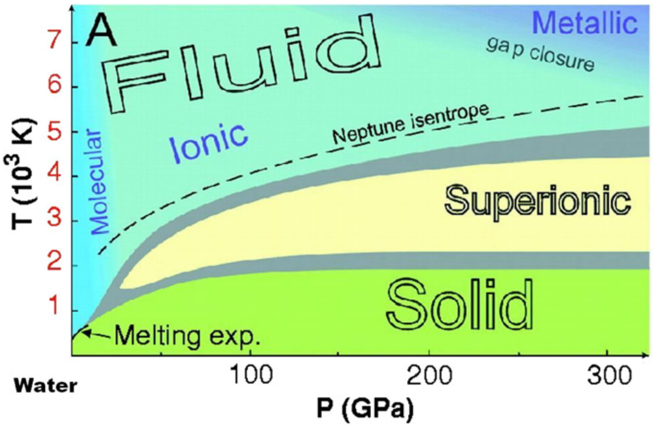

Figure 11. Phase diagram of water under conditions relevant to the interior of Uranus and Neptune. Reprinted with permission from [61]. Copyright 1999 American Association for the Advancement of Science.

Rather extensive first-principles m.d. work on both H_2O and NH_3 under extreme conditions by Cavazzoni *et al* [61] gives a valuable overview of the phase diagram, covering the pressure range 30–300 GPa and the temperature range 300–7000 K (see figure 11). In the lower part of the temperature range, both H_2O and NH_3 are solid. In the case of H_2O, the system is in the ice-VII structure at 30 and 60 GPa and in the ice-X structure at 150 and 300 GPa. In both structures, the O ions form a bcc lattice, the protons being on the O–O bond centres in ice-X and on off-centre positions in ice-VII. The simulations show that solid H_2O becomes a superionic conductor at temperatures above ~2000 K, with a very high proton diffusion coefficient of order $6 \times 10^{-8}\ \mathrm{m^2\ s^{-1}}$. This confirmed earlier first-principles work on superionic conduction in high-pressure ice [45].

A very recent study [109] combining Raman spectroscopy in a laser-heated diamond-anvil cell and first-principles m.d. simulation sheds further light on the dissociation of water at high p and T. The experiments confirm the solid–solid–liquid triple point, which is determined to be at 47 GPa and 1000 K, considerably lower in temperature than earlier measurements by Schwager *et al* [235]. At 30 GPa and 1200 K, just on the lower boundary of the superionic regime according to the first-principles work of Cavazzoni *et al* [61], they find that molecular H_2O is the dominant species. Their first-principles simulations indicate that superionic conduction sets in at a pressure of ~ 43 GPa, very close to that of the triple point.

The fate of CH_4 in the interiors of Uranus and Neptune is outstandingly important for a very interesting reason. If the temperatures and pressures are high enough to induce dissociation into carbon and hydrogen, then the C and H are expected to be in the form of neutral atoms, rather than ions, because the electronegativities of H and C are similar. This means that the C might precipitate out as diamond. The notion that Uranus and Neptune might contain large amounts of diamond first emerged [226] from an analysis of shock experiments on CH_4, in which it was believed that dissociation had been observed. The importance of this notion is that the sinking of precipitated diamond towards the centre of the planet would release gravitational energy, which could help to explain the observed luminosity of Neptune.

First-principles m.d. simulations performed by Ancilotto *et al* [27] suggest that the situation may be more complex. This work was done nearly 10 years ago, and at that time it was possible to treat only a rather small system of 16 CH_4 molecules. Nevertheless, qualitatively correct results would be expected even for such a small system. The main simulations were done at the two states 100 GPa, 4000 K and 300 GPa, 5000 K, which were chosen to lie roughly on the planetary isentropes, the first state corresponding to a distance of about two-thirds of a planetary radius from the centre. The simulations showed that the CH_4 molecules initially present undergo very rapid breakdown, within ∼1 ps, but surprisingly this breakdown does not yield carbon and hydrogen at either of the two states. The simulation at 100 GPa and 4000 K revealed the formation of ethane (C_2H_6) and butane (C_3H_8) molecules, the hydrogen produced in this process being mainly in molecular form. Analysis of the energetics of the system showed that under these conditions, the formation of higher hydrocarbons, rather than diamond, is thermodynamically favoured. At 300 GPa and 5000 K, longer alkane chains were formed, but there was still no evidence for the formation of diamond. The authors proposed that diamond precipitation might still occur in the deep interior of Uranus and Neptune at pressures beyond 300 GPa, but to the best of our knowledge direct evidence for this from first-principles simulation is still lacking. It is worth noting that at both the thermodynamic states examined there is a large electronic band gap of ∼6 eV, so the C/H system certainly cannot contribute to the electrical conductivity of the planets. Experimental evidence for the formation of polymeric hydrocarbons has appeared since the simulations were reported [44].

6.4. *Iron in the Earth's inner core*

We noted earlier some of the questions concerning the Earth's inner core, including the estimation of the impurity content, the question of whether the crystal structure is hcp, as has generally been assumed, and the correct explanation of the observed seismic anisotropy. First-principles m.d. has already started to give partial answers to these questions, though much remains to be done.

The simulations of Laio *et al* [159] provide useful information about the elastic properties of hcp Fe under inner-core conditions. (These simulations are based on an embedded-atom model fitted to high-temperature first-principles data but are expected to be reliable for high-temperature elastic properties; this is confirmed by their good agreement with full first-principles and DAC results for low-temperature elastic constants.) An important finding from their work is that the shear modulus decreases very rapidly with increasing temperature at the ICB pressure of 330 GPa. The calculated value at temperatures just below T_m is ∼70% lower than the $T = 0$ K value, and this strong reduction is consistent with the very high value of Poisson's ratio in the core of ∼0.44 deduced from seismic measurements [89].

We mentioned in section 5.6 the controversy over the temperature variation of c/a in hcp Fe under inner-core conditions. The indication from quasi-harmonic DFT calculations [103] and from recent experiments [167] that c/a under these conditions differs little from its $T = 0$ value is confirmed by first-principles DFT m.d. simulations [103], in which the thermal average deviatoric stress is calculated as a function of c/a at inner-core temperatures.

The question of the most stable crystal structure of Fe at inner-core p and T appeared to have been settled in favour of hcp by the quasi-harmonic free-energy calculations summarized in section 5.6. Indeed, the bcc structure seemed to be ruled out by three separate arguments: (i) the elastic instability of bcc at pressures above 150 GPa; (ii) the vibrational instability also revealed by first-principles calculations at high pressure; (iii) the fact that the enthalpy of bcc is considerably higher than that of hcp. However, these arguments are not conclusive, because they are based on zero-temperature or lattice-dynamical first-principles calculations, which cannot be used to determine entropic effects in highly anharmonic solids at high temperatures. Remarkably, recent first-principles m.d. simulations [269] on bcc Fe have shown that entropy effects remove the low-temperature instabilities. (See also closely related work based on an

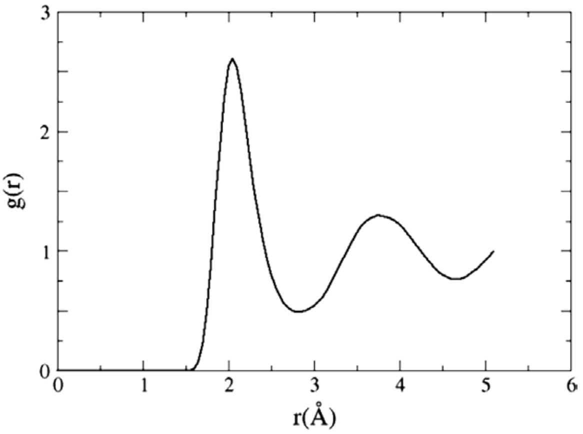

Figure 12. Radial distribution function of liquid Fe from first-principles simulation at pressure of 330 GPa and temperature of 6000 K.

embedded-atom model parametrized using first-principles calculations [43].) The entropic stabilization was shown by three different kinds of analysis: the anisotropic stresses were monitored for evidence of elastic instability; the atomic positions relative to those of a perfect bcc structure were analysed for evidence of vibrational instability; and the structure factors were computed to help identify the crystal phases present. As we shall note in section 7, calculations of the relative free energy of hcp and bcc structures indicate that hcp is still marginally more stable than bcc, but the difference is small enough so that the balance could be tipped by as little as 5 mole% of Si and/or S in the inner core. First-principles calculations of chemical potentials in the inner core (section 9) suggest that at least this level of impurity content is highly probable.

6.5. Liquid iron at Earth's core conditions

Since the Earth's liquid outer core is believed to consist mainly of iron, the properties of liquid iron (we abbreviate to *l*-Fe) under relevant conditions ($135 < p < 330$ GPa and $4000 < T < 6000$ K) are of fundamental importance to an understanding of the Earth's interior. Direct experiments on the properties of *l*-Fe under these extreme conditions are not possible at present. However, the excellent agreement of DFT predictions with experimental data for the low-temperature equation of state of Fe up to pressures of 300 GPa (section 4.1.4) and for the phonon density of states up to 150 GPa (section 5.5) provides grounds to believe that first-principles m.d. should be capable of giving an accurate picture of *l*-Fe under these conditions.

Direct first-principles m.d. simulations of *l*-Fe have been reported over a wide range of thermodynamic conditions [13, 18, 23, 82], and there have also been simulations based on empirical potentials fitted to DFT data under the relevant conditions [42, 159]. Even with direct first-principles m.d., it is possible at present to simulate systems of up to several hundred atoms, and careful studies of size errors [17] show that this is more than large enough to make size errors very small for most quantities.

To illustrate the predicted structure of *l*-Fe under core conditions, we show the r.d.f. at 330 GPa and 6000 K (figure 12). This $g(r)$ is extremely similar to those found in simple liquids in which the structure is dominated by short-range interatomic repulsion, as might be expected for a liquid under very high pressure. This similarity suggests that the properties of the liquid might be accurately modelled by a system of atoms interacting *via* a short-range repulsive pair potential. It has been shown that this idea can be used to provide an extremely simple description of *l*-Fe over the whole range of conditions of the Earth's core, as will be outlined later in this section.

The m.d. simulations of *l*-Fe also yield thermodynamic information. The most straightforward quantities to calculate are the internal energy and pressure for a given volume and temperature. An alternative and more general approach is to calculate all thermodynamic quantities from the free energy, as we describe in section 7.

A property of great interest for understanding the core is the self-diffusion coefficient D of the atoms in *l*-Fe. There has been enormous controversy in the past about the viscosity of the outer core, with estimates ranging over 12 orders of magnitude [219]. The high estimates, if correct, would have important implications for understanding the generation of the Earth's magnetic field. A very high viscosity would imply a very low diffusion coefficient D, so a simple way of addressing the controversy is to calculate D. The standard technique for doing this in m.d. [26, 101] is to compute the mean square time dependent displacement $f(t) = \langle |\boldsymbol{r}_i(t+t_0) - \boldsymbol{r}_i(t_0)|^2 \rangle$ for individual atoms, i.e. the mean square distance travelled by an atom in time t. In simple liquids, $f(t)$ increases linearly in t once t is greater than a few tenths of a picosecond, and the slope of $f(t)$ in this regime is $6D$. First-principles m.d. calculations of D at core conditions [18, 82] give values in the region of $5 \times 10^{-9}\ \mathrm{m^2\,s^{-1}}$. To put this in context, the self-diffusion coefficient of water at ambient pressure and room temperature is $\sim 2 \times 10^{-9}\ \mathrm{m^2\,s^{-1}}$ [108]. The important conclusion is that, in spite of the enormous pressure in the core, Fe atoms diffuse at a rate that is not much different from that of molecules in water under everyday conditions, so that there is no support for a very high viscosity in the core.

This conclusion about the rather low viscosity of the liquid outer core has been confirmed by first-principles m.d. simulations in which the viscosity itself is computed [9, 18, 245]. It might be thought that this would necessitate a simulation of *l*-Fe under shear flow. However, a far more effective way of calculating the viscosity in m.d. is to make use of the fluctuation–dissipation theorem, which provides a formula for η in terms of the autocorrelation function of spontaneous stress fluctuations in an unperturbed system in thermal equilibrium:

$$\eta = \frac{V}{k_B T} \int_0^\infty \mathrm{d}t \, \langle \sigma_{xy}(t) \sigma_{xy}(0) \rangle, \tag{38}$$

where σ_{xy} is the off-diagonal component of the stress tensor and V is the volume of the system. First-principles calculations of η of *l*-Fe at core conditions using this approach [18] give viscosity values in the region of 5 mPa s (the values depend somewhat on thermodynamic state). Similar values of the viscosity of Fe under core conditions emerge from the first-principles-based simulations of Laio *et al* [159]. For comparison, the viscosity of water under ambient conditions has the rather similar value of 1 mPa s.

We now return briefly to the modelling of *l*-Fe using empirical potentials [18]. There are three motivations for considering this. The first is that m.d. based on empirical models is many orders of magnitude faster than first-principles m.d., so an accurate empirical model can be used for much longer simulations on much larger systems. The second motivation is that an empirical model can give revealing insights into the first-principles results. The third is that empirical models play a key role in first-principles calculations of free energies (see section 7). The important thing to appreciate here is that if an empirical total energy function $U_{\mathrm{emp}}(\boldsymbol{r}_1, \ldots \boldsymbol{r}_N)$ can be found which accurately mimics the DFT function $U_{\mathrm{DFT}}(\boldsymbol{r}_1, \ldots \boldsymbol{r}_N)$ for all relevant configurations of the atoms, then by definition simulations using the empirical model will give essentially the same results for structural quantities such as $g(r)$ and dynamical quantities such as D and η. The remarkable fact is that this can be achieved for high-pressure *l*-Fe with extremely simple models for $U_{\rm emp}$.

The relevant criterion for a model $U_{\rm emp}$ to accurately reproduce $U_{\rm DFT}$ is that the difference $\Delta U = U_{\rm emp} - U_{\rm DFT}$ should be essentially constant for all relevant configurations $\{\boldsymbol{r}_i\}$. Specifically, the requirement is that the fluctuations $\delta\Delta U \equiv \Delta U - \langle\Delta U\rangle$ should be as small as possible, $\langle\Delta U\rangle$ being the thermal average of ΔU. The straightforward way to apply this criterion is to adopt a parametrized model $U_{\rm emp}$ and to minimize the quantity $\langle|\delta\Delta U|^2\rangle$. For *l*-Fe under core conditions, the first-principles m.d. simulations [18] show that a pair potential model for $U_{\rm emp}$, with a repulsive inverse power pair potential, can achieve extremely small values of $\langle|\delta\Delta U|^2\rangle$. Explicit comparison of the radial distribution function, diffusion coefficient and viscosity of the empirical and first-principles systems confirms that the model does mimic the first-principles system very closely. This leads to a major simplification in describing the system, since with an inverse power model all the properties of the system depend non-trivially only on a single thermodynamic parameter, instead of on both pressure and temperature.

The finding that an inverse power model reproduces the first-principles properties of *l*-Fe also illuminates the inderlying physics of the system. The energetics of a transition metal clearly must be governed both by Pauli exclusion repulsion of the electron distributions on the atoms, and by d-band bonding. What the pair model reveals is that, although d-band bonding is clearly crucial in determining the total energy of the system, the structure and dynamics are completely determined by the short-range repulsion [21].

Before leaving *l*-Fe, we mention that quite extensive first-principles m.d. work has also been reported on liquid alloys of Fe with sulfur and oxygen [10, 19, 270]. These are very relevant to the Earth's core, since S and O are two of the three main candidates for light impurities in the core (the other is Si). In the first of these papers, the main aim was to test suggestions that the viscosity might be considerably greater than that of pure Fe because the liquid structure could be changed by the presence of S. The simulations showed that in fact the structure is changed very little, and the self-diffusion coefficients of both Fe and S are very similar to their values in pure liquid Fe at the same pressure and temperature. At much lower pressures than those of the core, direct comparisons between first-principles simulations of liquid Fe–S alloys and experimental measurements of diffusivity and viscosity confirm the reliability of the simulations [270]. At low pressures, the solubility of O in liquid Fe is very small, and it has been doubted whether significant amounts of O could be held in solution in the Earth's core. First-principles simulations of liquid Fe–O alloys under core conditions [19] indicate that they are stable against precipitation of FeO up to at least 30 mol% of O.

6.6. Materials of the Earth's mantle

Some of the earliest first-principles m.d. work on $MgSiO_3$ perovskite at high T was reported by Oganov *et al* [200, 201]. They showed that the calculated thermal expansivity of $1.86 \times 10^{-5}\,{\rm K}^{-1}$ at ambient conditions agreed with the available experimental data in the range $1.5\text{–}2.2 \times 10^{-5}\,{\rm K}^{-1}$. They obtained the high-T elastic constants by directly calculating the thermal average stress tensor as a function of strains of chosen symmetries. An important result from this work concerned the quantity $R \equiv d\log V_{\rm s}/d\log V_{\rm p}$, where $V_{\rm p}$ and $V_{\rm s}$ are the longitudinal and transverse velocities. Their calculated value $R \simeq 2.0$ is significantly below the value of at least 2.5 observed in the deepest parts of the lower mantle. They concluded that the mantle must be either chemically heterogeneous or significantly anelastic. First-principles m.d. has also been used very recently by Stackhouse *et al* [244] to study the seismic anisotropy of $MgSiO_3$ perovskite and post-perovskite, and they confirmed that the observed seismic anisotropy in the D″ zone is consistent with the post-perovskite phase. Work by Oganov *et al* [205] using first-principles metadynamics (see below section 6.7) has suggested the existence of polytypic structures intermediate between $MgSiO_3$ perovskite and post-perovskite. This work has also indicated the mechanism of plastic deformation of these structures and has led to a new interpretation of the anisotropy of the D″ zone.

We noted in section 2.1 that the Grüneisen parameter γ is crucial in understanding the temperature distribution in the Earth's lower mantle. Since the lower mantle is convecting, the relation between temperature and pressure is expected to be adiabatic, so that the variation of temperature with radius is given by equation (1), which contains γ as the only unknown quantity. In the first-principles work of Oganov *et al* [200, 201] on $MgSiO_3$ perovskite, γ was determined using the definition $\gamma = V(\partial p/\partial U)_V$. For a given fixed V, m.d. calculations at high T were used to calculate the thermal average pressure p. It was found that to a very good approximation the temperature dependence of U is given by the classical harmonic expression $U = U_{\mathrm{perf}} + 3Nk_{\mathrm{B}}T$, with U_{perf} the total energy of the static perfect lattice and N the number of ions in the system. From this, $(\partial p/\partial U)_V$ and hence γ was estimated using the ratio of finite differences. It turned out that at ambient pressure γ for $MgSiO_3$ perovskite is 1.51 ± 0.1, close to earlier estimates [30]. With increasing pressure, γ decreases smoothly to a value of ~ 1.2 at the pressure of the base of the mantle. The mantle geotherm then follows by integration of equation (1), allowing also for the presence of MgO, whose γ values are very close to those of $MgSiO_3$ perovskite. The resulting temperature at the base of the mantle was estimated as ~ 2600 K [199]. These first-principles m.d. calculations also yielded estimates for the lateral variations of temperature in the mantle. Confirmation for the geotherm obtained in this way comes from recent experimental studies of the post-perovskite transition compared with seismological observation [210].

6.7. The first-principles determination of crystal structures

In section 4, we described how static first-principles calculations can be used to compare the stability of different crystal structures. This straightforward approach has been very important, but suffers from a serious limitation. If we wish to determine what is the most stable crystal structure at a given pressure and temperature, in principle we should compare the Gibbs free energies of all possible structures to find out which has the lowest value. But in practice, it may not be obvious which candidate structures should be examined. What is needed is automatic procedures to do the search for us. In the past few years, such procedures have started to appear, and we summarize two of these here.

The first approach is known as metadynamics, which is a modified form of molecular dynamics simulation. The general ideas of metadynamics [160] are motivated by the fact that simulated condensed-matter systems (or indeed systems in the real world) can be trapped for long periods of time in particular regions of configuration space. A common example of this is thermally activated processes, when the system has to overcome a substantial energy barrier in order to cross from one region of configuration space to another. In such situations, normal molecular dynamics simulation explores only a limited region of configuration space, and other regions may never be reached, even if they are associated with lower free energies. Metadynamics [160] is a strategy for overcoming this, by discouraging the simulated system from revisiting regions of configuration space that have already been visited. In practice, the way this is done is to identify a small set of dynamical variables, which are the 'slow' variables, i.e. the ones responsible for the trapping of the system. A record is kept of the trajectory of these variables, and repulsive potentials are planted on earlier points of the trajectory. This device obliges the slow variables to move away from their original values, until they attain values at which a transition to another region of configuration space spontaneously occurs. The process is then repeated.

This rather general idea has recently started to be applied to the problem of determining crystal structures [172]. To do this, the cell vectors $\boldsymbol{a}$, $\boldsymbol{b}$ and $\boldsymbol{c}$ of the m.d. simulation cell are themselves treated as dynamical variables, following well-established techniques pioneered by Parrinello and Rahman [214]. These cell vectors are used as the 'slow' variables in a metadynamics simulation. The effect of this is to generate a wide-ranging exploration of different cell shapes and volumes, and hence crystal structures, associated with specified p and T. An attractive general feature of metadynamics is that it automatically provides an estimate of the free energies associated with the various regions explored, and this allows one to identify which of the crystal structures explored is the most stable at the given (p, T) conditions.

The metadynamics approach to the determination of crystal structures has recently been used in a re-examination of $MgSiO_3$ under the conditions of the lower mantle [206]. This work has fully confirmed that the post-perovskite structure (figure 5) is indeed the structure of lowest free energy at the p and T of the D'' zone. Furthermore, the simulation revealed the existence of polytypes having structures intermediate between those of perovskite and post-perovskite, these structures being surprisingly different from what had been expected.

A second approach to the crystal-structure problem is the use of so-called evolutionary algorithms. As the name implies, evolutionary algorithms are based on ideas that mimic the Darwinian principle of natural selection. In general, one searches for 'individuals' having the greatest 'fitness'. An individual is characterised by a set of parameters, and the fitness of an individual is specified by the degree to which its set of parameters satisfies some chosen criterion. The evolutionary approach works with groups ('populations') of individuals and produces successive generations by mutating and recombining the parameters in a predetermined way. This general approach has been applied many times in the past to the search for the most stable crystal structures. In this context, an 'individual' is a given crystal structure and the 'parameters' are the numbers – lattice parameters, atomic positions – specifying the crystal structure. The 'fitness' is the energy, or more generally, free energy, and optimisation of the fitness means searching for the lowest energy or free energy.

Over the past 10 years, the evolutionary approach to the determination of crystal structures has had rather mixed success. But very recently, a new version of this idea has been reported, which appears to be very successful [205]. This technique, embodied in a computer code known as USPEX, has been applied with first-principles calculations in a combined theoretical–experimental study of the high-pressure phases of $CaCO_3$ [205], which is important in the global carbon cycle of the Earth. It seems likely that this approach, as well as the metadynamics method, will play an increasingly important role in first-principles mineral crystallography.

7. Free energy

7.1. Basic concepts

When two phases of a pure material, for example solid and liquid, co-exist in thermodynamic equilibrium at a given pressure and temperature, their Gibbs free energies must be equal. If the material contains two or more chemical components, then the chemical potentials of each component in the coexisting phases must be equal. Because thermodynamic and chemical equilibria play such a key role in the Earth and the planets, the first-principles calculation of free energies and chemical potentials is extremely important. As we remarked in section 5, for harmonically vibrating solids, these quantities can be calculated from the first-principles vibrational frequencies. However, for anharmonic solids and for liquids this approach is powerless. In spite of this, it has become clear in the last few years that free energies and chemical potentials of both anharmonic solids and liquids can be calculated with essentially the same first-principles precision as they can be for harmonic solids. Free energies will be discussed in this section, chemical potentials being deferred to section 9.

We recall some elementary ideas. The Helmholtz free energy F is defined as $F = E - TS$ and the Gibbs free energy as $G = F + pV$, where E and S are the internal energy and entropy. Since a system is in full thermodynamic equilibrium at a given p and T when G is a minimum, phase 1 is more stable than phase 2 when $G_1 < G_2$, and the reverse is true when $G_2 < G_1$. The two phases are in stable coexistence when $G_1(p, T) = G_2(p, T)$, and this condition determines the relation between p and T on the phase boundary. It is useful to note also that since pressure is given by $p = -(\partial F/\partial V)_T$, the Gibbs free energy can readily be calculated from the Helmholtz free energy. The isothermal infinitesimal form of this relation, $\mathrm{d}F = -p\,\mathrm{d}V$, expresses the fact that dF is the isothermal reversible work done in an infinitesimal change of volume. The Helmholtz free energy is also important because it provides a very general route from statistical mechanics to thermodynamics, *via* the relation $F = -k_\mathrm{B}T \ln Z$, where Z is the partition function.

7.2. *Statistical mechanics: thermodynamic integration*

For most problems in the earth and planetary sciences, the temperature is well above the Debye temperature, so that the nuclei can be treated as classical particles. This means that it is a good approximation to use the standard formula of classical statistical mechanics for the partition function Z.

For a solid or liquid containing only a single type of nucleus, the first-principles Helmholtz free energy with the nuclei treated classically is given by

$$F = -k_\mathrm{B}T \ln Z = -k_\mathrm{B}T \ln \left\{ \frac{1}{N!\Lambda^{3N}} \int \mathrm{d}\boldsymbol{r}_1 \ldots \mathrm{d}\boldsymbol{r}_N \exp\left[-\beta U(\boldsymbol{r}_1, \ldots \boldsymbol{r}_N)\right] \right\}, \tag{39}$$

where Λ is the thermal wavelength of the nuclei and $U(\boldsymbol{r}_1, \ldots \boldsymbol{r}_N)$ is the first-principles total energy of the system in which the N nuclei are fixed at positions $\boldsymbol{r}_1, \ldots \boldsymbol{r}_N$. (As usual, if thermal excitation of the electrons is included *via* finite-temperature DFT, $U(\boldsymbol{r}_1, \ldots \boldsymbol{r}_N)$ represents a *free* energy.) In the multiple integral, all nuclear positions range over the entire volume of the system. For the periodic systems generally used for DFT calculations on condensed matter, U means the first-principles total energy per repeating cell, and the integrations go over the volume of the cell. In order to obtain the true thermodynamic free energy, we must take the 'thermodynamic limit', in which the volume of the cell and the number of atoms are taken to infinity at constant density.

The formula for F cannot be evaluated as it stands, because there is no known way of evaluating the multi-dimensional integral. Fortunately, this problem was studied and solved many years ago in the context of classical simulations. (A wide-ranging exposition of techniques for calculating free energies in simulation based on empirical interaction models can be found in [101].) The key idea is to use the defining equation (39) for F to derive relations that allow one to calculate the change of free energy when certain well-defined operations are performed on the system.

The simplest of such operations are changes of volume and temperature. Since $p = -(\partial F/\partial V)_T$, and since p can be calculated as a thermal average in first-principles m.d. (section 6.1), the difference of F for two volumes can be calculated by numerical integration of p. Similarly, the Helmholtz formula $E = (\partial(F/T)/\partial(1/T))_V$ can be used to find the difference in values of F/T at two temperatures by the integration of internal energy E, with E computed as a thermal average. In classical simulation, one of the standard ways of calculating the free energy of an anharmonic solid is to integrate E from the low-temperature regime where the quasi-harmonic approximation is accurate (see e.g. [56]).

Many other operations are also possible. The basic idea of the technique known as ‘thermodynamic integration’ (see [101] and references therein) is to study the change of free energy of a system when its total energy function $U(\boldsymbol{r}_1, \ldots \boldsymbol{r}_N)$ is changed. This idea is important, because it allows one to relate the unknown first-principles free energy to the known free energy of some other system. The derivation of the general formula for the free-energy change due to a change in U is simple enough to be given here.

Suppose we have two systems, both having the same number N of atoms contained in the same volume, whose total energy functions are $U_0(\boldsymbol{r}_1, \ldots \boldsymbol{r}_N)$ and $U_1(\boldsymbol{r}_1, \ldots \boldsymbol{r}_N)$. Call their free energies F_0 and F_1. In the present context, ‘1’ will be identified as the first-principles system and ‘0’ as the model system, whose free energy F_0 we know. Now imagine that we transform system ‘0’ continuously into system ‘1’ by changing the total energy function. To describe this, let the total energy function $U_\lambda(\boldsymbol{r}_1, \ldots \boldsymbol{r}_N)$ depend on a parameter λ which varies smoothly from 0 to 1; when $\lambda = 0$, we have the total energy function of system 0, and when $\lambda = 1$ we have the total energy function of system 1. Denote by F_λ the free energy associated with U_λ.

A formula for the variation of F_λ with λ can be derived from the definition of F_λ:

$$\begin{aligned}\frac{\mathrm{d}F_\lambda}{\mathrm{d}\lambda} &= -k_\mathrm{B}T\frac{\mathrm{d}}{\mathrm{d}\lambda}\ln\left\{\frac{1}{N!\Lambda^{3N}}\int \mathrm{d}\boldsymbol{r}_1\ldots\mathrm{d}\boldsymbol{r}_N\mathrm{e}^{-\beta U_\lambda}\right\}\\ &= \int \mathrm{d}\boldsymbol{r}_1\ldots\mathrm{d}\boldsymbol{r}_N\,\frac{\mathrm{d}U_\lambda}{\mathrm{d}\lambda}e^{-\beta U_\lambda}\bigg/\int \mathrm{d}\boldsymbol{r}_1\ldots\mathrm{d}\boldsymbol{r}_N e^{-\beta U_\lambda}\,. \qquad (40)\end{aligned}$$

But $\exp(-\beta U_\lambda(\boldsymbol{r}_1, \ldots \boldsymbol{r}_N))$ is the probability of finding the atoms at positions $\boldsymbol{r}_1, \ldots \boldsymbol{r}_N$, apart from a normalization constant. It follows that

$$\mathrm{d}F_\lambda/\mathrm{d}\lambda = \langle \mathrm{d}U_\lambda/\mathrm{d}\lambda\rangle_\lambda, \qquad (41)$$

where $\langle \mathrm{d}U_\lambda/\mathrm{d}\lambda\rangle_\lambda$ denotes the thermal average of $\mathrm{d}U_\lambda/\mathrm{d}\lambda$ for the system whose total energy function is U_λ. Integrating this relation, we get

$$F_1 = F_0 + \int_0^1 \langle \mathrm{d}U_\lambda/\mathrm{d}\lambda\rangle_\lambda\,\mathrm{d}\lambda. \qquad (42)$$

This is the thermodynamic integration formula, which gives the change of free energy $F_1 - F_0$ caused by changing the total energy function. The crucial point to note is that $\langle \mathrm{d}U_\lambda/\mathrm{d}\lambda\rangle_\lambda$, being a thermal average, can be evaluated by m.d. simulation. The idea is then to perform a sequence of m.d. simulations with values of λ going from 0 to 1, calculate the value of $\langle \mathrm{d}U_\lambda/\mathrm{d}\lambda\rangle_\lambda$ in each of these simulations and then apply equation (42) by numerical evaluation of the integral. This yields the required difference between the first-principles free energy F_1 and the known free energy of the system ‘0’. In practice, the usual choice of U_λ is $U_\lambda = (1-\lambda)U_0 + \lambda U_1$, in which case $\mathrm{d}U_\lambda/\mathrm{d}\lambda = \Delta U \equiv U_1 - U_0$. With this choice, the formula becomes

$$F_1 = F_0 + \int_0^1 \langle \Delta U\rangle_\lambda\,\mathrm{d}\lambda.. \qquad (43)$$

When the thermodynamic integration formula (43) is used to calculate first-principles free energies, the system ‘0’ whose free energy F_0 is already known is called the ‘reference’ system. There are many different ways of choosing reference systems. In one of the first applications of these methods [257], first-principles calculations of the free energies of solid and liquid silicon were used to compute its melting properties. In that work, the reference system for the solid was chosen to be the Einstein model of atoms bound to lattice sites by harmonic springs;

for the liquid, the reference system was the Stillinger–Weber model [249], whose free energy was known from an earlier work [56].

The efficiency of thermodynamic integration is enormously improved if the reference system is close to the first-principles system, in the sense that $U_0(\boldsymbol{r}_1, \ldots \boldsymbol{r}_N)$ strongly resembles the first-principles total energy $U_1(\boldsymbol{r}_1, \ldots \boldsymbol{r}_N)$. There are two reasons for this: first, the number of λ values at which one needs to evaluate $\langle \Delta U \rangle_\lambda$ is reduced; second, the duration of the m.d. runs needed to evaluate $\langle \Delta U \rangle_\lambda$ is also reduced. If the fluctuations of ΔU are small enough, F_1 is accurately given by the leading terms of thermodynamic perturbation theory (see, e.g. Chandler [63])

$$F_1 = F_0 + \langle \Delta U \rangle_0 - \langle (\delta \Delta U)^2 \rangle_0 \big/ (2 k_{\mathrm{B}} T) + \ldots . \tag{44}$$

Here, the thermal averages $\langle \Delta U \rangle_0$ and $\langle (\delta \Delta U)^2 \rangle_0$ are taken in the reference system, and $\delta \Delta U$ is the fluctuating part of U in this system, namely $\delta \Delta U \equiv \Delta U - \langle \Delta U \rangle_0$. Further details about these ideas can be found in [21, 23].

In order to achieve good efficiency, careful thought should always be given to the design of reference systems. Fortunately, good empirical total-energy functions are already available for many types of material. For example, rigid-ion and polarizable-ion models of the Born–Mayer type for oxides, silicates and many other minerals have been in widespread use for many years [59, 94]. A large amount of work has gone into the development of total-energy models for water and other molecular systems (see e.g. [225] and references therein). For transition metals, there is a whole hierarchy of models that can be used, ranging from tight-binding schemes, through embedded-atom models, to simple pair potentials [94]. This means that there is often a ready-made reference system that can be used. But the quality of the reference system can almost always be improved by ‘tuning’ the model parameters to minimize the ‘distance’ between the reference and first-principles systems.

Thermodynamic integration provides a scheme for evaluating the first-principles free energy F. Before discussing practical applications, we return to our statement that the scheme allows F to be evaluated with high precision. It will become clear that for some applications, a precision as high as 10 meV/atom is needed. Since we require F in the thermodynamic limit of infinite system size, this implies, for example, that the average $\langle \Delta U \rangle_\lambda$ appearing in the thermodynamic integration formula must be evaluated with sizes of simulated system large enough to reduce the size error below the specified tolerance. The duration of the m.d. simulation runs must also be long enough to reduce statistical errors below this tolerance. In addition, for any given system size, checks need to be made on the errors due to electronic k-point sampling. Furthermore, the number of λ-points used for the numerical evaluation of the integral of equation (43) must be large enough. All this means that considerable attention to detail is needed to make the scheme work correctly. These issues of precision are discussed in detail in [21, 23], where it is shown that, at least for some systems, a precision approaching 10 meV/atom is indeed achievable.

7.3. The practical calculation of first-principles free energies

In section 5, we summarized work by several groups on the first-principles calculation of free energy in the quasi-harmonic approximation for various materials. More recently, thermodynamic integration has been used in extensive calculations of the free energy of iron in several crystal structures, at least one of which is highly anharmonic, and also in the liquid state [21, 23, 269]. We outline this work here.

We have seen that the simple inverse-power model serves as an excellent reference system for first-principles l-Fe over a wide range of conditions (section 6.5). This reference system has convenient scaling properties, which make it possible to represent its free energy almost exactly with a simple analytic formula. The differences between the reference and first-principles systems are small enough that they are accurately given by the second-order expansion, equation (44). These differences are weakly varying functions of density and temperature, which can be represented as low-order polynomials. The result is that the first-principles free energy of the liquid is given to a precision approaching 10 meV/atom by a simple analytical formula [23]. Remarkably, it turns out that almost the same inverse-power model is a very good reference system also for the high-temperature Fe solid near the melting line. This is unexpected, since the harmonic phonon frequencies of the model reproduce the first-principles frequencies only qualitatively [21]. Thermodynamic integration then leads to a simple but accurate formula for the first-principles free energy of the hcp solid, with full inclusion of anharmonicity [21]. From these free energies, it is straightforward to determine the melting curve and other melting properties, as we describe in section 8.

The formulae for the first-principles free energies can be used to derive formulae for all other thermodynamic quantities, such as expansion coefficient, bulk modulus and Grüneisen parameter as a function of thermodynamic state, and results for these and other quantities can be found in the original papers [21, 23, 268]. As we noted in section 2.1, these thermodynamic calculations are crucial for constraining the temperature distribution in the Earth's liquid outer core. Since this part of the core is convecting, the variation of temperature with pressure is expected to follow an adiabat, as in equation (1). The computation of the p–T adiabats and other thermodynamic quantities for l-Fe is described in detail in [268]. We return in section 9.3 to the use of these adiabats to estimate the temperature at the core–mantle boundary. An assumption that has been made in the past in estimations of the core geotherm is that the Grüneisen parameter of l-Fe under core conditions has an almost constant value of $\sim$1.5. The first-principles results confirm that this is an accurate assumption.

The correctness of the thermodynamic results can be checked by comparing with shock measurements on solid and liquid Fe. We show in figure 13 the first-principles prediction for the principal Hugoniot for Fe compared with experimental data. The excellent agreement gives direct confirmation of the accuracy of the first-principles methods. This goes beyond the low-temperature cold-compression comparisons presented in section 4, because here the comparison depends crucially on thermal pressure effects. The same calculations yield T as a function of p on the Hugoniot (figure 14). The theoretical results in these two plots agree closely with earlier first-principles-based free-energy calculations [254, 279] obtained using a statistical-mechanical approximation known as the particle-in-cell method. We recall that temperature is notoriously difficult to measure in shock experiments. The substantial difference between the experimentally estimated temperature and the two independent first-principles predictions (figure 14) confirms evidence from other work that the experimental temperature is probably unreliable.

We conclude our illustrations of first-principles free-energy calculations by returning to the question of the thermodynamically most stable structure of Fe under inner-core conditions. We saw in section 6.4 that under these conditions bcc Fe is entropically stabilized against the low-temperature elastic and vibrational instabilities from which it suffers at high pressures. However, to know whether it might be present in the inner core, its free energy must be compared with that of hcp Fe. Remarkably, it turned out once again that the inverse-power model provides an excellent reference system, also for the high-temperature bcc phase, so that first-principles free-energy calculations can be performed with only short integration paths [269]. This is an intriguing example of the technique, because it is a case where quasi-harmonic calculations on a crystal are completely meaningless. Near the melting curve, the Gibbs free energy of the bcc structure is predicted to lie above that of hcp by only

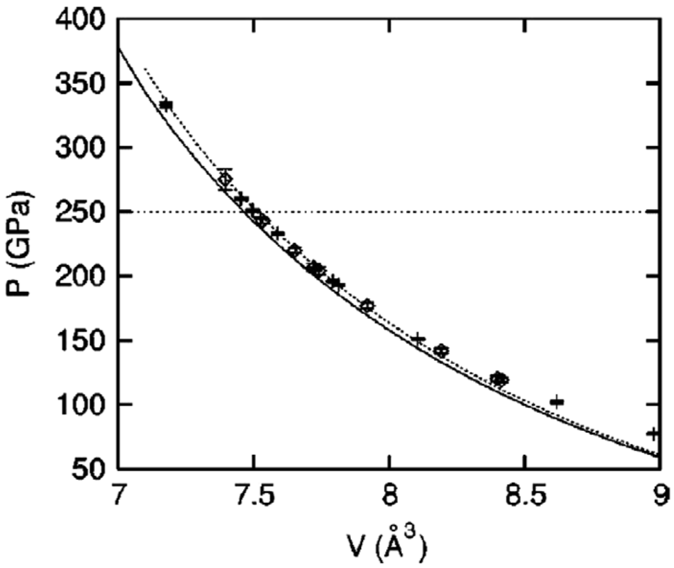

Figure 13. Experimental and first-principles Hugoniot pressure p of Fe as a function of atomic volume V. Symbols show the measurements of Brown and McQueen [57]. Solid curve is first-principles pressure obtained when calculated equilibrium volume of bcc Fe is used in the Hugoniot–Rankine equation; dotted curve is the same, but with experimental equilibrium volume of bcc Fe. The comparison is meaningful only up to a pressure of ca 250 GPa (······), at which point the experiments indicate melting. Reprinted with permission from [21]. Copyright 2001 American Physical Society.

$\sim$35 meV/atom, which is comparable with the precision of the calculations. The message that emerges from this is that bcc must, after all, be regarded as a serious candidate for the structure of the inner core. This conclusion is reinforced when one considers the likely effect of light impurities in the core, a theme to which we return in section 9.

8. Melting

Solid–liquid equilibria are important throughout the earth and planetary sciences. An example of the key importance of knowing melting properties is the equilibrium between the liquid outer core and the solid inner core of the Earth. This equilibrium provides one of the few ways of determining the temperature at the inner-core/outer-core boundary (ICB), which in turn gives estimates of the temperature distribution throughout the core. If, as a zeroth approximation, we assume that the core consists of pure Fe, then the temperature at the ICB must be equal to the melting temperature of Fe at the ICB pressure of 330 GPa. Furthermore, the magnitudes of both the density discontinuity at the boundary and the latent heat are crucial quantities in understanding the heat released by the growth of the inner core due to crystallization of the liquid outer core [117, 118].

These questions have stimulated a strong experimental effort to measure the melting properties of geological materials up to high pressures. There are two main approaches. One consists of static compression, performed by multi-anvil or diamond-anvil techniques, which at present are able to reach pressures of $\sim$ 30 GPa and $\sim$200 GPa respectively. The other approach exploits the fact that phase transitions, including melting, can be detected in shock experiments. Although discontinuities in the principal Hugoniot $p(V)$ itself are extremely weak, sound

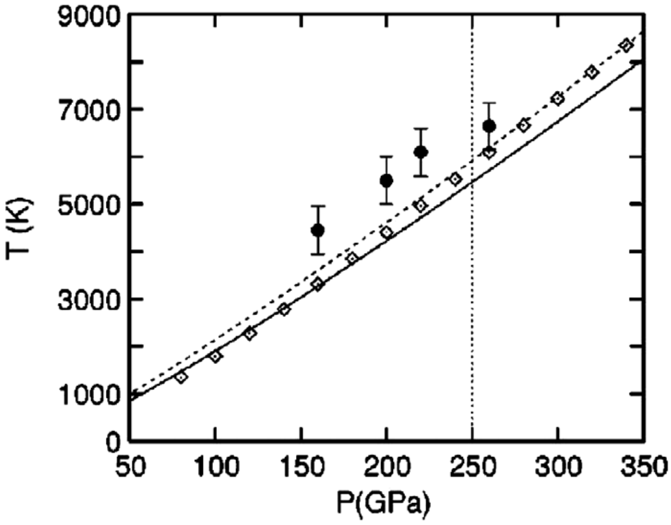

Figure 14. Experimental and first-principles Hugoniot temperature T of Fe as a function of pressure. Black circles with error bars and white diamonds are measurements due to Yoo *et al* [290] and estimates due to Brown and McQueen [57], respectively. Solid and dashed curves are first-principles results obtained using theoretical and experimental initial bcc volumes. The comparison is meaningful only up to a pressure of ca 250 GPa (vertical dotted line), at which point the experiments indicate melting. Reprinted with permission from [21]. Copyright 2001 American Physical Society.

waves can be detected in shock experiments, and discontinuities in the sound-wave velocity can provide prominent indications of melting. However, the experiments are challenging, and their interpretation may not always be straightforward. The determination of melting in static-compression experiments is not unambiguous, and the temperature distribution in the samples may be significantly non-uniform. The difficulty of measuring temperature in shock experiments is a major problem, though the possibility of taking the missing information from first-principles calculations gives one away forward. One symptom of the problems is the large disagreements between melting curves deduced from static-compression and shock measurements, which can be as much as several thousand kelvin [91]. Furthermore, none of the techniques can give reliable values for the volume and latent heat of melting at high pressures.

8.1. The three approaches to first-principles melting

In the last few years, it has become clear that the first-principles calculation of melting properties is not only feasible but also reliable, provided enough attention is paid to technical sources of error [24]. This means that first-principles calculations are beginning to be taken seriously as a way of resolving conflicts between experimental measurements [87]. There are currently three broad approaches to the first-principles treatment of melting. The first is the direct application of the free-energy methods outlined in section 7. The Gibbs free energies of solid and liquid are calculated as a function of p and T, and the melting temperature at any p is the T at which the free energies cross (figure 15) [13, 81, 257]. The second approach is to use first-principles calculations to construct an empirical total-energy model which accurately mimics both solid and liquid. This model is then used to determine the melting curve and other melting

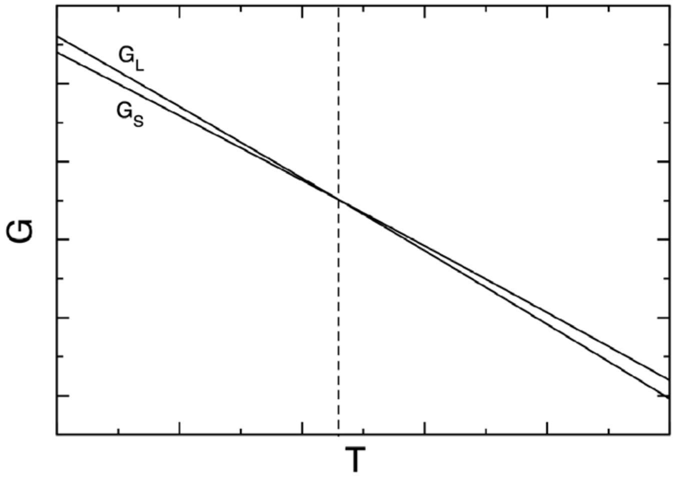

Figure 15. Schematic plot of Gibbs free energies G_S and G_L of solid and liquid as function of temperature at fixed pressure, to illustrate the procedure for determining melting temperature.

properties, either by free-energy methods [184,183] or by performing m.d. simulations on large systems containing solid and liquid in coexistence [22,42,159]. The third approach is to use direct first-principles m.d., without the use of any auxiliary models, to simulate a large system containing coexisting solid and liquid [7,17,50,209]. Needless to say, provided all technical errors are eliminated or corrected for, then for a given exchange-correlation functional, all three approaches should give the same melting curves and other melting properties.

In this section, we survey the three approaches in turn, summarizing the methods that have been developed and the practical results obtained and comparing the results of the methods with each other and with experiment wherever possible.

8.2. Free-energy approach

The idea of using free energies calculated from first principles to predict the melting properties of materials was first described by Sugino and Car [257], who used it to study the melting of Si at ambient pressure. The volume and entropy changes on melting that they found were in good accord with experiment, but the melting temperature itself was about 20% too low. It has since been shown [11] that the discrepancy in T_m was almost certainly due to their use of the LDA, which gives non-cancelling errors in the energies of the semi-conducting solid and the metallic liquid. Their work was followed by first-principles free-energy calculations of the melting properties of Al at ambient pressure [81], which agreed rather closely with experiment. The first extensive melting curve of any material by this first-principles approach was reported by Alfè *et al* [13], who evaluated the DFT–GGA melting curve of Fe for the pressure range 100–350 GPa relevant to the Earth's core. The same approach has also been used to compute the melting curve of Al from ambient pressure up to 150 GPa [266]. Although the melting of Al has no direct relevance to the earth sciences, the latter work is significant here because the close agreement with the rather complete experimental data provides valuable evidence for the reliability of the methods. The technical details of the calculations on Al and Fe are very similar, and we summarize them briefly here.

The construction of reference models for use as a starting point for thermodynamic integration has been described in section 7, where we saw that for liquid Fe an inverse power pair potential is surprisingly effective. We noted there the simplifications that follow from

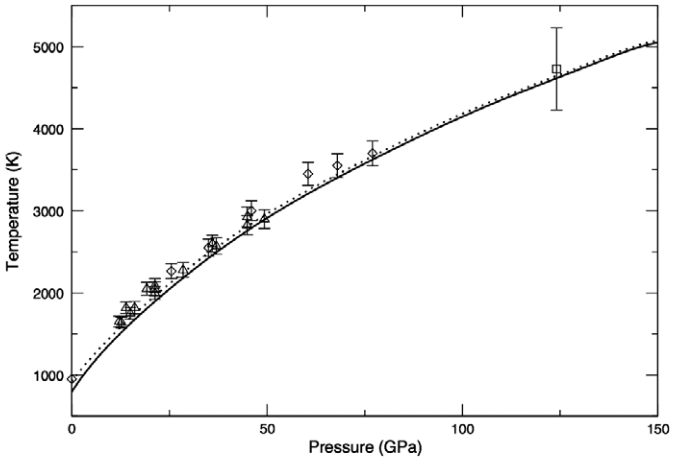

Figure 16. Melting curve of Al from DFT calculations (solid and dotted curves), compared with experimental results (diamonds and triangles: DAC measurements, square: shock measurement). Reproduced with permission from [266]. Copyright 2002 American Physical Society.

the scaling properties of this simple reference model. The same kind of inverse-power model turns out to be equally effective for liquid Al [266]. The free-energy difference between the first-principles and model systems for liquid Al was calculated by thermodynamic integration in the same way as for liquid Fe. For high-temperature solid Al, the free-energy calculations again followed the methods used for Fe, the reference model here being a linear combination of the inverse-power model and the quasi-harmonic part of the first-principles total energy function. Estimates of the technical errors in the computed melting curves can be obtained by combining the errors from the various parts of the free-energy calculations. For Al, errors of ±50 K are estimated at low pressure, increasing to ±100 K at 150 GPa. For Fe, the technical errors are estimated to be ±300 K across the whole pressure range.

In figure 16, we compare the first-principles melting curve of Al with the available experimental data. The agreement looks good over the whole pressure range reported. However, closer inspection shows that at ambient pressure the predicted melting temperature is ∼15% below the experimental one (786 K compared with 933 K), though this discrepancy diminishes rapidly with increasing pressure. It turns out that the main reason for this discrepancy is that GGA does not give very accurate phonon frequencies, so that the thermal pressure predicted by GGA is in error. Recognizing this shortcoming of GGA, it is possible to correct for it, and once this is done, the corrected melting curve comes into very close agreement with the experimental data across the entire pressure range.

Figure 17 shows the melting curve of Fe obtained from first-principles free-energy calculations, compared with experimental data from both DAC and shock experiments; also shown are predictions from the approach in which a model is fitted to first-principles data, which will be described in more detail in section 8.3. The early experimental data of Williams *et al* [285] lie considerably above those of other groups and are now generally discounted. This still leaves a range of ∼ 700 K in the experimental values of T_m at 100 GPa. The first-principles melting curve obtained by the free-energy approach lies above the surviving DAC curves, with our T_m being above that that of Ma *et al* by ∼400 K at 100 GPa. However, we need to consider corrections due to known errors of GGA for Fe. We know from the low-temperature equation of state (section 4.1.4) that for Fe the GGA underestimates the pressure by ∼8 GPa at core

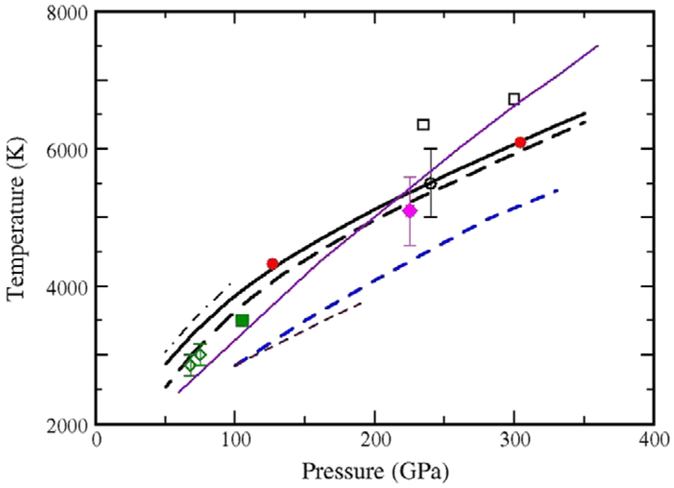

Figure 17. Comparison of melting curve of Fe from DFT calculations and experimental data: black solid and dashed curves: first-principles results of [23] without and with free-energy correction (see text); red filled circles: corrected coexistence results from [24]; blue dashed curve: first-principles based results of [159]; purple curve: first-principles based results of [42]; black chained and maroon dashed curves: DAC measurements of [285] and [49]; green diamonds and green filled square: DAC measurements of [238] and [167]; black open squares, black open circle and magenta diamond: shock experiments of [290], [57] and [193]. Error bars are those quoted in original references.

conditions. If we correct for this in the same way as was done for Al, we find that this lowers the melting curve by ∼350 K at pressures of 50 GPa and by ∼70 K in the region of 300 GPa. The corrected melting curve is also given in figure 17 and shows excellent agreement with the recent DAC data of Shen *et al* [238] and Ma *et al* [167], which lie considerably above the older measurements due to Boehler [49].

In the pressure region above 200 GPa, the only available experimental data come from shock experiments, which should be able to fix a point on the melting curve at the state where melting first occurs on the Hugoniot. However, the difficulties of measuring temperature in shock experiments are well known. The temperatures obtained by Yoo *et al* [290] using pyrometric measurements are generally regarded as too high by at least 1000 K; the strong evidence in support of this statement from first-principles calculations of the Hugoniot temperature was noted in section 7.3. We therefore disregard their point on the curve. In the shock experiments of Brown and McQueen [57] and Nguyen and Holmes [193], no attempt was made to measure temperature, which was estimated using models for the specific heat and the Grüneisen parameter; the approximate validity of these models is supported by first-principles calculations on hcp Fe. We note that the melting curve from free-energy calculations, especially when corrected as described above, agrees well with the shock data of both Brown and McQueen and of Nguyen and Holmes. A useful review of synchrotron techniques, including a section on their contribution to the determination of the melting curve of Fe has recently been published [87].

The large differences between the melting curve from the first-principles free-energy approach and the model fitting approaches will be discussed further after we have described the latter approach in more detail in the next section. We shall show there that these differences are almost certainly due not to errors in the first-principles calculations themselves but to imperfect fitting of empirical models.

8.3. The model-fitting approach

There are many different ways of implementing the general model-fitting approach for the determination of melting properties. For a given material, there may be models having different levels of sophistication and realism. Once a given parametrized model has been chosen, the next problem will be the adjustment of the parameters so that the model mimics the first-principles solid and liquid as accurately as possible. This apparently simple statement conceals a number of puzzles: what first-principles data should be used in fitting the model? By what criteria should one judge the goodness of the fit? How does one decide whether the model is good enough to yield accurate melting properties? Even when these matters are settled, there are several ways of using coexistence simulations to determine the melting properties of the model. We review briefly how these general questions have been approached by different researchers, before outlining the work that has been reported.

Much of the first-principles work reported on melting has been done on transition metals, and these metals give a good illustration of the issues involved in constructing models. A large component in the energy of these metals comes from the quantum mechanics of d-band bonding, so it is natural to work with models that represent this bonding explicitly. The most complete empirical way of doing this is through tight-binding (TB) schemes, which allow the explicit calculation of the electronic band-structure as a function of ionic positions [71, 115, 212, 258]. TB models parametrized against DFT calculations have been widely used in the modelling of transition metals and have been used in the study of Fe at Earth's core conditions [72]. However, even more widely used are simplifications of TB that still include the dependence of d-band energy on ionic positions. The most sophisticated of these is the generalized pseudopotential theory of Moriarty and co-workers [184, 182], which retains a description of the angular dependence of d-band energy. The method variously known as 'second moment', 'embedded atom' [77, 78], and Finnis–Sinclair [93] is a simpler description in which the d-band energy is assumed to depend only on the second moment of the d-band electronic density of states. (The effective medium theory of Nørskov *et al* [131] is closely related.) Simpler yet are models which include the dependence on ionic positions only of the energy of repulsion between the atoms, the d-band energy being represented as a term depending solely on the overall density (and possibly temperature) of the system. The way all these methods have been used in the modelling–fitting approach to the melting of transition metals will be summarized below.

In the earliest first-principles based calculations of melting properties, the model fitting was performed using only first-principles data for the solid. Examples of this approach are the work of Moriarty and co-workers on Al [185], Cu [181], Mo [183] and Ta [184], using generalized pseudopotential models. Comparisons with subsequent first-principles work [271] and with shock data indicate that the predictions were accurate. In more recent work, first-principles data on the high-temperature solid and the liquid are used to parametrize the empirical model. A popular way of doing this is known as 'force matching' [90, 159]. An m.d. simulation of the material at the thermodynamic state of interest is performed and representative atomic configurations are drawn from the simulation. For these configurations, the parameters of the model are adjusted so as to minimize the mean square difference of first-principles and model forces, i.e. the quantity

$$W = \sum_n \sum_{i=1} \left| \boldsymbol{F}_i^{\mathrm{mod}}(t_n) - \boldsymbol{F}_i^{\mathrm{FP}}(t_n) \right|^2, \tag{45}$$

where the sums go over the configurations and over the ions and $\boldsymbol{F}_i^{\mathrm{mod}}(t_n)$ and $\boldsymbol{F}^{\mathrm{FP}}(t_n)$ denote the model and first-principles forces at the time-step t_n. Sometimes, other quantities, such as the pressure, are also included in the fit [159]. Clearly, as W tends to zero, the model system becomes identical to the first-principles system in its energetics and dynamics, except perhaps for an overall constant shift of the total energy of one system with respect to the other. A different, but not unrelated, method is the one described in section 7.2 for the tuning of reference systems. In this method, the parameters are adjusted to minimize the strength of the fluctuations of the difference of the model and first-principles total energies. Other fitting methods have also been proposed. Naturally, for a given parametrized model, different fitting strategies will yield different 'optimized' parameters. Given that no method can give a perfect fit, it is then relevant to ask if one method is preferable to another for the calculation of melting properties. We return to this question later.

If explicit simulation of solid and liquid is used to determine the melting properties of the empirical model, various approaches are possible, the approaches differing mainly in the quantities that are held fixed. In the work of Morris *et al* on the melting of Al [186], which were based on an embedded-atom total-energy model, coexisting solid and liquid Al were simulated with the total number of atoms N, volume V and internal energy E fixed. They showed that, provided V and E are appropriately chosen, the two phases coexist stably over long periods of time, and the average p and T give a point on the melting curve. In the work of the UCL group [22], the constant quantities are chosen to be (N, V, T). A further approach was used in the simulations of Laio *et al* [159] on the high-pressure melting of Fe, which were performed at constant stress, with enthalpy almost exactly conserved. The approach of Belonoshko *et al* [41,42] is different again. Here, the (N, p, T) ensemble is used. the system intially contains coexisting solid and liquid, but since p and T generally do not lie on the melting curve the system ultimately becomes entirely solid or liquid. Unlike the other approaches, this one does not yield directly points on the melting curve but instead provides upper or lower bounds, so that a series of simulations is needed to locate the transition. For a given model, the various methods will yield the same results, provided the simulations are done on large enough systems.

There have been three independent first-principles studies of the high-pressure melting of Fe based on the model-fitting approach, all three of which were based on the embedded-atom model. In the work of Laio *et al* [159], the model used for the coexistence simulations was constructed using force matching, and efforts were made to ensure that the model forces accurately reproduced the first-principles forces both in the solid and the liquid. The work of Belonoshko *et al* [42] was based on fitting the energies of the model to those calculated with FPLMTO at one thermodynamic state in the liquid. The fitting of the embedded-atom model used by the UCL group [22] was performed by minimizing the fluctuations in the energy differences between the model and first-principles in simulations on the liquid state. The predicted melting curves from these studies are shown in figure 17, in comparison with the results from the free-energy method reported by the UCL group [23], together with the experimental static-compression and shock measurements.

The large differences between the theoretical melting curves are of about the same general magnitude ($\sim$1000 K) as the differences between the experimental data. Like the free energy results described in section 8.2, the model-fitting melting curves of Belonoshko *et al* and of the UCL group are consistent with the shock results and with recent static compression data, including those of Ma *et al* [167], but lie well above the Boehler static-compression values [49]; the model-fitting results of Laio *et al* are much lower and agree quite well with the Boehler data. The UCL results for the volume and the entropy change on melting at the pressure of the inner-core boundary were 1.8% and 1.03 k_B.

Obviously, the large disagreements between theoretical results seen in figure 17 are completely unacceptable. Although the electronic structure methods are not identical in the three studies, we would not expect the small difference between them to lead to such large discrepancies in the predictions. A far more likely cause is the differences in the fitting of the empirical models. Clearly, the models have to fit the first-principles data very exactly, otherwise the calculated melting curves cannot be regarded as truly first-principles. If high quality fits cannot be achieved, then at least corrections should be made for the imperfections of the models. To investigate this question, the UCL group developed a scheme [22] for correcting the predicted melting curves for the differences between the fitted models and the first-principles data. The basic idea of this scheme is that differences between the fitted model and the first-principles data represent differences between the model and first-principles free energies. The resulting shifts of the free-energy curves result in a shift of the melting temperature, for which a simple formula can be obtained.

When this correction scheme is applied to our own model-fitting predictions, the resulting melting curve is in essentially perfect agreement with the free-energy predictions. From the information published by Belonoshko *et al* [42] about the parameters of their embedded-atom model, it is possible to estimate the corrections that should be applied to their melting curve. With these corrections, their melting curve comes into good agreement with ours. For the model used by Laio *et al* [159], insufficient information has been published to support a full analysis, but the information that is available suggests that the correction that should be applied to their melting curve is of the correct sign and magnitude to account for the discrepancies.

The conclusion from these comparisons is therefore that the model-fitting approach to the first-principles calculation of melting properties is capable of being reliable but only if corrections are made for deficiencies of the fit of the model to the first-principles data.

8.4. First-principles coexistence approach

The direct simulation of solid and liquid in coexistence has been one of the standard methods for studying melting properties with empirical interaction models. Extensive work of this kind has shown that reliable results are generally obtained provided the entire system (solid and liquid together) contains ∼500 atoms or more. Recent advances of computer power have brought this size of system within range of first-principles m.d., so it is now an attractive possibility to study melting by direct first-principles coexistence simulations. This brings the enormous advantage of entirely avoiding the need for empirical models. The first studies of melting by first-principles coexistence were reported by Alfè [6], and shortly after by Ogitsu *et al* [209] The work of Alfè was on the melting properties of Al, whose melting curve up to 150 GPa had already been studied by the first-principles free-energy method. The first-principles work of Ogitsu *et al* was on LiH. Neither material is of direct relevance to the earth sciences, but these studies are important for the present review because of the possibilities that they open up for the future.

The coexistence simulations of Alfè were done using the VASP code with the GGA for exchange and correlation, with the simulation cells having a range of sizes from 512 to 1728 atoms. We noted before that coexistence simulations can be done in a variety of ways. The approach used by Alfè was based on the (N, V, E) ensemble; the techniques used to bring solid and liquid into thermal equilibrium in the simulation cell with this approach are described in the original paper. Points on the melting curve obtained from the simulations with different sizes of system are compared in figure 18 with the melting curve from first-principles free-energy calculations using exactly the same exchange-correlation functional. The melting temperatures from the systems of 1000 and 1728 atoms agree to within ∼20 K, but with the system of 512 atoms, there are indications that the melting temperature may be overestimated by ∼50 K. Tests suggested that the error with the 512-atom system may be due to errors of

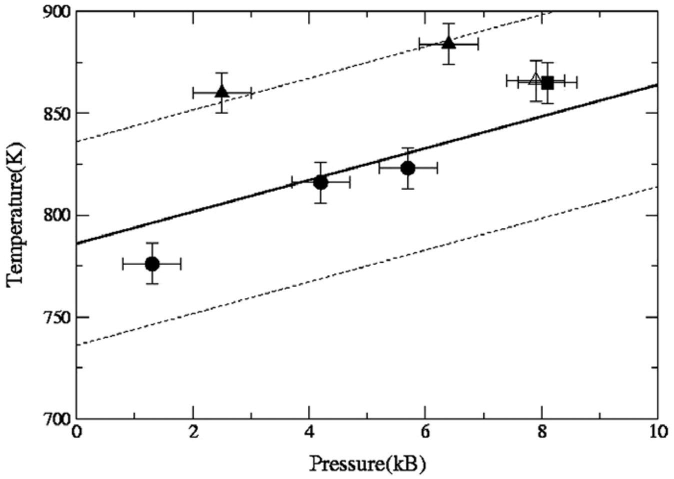

Figure 18. Temperatures and pressures at which liquid and solid Al coexist in first-principles simulations containing 1728 atoms (■), 1000 atoms (●) and 512 atoms (▲) performed using Γ-point sampling only. Open triangle is the result from a $2 \times 2 \times 1$ ***k***-point grid on a 512-atom system. The solid line is the lower end of the melting curve calculated using the free-energy approach in [266] (see text), light dashed lines represent error bars. Results from simulations using Γ-point sampling only for 1000- and 512-atom systems were reported in [6].

k-point sampling, rather than the limited size of the system itself. The overall conclusion is that direct first-principles coexistence is a viable and accurate method for determining melting curves. Other melting curves obtained with the first-principles coexistence method include those of LiH [209] and H [50].

Very recently, first-principles coexistence simulations have been reported on MgO [7]. We pointed out earlier the importance of understanding MgO, because of the presence of large amounts of Mg(Fe)O magnesiowüstite in the lower mantle. The melting curve of MgO at lower mantle pressures (up to 135 GPa) is particularly important, because it places constraints on the possibility of partial melting at the base of the mantle, which has sometimes been suggested. In fact, the melting curve of MgO has been highly controversial. The only experimental data up to lower mantle pressures are those of Zerr and Boehler [291], who found a rather low melting slope $\mathrm{d}T_\mathrm{m}/\mathrm{d}p \simeq 30\,\mathrm{K\,GPa^{-1}}$, which might be compatible with partial melting at the base of the mantle. However, several theoretical calculations based on empirical interaction models for MgO predict much higher melting slopes in the region of $70\,\mathrm{K\,GPa^{-1}}$ [69, 273].

Like the earlier work on Al, the recent first-principles calculations on MgO melting [7] were performed using the VASP code but used an LDA exchange-correlation functional, which gives a better value for the melting temperature at ambient pressure. The simulations were performed mainly on systems of 432 atoms but with one simulation of 1024 atoms as a control on size effects. The results, shown in figure 19, give strong support to the high melting slope predicted by earlier theoretical work and suggest that further experimental work on the melting of MgO would be very valuable.

We have tried to show in this section how the first-principles treatment of solid–liquid equilibria is helping to resolve important problems in the Earth and planetary sciences. However, the methods we have outlined so far are designed for pure single-component systems, and in the real world materials are rarely pure. We have stressed, for example, that the Earth's core cannot be pure Fe, so that the estimate of the temperature at the inner-core boundary from the

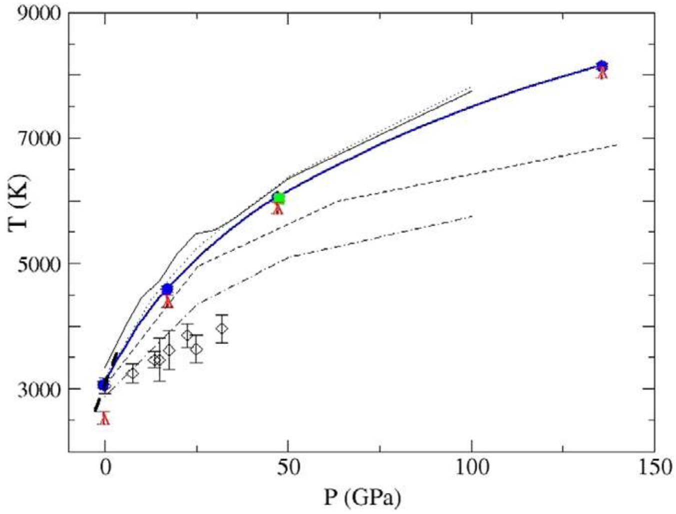

Figure 19. Melting curve of MgO obtained with DFT–LDA coexistence simulations performed on 432-atom cells (blue dots and heavy solid line), 1024-atom cell (green square) and DFT–GGA results (red triangles) [7], compared with experiments (◊) [291]. Other curves show results of earlier modelling work based on interaction models. Reproduced with permission from [7]. Copyright 2005 American Physical Society.

melting curve of pure Fe is only a zeroth approximation. To treat equilibria in multi-component systems, we need to consider chemical potentials, which are the subject of the next section.

9. Chemical potentials

9.1. First-principles methods

When two phases of a multi-component mixture are in thermodynamic equilibrium, the chemical potentials of each component are equal in the two phases: $\mu_i^a = \mu_i^b$, where i labels the chemical component and a, b denote the two phases. Generally, equality of the chemical potentials requires that the mole fractions c_i^a and c_i^b in the two phases are different. A familiar example of this is the use of salt to lower the freezing point of water when the roads are icy. The sodium and chlorine ions are accommodated more easily in liquid water than in ice, and equality of chemical potentials occurs when most of the salt is in the liquid phase. The lowering of the chemical potential of the water caused by this partitioning results in the well-known depression of freezing point. The analogous partitioning of light impurities in the liquid and solid parts of the Earth's core is crucial for an understanding of the core.

Our main interest here is in binary mixtures in which one component is fairly dilute, and we refer to the majority component as the 'solvent', denoted by A, and the minority component as the 'solute', denoted by X. Assuming that the two phases are solid and liquid, we denote the chemical potentials of solute and solvent in the phases by μ_X^s, μ_X^l and μ_A^s, μ_A^l, respectively. Then in thermodynamic equilibrium $\mu_X^s = \mu_X^l$ and $\mu_A^s = \mu_A^l$. The corresponding mole fractions in the two phases are denoted by c_X^s, c_X^l and c_A^s, c_A^l.

In the limit of infinite dilution $c_X \to 0$, the chemical potential μ_X diverges logarithmically, going as $k_B T \ln c_X$. It is convenient to separate out this logarithmic term by writing

$$\mu_X = k_B T \ln c_X + \bar{\mu}_X, \tag{46}$$

where the quantity $\bar{\mu}_{\mathrm{X}}$ is well behaved as $c_{\mathrm{X}} \to 0$. The condition $\mu^s_{\mathrm{X}} = \mu^l_{\mathrm{X}}$ then takes the form

$$k_{\mathrm{B}}T \ln c^s_{\mathrm{X}} + \bar{\mu}^s_{\mathrm{X}} = k_{\mathrm{B}}T \ln c^l_{\mathrm{X}} + \bar{\mu}^l_{\mathrm{X}}, \tag{47}$$

from which we immediately obtain a convenient formula for the ratio of mole fractions of solute in two phases:

$$c^s_{\mathrm{X}}/c^l_{\mathrm{X}} = \exp[(\bar{\mu}^l_{\mathrm{X}} - \bar{\mu}^s_{\mathrm{X}})/k_{\mathrm{B}}T_{\mathrm{m}}]. \tag{48}$$

Here, T_{m} is the temperature at which the two phases coexist at the given pressure and solute concentration. The quantities $\bar{\mu}^s_{\mathrm{X}}$ and $\bar{\mu}^l_{\mathrm{X}}$ themselves depend on c^s_{X} and c^l_{X}, but if these mole fractions are small, then it may be enough to approximate $\bar{\mu}^s_{\mathrm{X}}$ and $\bar{\mu}^l_{\mathrm{X}}$ by their infinite dilution values or to include their dependence on c^s_{X} and c^l_{X} to first order. It is shown in the textbooks (see e.g. Atkins [36]) that the depression of freezing point, i.e. the shift of coexistence temperature $T_{\mathrm{m}} - T^0_{\mathrm{m}}$, with T^0_{m} the melting temperature of pure solvent, is given to lowest order in the solute concentrations by

$$T_{\mathrm{m}} - T^0_{\mathrm{m}} = \frac{k_{\mathrm{B}}T}{\Delta s^0_{\mathrm{A}}}(c^s_{\mathrm{X}} - c^l_{\mathrm{X}}). \tag{49}$$

Here, $\Delta s^0_{\mathrm{A}} \equiv s^{0l}_{\mathrm{A}} - s^{0s}_{\mathrm{A}}$ is the entropy of fusion of pure solvent.

It follows from this that a first-principles treatment of coexisting phases of dilute solutions requires the calculation of the quantities $\bar{\mu}^s_{\mathrm{X}}$ and $\bar{\mu}^l_{\mathrm{X}}$. Practical methods for doing this have been developed in the last few years and, in combination with seismic data, have been used to obtain new constraints on the chemical composition of the Earth's core and the temperature at the ICB, from which the temperature distribution throughout the core can be estimated. Before explaining how these new constraints have been obtained, we summarize briefly the first-principles strategies for calculating the quantities $\bar{\mu}^s_{\mathrm{X}}$ and $\bar{\mu}^l_{\mathrm{X}}$, referring the reader to the published papers for full details [14–15].

In both the solid and the liquid, it is much easier to work with the difference of chemical potentials $\mu_{\mathrm{XA}} \equiv \mu_{\mathrm{X}} - \mu_{\mathrm{A}}$, rather than μ_{X} itself. If we can compute μ_{XA}, then μ_{X} can be obtained from a knowledge of μ_{A}. But at infinite dilution μ_{A} is simply the Gibbs free energy of the pure solvent, for which first-principles values are provided by the methods outlined in section 7. Since the variation of μ_{A} with mole fraction c_{X} is given by the Gibbs–Duhem relation $c_{\mathrm{A}}d\mu_{\mathrm{A}} + c_{\mathrm{X}}d\mu_{\mathrm{X}} = 0$, it follows that μ_{XA} as function of c_{X} gives us everything we need. Now μ_{XA} is the change of free energy of the system when a solvent atom is replaced by a solute atom. In the liquid state, it is convenient to calculate this free-energy change by thermodynamic integration, by computing the work done as an A atom is continuously transmuted into an X atom. Although such an 'alchemical' transmutation obviously cannot be accomplished in the real world, it is rigorously justified within first-principles statistical mechanics and provides an exact algorithm for calculating chemical potentials.

With this approach, the specific quantity that is evaluated by first-principles thermodynamic integration, denoted by $m(c_{\mathrm{X}})$, is defined by the equation:

$$\mu_{\mathrm{XA}} = k_{\mathrm{B}}T \ln \frac{c_{\mathrm{X}}}{1 - c_{\mathrm{X}}} + 3k_{\mathrm{B}}T \ln(\Lambda_{\mathrm{X}} / \Lambda_{\mathrm{A}}) + m(c_{\mathrm{X}}), \tag{50}$$

where Λ_{X} and Λ_{A} are the thermal wavelengths of solute and solvent. The thermodynamic integration formula for $m(c_{\mathrm{X}})$ is then

$$m(c_{\mathrm{X}}) = \int_0^1 \mathrm{d}\lambda \langle U_1 - U_0 \rangle_\lambda, \tag{51}$$

where $U_1(\boldsymbol{R}) \equiv U(N_{\mathrm{A}} - 1, N_{\mathrm{X}} + 1; \boldsymbol{R})$ and $U_0(\boldsymbol{R}) \equiv U(N_{\mathrm{A}}, N_{\mathrm{X}}, ; \boldsymbol{R})$. Here, $U(N_{\mathrm{A}}, N_{\mathrm{X}}; \boldsymbol{R})$ is the first-principles total energy function of the system consisting of N_{A} solvent atoms and N_{X}

solute atoms, while $U(N_A - 1, N_X + 1; \boldsymbol{R})$ is the same, but for $N_A - 1$ solvent atoms and $N_X + 1$ solute atoms. For any value of λ between 0 and 1, the simulated system is governed by the total energy function $U_\lambda(\boldsymbol{R}) \equiv (1-\lambda)U(N_A-1, N_X+1; \boldsymbol{R})+\lambda U(N_A, N_X; \boldsymbol{R})$. As the hybrid system evolves in time, the forces on atoms are the gradients of $U_\lambda(\boldsymbol{R})$, so that at each time step two separate first-principles calculations have to be performed, both for the same atomic positions $\boldsymbol{R} \equiv \{\boldsymbol{r}_i\}$. In one of these calculations, the total energy $U(N_A, N_X; \boldsymbol{R})$ and its associated forces $\boldsymbol{F}_i(N_A, N_X; \boldsymbol{R})$ are calculated, while in the other, the total energy $U(N_A - 1, N_X + 1; \boldsymbol{R})$ and its associated forces $\boldsymbol{F}_i(N_A - 1, N_X + 1; \boldsymbol{R})$ are calculated; the forces $\boldsymbol{F}_i(\boldsymbol{R})$ used to generate the time evolution in the equation of motion (see equation (37)) are then the linear combinations: $\boldsymbol{F}_i(\boldsymbol{R}) = (1 - \lambda)\boldsymbol{F}_i(N_A - 1, N_X + 1; \boldsymbol{R}) + \lambda \boldsymbol{F}_i(N_A, N_X; \boldsymbol{R})$. Exactly as in the thermodynamic integration techniques described in section 7, $m(c_X)$ is obtained by numerical integration of $\langle U_1 - U_0\rangle_\lambda$ computed at a chosen set of λ values. In practice, the statistical efficiency of this scheme is enhanced by switching more than a single atom at the same time.

In the solid state, thermodynamic integration is not the most appropriate way of calculating the chemical potential difference $\mu_{XA} \equiv \mu_X - \mu_A$. This is clear, because in the zero temperature limit, at infinite dilution $m(c_X \to 0)$ is simply the change in internal energy when one atom in the perfect lattice of solvent is replaced by a solute atom, the impurity system being relaxed to equilibrium. This requires only a static calculation of the type described in section 4. At finite temperatures in the infinite dilution limit, $m(c_X \to 0)$ can be obtained from the quasi-harmonic vibrational frequencies of the pure A system and the system containing a single X impurity. If anharmonic effects are significant, as they are in the case of O substituted in hcp Fe [20], thermodynamic integration can used to estimate the anharmonic effects. These methods can also be generalized to include the variation of $m(c_X)$ with c_X to linear order in c_X.

9.2. *Chemical composition and temperature of the core*

We turn now to the application of these techniques to constrain the chemical composition of the Earth's core. The strategy for doing this relies on two key facts already alluded to (section 2.1). The first is that the density of the liquid outer core, as deduced from seismic and free-oscillation measurements, is substantially lower than that of pure Fe at the relevant pressures and temperatures. The second key fact is that the density discontinuity across the ICB is rather well determined by free-oscillation measurements to be 6.5 ± 1.4%. The first-principles simulations of pure liquid Fe described in section 6.5 can be used to refine earlier quantitative estimates of the density deficit in the outer core, which is currently estimated to be $\sim 7\%$. If we assume initially that the core is a binary mixture of Fe and one of the impurities S, Si or O, then for a given candidate impurity, first-principles m.d. simulation can be used to determine what mole fraction of this impurity must be present in the outer core to lower the density to the seismically determined value. For example, if S is assumed to be the impurity, it turns out that 15 mol% must be present in the outer core to lower the density by the required amount. Calculations of the chemical potentials of the candidate impurity in the liquid and solid then determine the impurity mole fraction in the solid inner core. Finally, first-principles simulation can be used to deduce the density discontinuity across the ICB. This can then be compared with the known value of 6.5 ± 1.4%.

The way these comparisons turn out for the three main candidate impurities S, Si and O is displayed in figure 20, which shows the dependence of liquid density on the impurity mole fractions in the liquid outer core, the dependence of inner-core mole fractions on the outer-core mole fractions, and the resulting ICB density discontinuity as a function of the outer-core mole fractions. The comparisons reveal that S and Si partition only very weakly between the solid

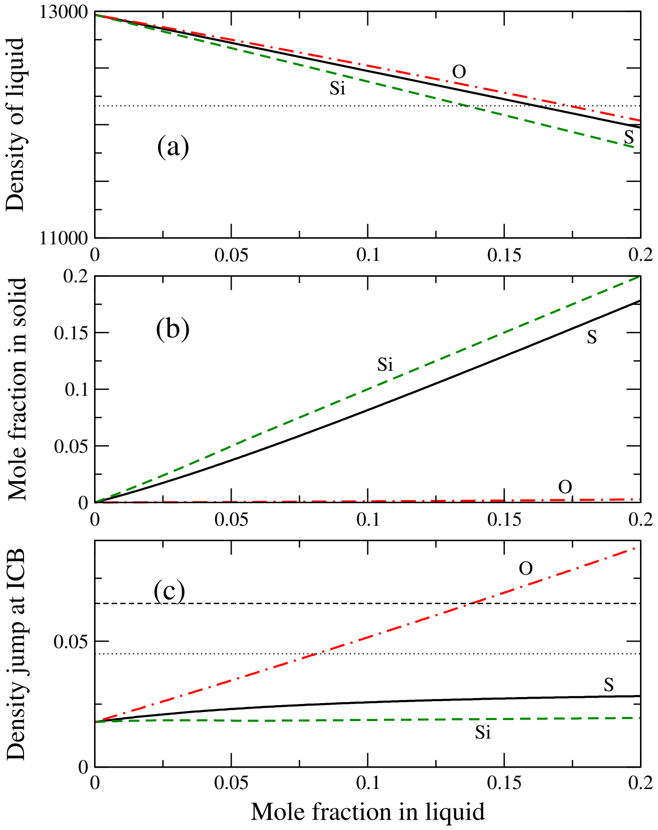

Figure 20. Liquid and solid impurity mole fractions c_X^l and c_X^s of impurities X = S, Si and O, and resulting densities of the inner and outer core of the Earth predicted by first-principles simulations. Solid, dashed and chain curves represent S, Si and O respectively. (*a*) liquid density ρ^l (kg m^{-3} units); horizontal dotted line shows density from seismic data [89, 173]; (*b*) mole fractions in solid resulting from equality of chemical potentials in solid and liquid; (*c*) relative density discontinuity $\delta\rho/\rho^l$ at the ICB; horizontal dotted and dashed lines are values from free-oscillation data [174, 173]. Adapted from [16], with permission.

Table 4. Estimated molar percentages of sulfur, silicon and oxygen in the solid inner core and liquid outer core of the Earth, obtained by combining first-principles calculations and seismic data. Sulfur/silicon entries refer to total percentages of sulfur plus silicon.

	Inner core	Outer core
Sulfur/Silicon	7.0 ± 2.5	8.0 ± 2.5
Oxygen	0.2 ± 0.1	13 ± 2.5

and the liquid, so that for the required outer-core mole fraction, the density discontinuity comes out much smaller than the seismically determined value. By contrast, the required fraction of O in the outer core leads to a density discontinuity that is much too big. It follows that the density discontinuity can only be explained if the outer core contains both O and also some mixture of S and Si, the ratio of the two latter elements being impossible to determine by this route. The chemical compositions of the inner and outer core deduced from these comparisons are summarized in table 4.

With these results for the impurity mole fractions, the depression of freezing point can then be estimated from equation (49). The value $T_m - T_m^0 = 900$ K thus obtained, taken together with the melting temperature $T_{\mathrm{m}}^{0} = 6200\,\mathrm{K}$ for pure Fe at the ICB pressure (section 8.2 and 8.3), leads to the estimate $T_{\mathrm{m}} = 5300\,\mathrm{K}$ as the ICB temperature.

9.3. *Implications for the core geotherm*

As we pointed out in section 2.1, this ICB temperature is crucially important, because it allows us to estimate the geotherm throughout the outer core. Since the outer core is convecting, the temperature T should follow an adiabat, so that the variation of T with radius is given by equation (1). We saw in section 7.3 that the adiabats for liquid Fe needed to obtain the geotherm can be derived from the first-principles free energy. The resulting geotherm [268] gives the temperature on the core side of the core–mantle boundary (CMB) in the range 3700–3950 K, in close agreement with other recent estimates [31]. This temperature happens to be very close to the estimated mantle solidus [292], so there may be partially molten regions at the base of the mantle. We note that the temperature range 3700–3950 K is substantially above the temperature of 2600 K estimated for the mantle side of the core–mantle boundary (see section 6.6). As we pointed out in section 2.1, this is expected, because of the thermal boundary layer at the base of the mantle.

10. Summary and prospects

We have tried to present a broad survey of the substantial contributions that first-principles modelling is now making to the understanding of materials in the deep interiors of the earth and the other planets. In early work of this kind, density-functional theory was used to make zero-temperature static calculations, but we have shown that it is now increasingly being used for high-temperature calculations, either by lattice dynamics in the quasi-harmonic approximation or by first-principles molecular dynamics simulation. As we have seen, the latter approach is indispensable for the many planetary science problems that concern liquids, such as the thermodynamics of the Earth's core and the electronic properties of hydrogen and water in the interiors of the giant planets. First-principles modelling has come into widespread use only in the past 15 years, and the methods are still evolving. We conclude this review by assessing the current accuracy of the methods, summarizing the scientific progress that has already been achieved and suggesting where future progress can be expected.

An important theme running through this review has been the continual interplay between first-principles modelling and experiment. Since first-principles methods necessarily involve approximations, the most important of which is the approximate form of the exchange-correlation functional, the importance of validating the calculations against experimental data is obvious. The most direct evidence for the reliability of DFT methods at the high pressures relevant to planetary interiors comes from cold compression curves, i.e. the $p(V)$ relation at room temperature. We have seen that cold compression curves of geologically important materials such as MgO, $\mathrm{MgSiO_3}$ and hcp Fe at pressures in the megabar region are predicted with excellent accuracy. In the case of hcp Fe, at a pressure of 300 GPa, which is near that of the inner-core/outer-core boundary, the error in the volume per atom is only $\sim 1\%$. Added weight is given to these comparisons by the facts that no adjustable parameters enter the calculations and that different DFT implementations (pseudopotential, projector augmented wave, augmented plane wave, etc) produce essentially identical results. We have also shown some of the abundant evidence for the accuracy of DFT for phonon frequencies at ambient pressures, the departures from experimental frequencies for materials such as MgO and Fe being generally no more than $\sim 3\%$. The existing rather sparse comparisons with high-pressure phonon densities of states, currently up to pressures $\sim 150\,\mathrm{GPa}$, give valuable evidence that

DFT maintains its accuracy at these pressures. Since it is exactly these phonon frequencies that enter free-energy calculations in the quasi-harmonic approximation, there is good reason to expect reliable thermodynamic results at high p and T. Further supporting evidence for this comes from the accurate predictions of $p(V)$ on the shock Hugoniot that we have illustrated.

We have seen that the situation with melting curves is not yet completely satisfactory. Particular controversy has surrounded the melting curve of Fe, which is crucial for constraining the temperature distribution in the Earth's core. In the past, there were major differences between the melting curves determined by static compression and by shock experiments. However, we have seen that the more recent static compression data is much more consistent with the shock data and also with the majority of first-principles results. This is an area where the combined efforts of experimentalists and modellers are leading to major progress. However, in other cases (we have noted particularly the case of MgO), there remains a strong need to reconcile the theoretical and experimental results.

The work that we have reviewed shows that first-principles modelling has already made substantial contributions to our understanding of the Earth's interior. A striking example of this is the role of first-principles calculations in the discovery of the post-perovskite phase of $MgSiO_3$ and the use of the calculations to explore its properties. We have seen how the calculated elastic properties of post-perovskite and the calculated Clapeyron slope of the perovskite/post-perovskite phase boundary explain the previously mysterious properties of the D'' zone. First-principles work has also contributed significantly to our understanding of the temperature distribution within the Earth and the chemical composition of the core. We have shown how these advances were made possible by the development of first-principles techniques for calculating melting properties and chemical potentials in solid and liquid phases. These developments have led to completely new ways of placing constraints on the temperature at the inner-core/outer-core boundary and on the fractions of light impurities in the inner and outer core. An important conclusion is that oxygen seems to be an essential impurity in the core, without which the seismically observed density difference between inner and outer core cannot be explained.

For the giant planets, the interplay between first-principles modelling and experiment has been particularly fruitful. A central concern of much of the modelling work on hydrogen has been the demonstration that the DFT simulations really do reproduce the available data. We have seen that the simulations reproduce at least semi-quantitatively the location of the insulator–metal transition in the p–T plane and particularly the strong dependence of the transition line on temperature. The values of the electrical conductivity measured in shock experiments are also approximately reproduced. This gives confidence in the predictions of modelling at thermodynamic states that cannot be probed experimentally. The simulations have provided firm evidence for the existence of a first-order fluid–fluid transition associated with the insulator–metal transition but so far only in a region that is rather far from the conditions inside Jupiter and Saturn. It seems that there is much more to be done in mapping this transition line and particularly in determining where it ends. In the case of Uranus and Neptune, where the problems concern the properties of water, methane and ammonia under extreme conditions, first-principles modelling has given completely new insights. In particular, it has given much firmer quantitative guides to the thermodynamic regions in which ionic and electronic electrical conductivity should be expected. It has also played an important role in mapping the phase diagram of water, methane and ammonia over a wide range of thermodynamic conditions.

In spite of the large amount of first-principles work already reported, there remains ample scope both for investigating new physical quantities and for improving the calculation of those already studied. For example, rather little attention has been paid to the calculation of electrical and thermal conductivities at high pressures and temperatures, apart from old work using empirical potentials. The calculation of the electrical conductivity of hydrogen using the Kubo–Greenwood approximation has been described in this review. Similar calculations on liquid Fe under Earth's core conditions would be valuable because the value of the conductivity is crucial in all geodynamo models but up to now has simply been estimated by intelligent guesswork. The existing comparisons of electrical conductivity with shock compression values for hydrogen show only qualitative agreement, and there is a clear need for improvement in this area.

Finally, we note some current developments which look likely to improve still further the accuracy of first-principles calculations. The first is the rapid improvement of quantum Monte Carlo methods for calculating the total energy of condensed matter. We have described how these have already played a significant role in studying the metallization of hydrogen (section 6.2). Recent work has shown the feasibility of applying them to oxide materials [8,12], and there are hopes that they may be able to overcome the deficiencies of DFT for highly correlated materials such as transition-metal oxides. An alternative route to the treatment of highly correlated materials is dynamical mean field theory. Here again, recent progress on transition–metal oxides is encouraging [230]. A third area of technical progress is the calculation of total energies by computational many-body theory (see e.g. [229]). This approach provides a way of accounting for the dependence of exchange-correlation energy on local electronic structure, and is known to be superior to existing DFT methods for some materials. The rapid pace of technical advances means that we can look forward to even more accurate and reliable quantum-based modelling methods in the future.

Acknowledgments

The authors are grateful to NERC and EPSRC for national supercomputer resources. DA, JB and LV acknowledge support from the Royal Society, and DA acknowledges support from the Leverhulme Trust. LV thanks D Fortes, I Wood and D Dobson for useful discussions.

References

[1] Ahrens T J, Holland K G and Chen C Q 2002 Phase diagram of iron, revised core temperatures *Geophys. Res. Lett.* **29** 1150

[2] Alder B J, Frankel S P and Lewinson V A 1955 Radial distribution function calculated by the Monte Carlo method for a hard-sphere fluid *J. Chem. Phys.* **23** 417–19

[3] Alder B J and Wainwright T E 1959 Studies in molecular dynamics: I. General method *J. Chem. Phys.* **31** 459–66

[4] Alfè D 1998 PHON: a code to calculate phonons using the small displacement method http://chianti.geol.ucl.ac.uk/~dario/

[5] Alfè D 1999 Ab initio molecular dynamics, a simple algorithm for charge extrapolation *Comput. Phys. Commun.* **118** 31–3

[6] Alfè D 2003 First-principles simulations of direct coexistence of solid and liquid aluminum *Phys. Rev.* B **68** 064423

[7] Alfè D 2005 The melting curve of MgO from first-principles simulations *Phys. Rev. Lett.* **94** 235701 1–4

[8] Alfè D, Alfredsson M, Brodholt J, Gillan M J, Towler M D and Needs R J 2005 Quantum Monte Carlo calculations of the structural properties and B1-B2 phase transition of MgO

[9] Alfè D and Gillan M J 1998 First-principles calculation of transport coefficients. *Phys. Rev. Lett.* **81** 5161–4

[10] Alfè D and Gillan M J 1998 First-principles simulations of liquid Fe-S under Earth's core conditions *Phys. Rev.* B **58** 8248–56

[11] Alfè D and Gillan M J 2003 Electron correlation and the phase diagram of silicon *Phys. Rev.* B **68** 205212 1–5

[12] Alfè D and Gillan M J 2005 The Schottky defect formation energy in MgO calculated by diffusion Monte Carlo *Phys. Rev.* B **71** 220101 1–3

[13] Alfè D, Gillan M J and Price G D 1999 Melting curve of iron at Earth's core pressures from *ab initio* calculations *Nature* **401** 462–4

[14] Alfè D, Gillan M J and Price G D 2000 Constraints on the composition of the Earth's core from *ab initio* calculations *Nature* **405** 172–5

[15] Alfè D, Gillan M J and Price G D 2002 *Ab initio* chemical potentials of solid and liquid solutions and the chemistry of the Earth's core *J. Chem. Phys.* **116** 7127–35

[16] Alfè D, Gillan M J and Price G D 2002 Composition and temperature of the Earth's core constrained by combining *ab initio* calculations and seismic data *Earth Planet. Sci. Lett.* **195** 91–8

[17] Alfè D, Gillan M J and Price G D 2003 Thermodynamics from first principles: temperature and composition of the Earth's core *Mineral. Mag.* **67** 113–25

[18] Alfè D, Kresse G and Gillan M J 2000 Structure and dynamics of liquid iron under Earth's core conditions *Phys. Rev.* B **61** 132–42

[19] Alfè D, Price G D and Gillan M J 1999 Oxygen in the Earth's core: a first-principles study *Phys. Earth Planet. Inter.* **110** 191–210

[20] Alfè D, Price G D and Gillan M J 2000 Thermodynamic stability of Fe/O solid solution at inner-core conditions *Geophys. Res. Lett.* **27** 2417–20

[21] Alfè D, Price G D and Gillan M J 2001 Thermodynamics of hexagonal-close-packed iron under Earth's core conditions *Phys. Rev.* B **64** 045123 1–16

[22] Alfè D, Price G D and Gillan M J 2002 Complementary approaches to the *ab initio* calculation of melting properties *J. Chem. Phys.* **116** 6170

[23] Alfè D, Price G D and Gillan M J 2002 Iron under Earth's core conditions: liquid state thermodynamics and high pressure melting curve from ab-initio calculations *Phys. Rev.* B **65** 165118

[24] Alfè D, Vočadlo L, Price G D and Gillan M J 2004 Melting curve of materials: theory versus experiment *J. Phys. Condens. Matter* **16** S973

[25] Allègre C J, Poirier J-P, Humler E and Hofmann A W 1995 The chemical composition of the earth *Planet. Sci. Lett.* **134** 515–26

[26] Allen M P and Tildesley D J 1987 *Computer Simulation of Liquids* (Oxford: Oxford University Press)

[27] Ancilotto F, Chiarotti G L, Scandolo S and Tosatti E 1997 Dissociation of methane into hydrocarbons at extreme (planetary) pressure and temperature *Science* **275** 1288–90

[28] Andersen H C 1980 Molecular dynamics at constant pressure and/or temperature *J. Chem. Phys.* **72** 2384–93

[29] Anderson D L 1989 *Theory of the Earth* (Boston: Blackwell Scientific)

[30] Anderson O L 1995 *Equations of State of Solids for Geophysics and Ceramic Science* (Oxford: Oxford University Press)

[31] Anderson O L 2003 The three dimensional phase diagram of iron *Geodynamics Series* vol 31 ed V Dehant *et al* (Washnington, DC: American Geophysical Union)

[32] Andrault D, Fiquet G, Charpin T and Le Bihan T 2000 Structure analysis and stability field of iron at high pressure and temperature *Am. Mineral.* **85** 364–71

[33] Andrault D, Fiquet G, Kunz M, Visocekas F and Hausermann D 1997 The orthorhombic structure of iron: an in situ study at high temperature and high pressure *Science* **278** 831–4

[34] Aoki K, Yamawaki H, Sakashita M and Fujihisa H 1996 Infrared absorption study of the hydrogen-bond symmetrization in ice to 110 GPa *Phys. Rev.* B **54** 15673–7

[35] Ashcroft N W 2004 Bridgman's high-pressure atomic destructibility and its growing legacy of ordered states *J. Phys.: Condens. Phys.* **16** S945–52

[36] Atkins P W 1994 *Physical Chemistry* (Oxford: Oxford University Press)

[37] Baroni S, de Gironcoli S, Dal Corso A and Gianozzi P 2001 Phonons and related crystal properties from density-functional perturbation theory *Rev. Mod. Phys.* **73** 515–62

[38] Baroni S, Gianozzi P and Testa A 1987 Green's-function approach to linear response in solids *Phys. Rev. Lett.* **58** 1861–4

[39] Barron T H K and Klein M L 1965 Second-order elastic constants of a solid under stress *Proc. Phys. Soc.* **85** 523–32

[40] Bell D R and Rossman G R 1992 Water in the Earth's mantle—the role of nominally anhydrous minerals *Science* **255** 1391–7

[41] Belonoshko A B, Ahuja R, Eriksson O and Johansson B 2000 Quasi-*ab initio* molecular dynamic study of Cu melting *Phys. Rev.* B **61** 3838

[42] Belonoshko A B, Ahuja R and Johansson B 2000 Quasi-*ab initio* molecular dynamics study of Fe melting *Phys. Rev. Lett.* **84** 3638

[43] Belonoshko A B, Ahuja R and Johansson B 2003 Stability of body-centred-cubic phase of iron in the Earth's inner core *Nature* **424** 1032–4

[44] Benedetti L R, Nguyen J H, Caldwell W A, Liu H, Kruger M and Jeanloz R 1999 Dissociation of methane at high pressures and temperatures: diamond formation in giant planet interiors? *Science* **286** 100–2

[45] Benoit M, Bernasconi M, Focher P and Parrinello M 1996 New high-pressure phase of ice *Phys. Rev. Lett.* **76** 2934–7

[46] Benoit M, Marx D and Parrinello M 1998 Tunnelling and zero-point motion in high-pressure ice *Nature* **392** 258–61

[47] Bernasconi M, Silvistrelli P L and Parrinello M 1998 *Ab initio* infrared absorption study of the hydrogen-bond symmetrization in ice *Phys. Rev. Lett.* **81** 1235–8

[48] Blöchl P E 1994 Projector augmented-wave method *Phys. Rev.* B **50** 17953–79

[49] Boehler R 1993 Temperatures in the Earth's core from melting point measurements of iron at high static pressures *Nature* **363** 534

[50] Bonev S A, Schwegler E, Ogitsu T and Galli G 2004 A quantum fluid of metallic hydrogen suggested by first-principles calculations *Nature* **431** 669–72

[51] Born M and Huang K 1954 *Dynamical Theory of Crystal Lattices* (Oxford: Oxford University Press)

[52] Braithwaite J S, Wright K and Catlow C R A 2003 A theoretical study of the energetics and IR frequencies of hydroxyl defects in forsterite *J. Geophys. Res. Solid Earth* **108** 2284

[53] Brodholt J P 1997 *Ab initio* calculations on point defects in forsterite (Mg_2SiO_4) and implications for diffusion and creep *Am. Mineral.* **82** 1049–53

[54] Brodholt J P 2000 Pressure induced changes in the compression mechanism of aluminous perovskite in the earth's mantle *Nature* **407** 620–2

[55] Brodholt J P and Refson K 2000 An *ab initio* study of hydrogen in forsterite and a possible mechanism for hydrolytic weakening *J. Geophys. Res.* **105** 18977

[56] Broughton J Q and Li X P 1987 Phase diagram of silicon by molecular dynamics *Phys. Rev.* B **35** 9120–7

[57] Brown J M and McQueen R G 1986 Phase transitions, Grüneisen parameter and elasticity for shocked iron between 77 GPa and 400 GPa *J. Geophys. Res.* **91** 7485–94

[58] Car R and Parrinello M 1985 Unified approach for molecular dynamics and density-functional theory *Phys. Rev. Lett.* **55** 2471–4274

[59] Catlow C R A (ed) 2003 *Computational Materials Science* vol 187 *NATO Science Series: Computer and Systems Sciences* (Amsterdam: IOS Press)

[60] Catlow C R A and Norgett M J 1973 Shell model calculations of the energies of formation of point defects in alkaline earth fluorides *J. Phys. C: Solid State Phys.* **6** 1325–39

[61] Cavazzoni C, Chiarotti G L, Scandolo S, Tosatti E, Bernasconi M and Parrinello M 1999 Superionic and metallic states of water and ammonia at giant planet pressures *Science* **283** 44–6

[62] Chabrier G, Saumon D, Hubbard W B and Lunine J I 1992 The molecular-metallic transition of hydrogen and the structure of Jupiter and Saturn *Astrophys. J.* **391** 817–26

[63] Chandler D 1987 *Introduction to Modern Statistical Mechanics* (Oxford: Oxford University Press)

[64] Chang K J and Cohen M L 1984 High-pressure behavior of MgO: structural and electronic properties *Phys. Rev.* B **30** 4774–81

[65] Chau R, Mitchell A C, Minich R W and Nellis W J 2001 Electrical conductivity of water compressed dynamically to pressures of 70–180 GPa (0.7–1.8 Mbar) *J. Chem. Phys.* **114** 1361–5

[66] Cheung P S Y and Powles J G 1975 The properties of liquid nitrogen: a computer simulation *Molec. Phys.* **30** 921–36

[67] Cohen M L and Heine V 1970 *Solid State Phys.* ed H Ehrenreich *et al* (New York: Academic Press) p 37

[68] Cohen R E 1987 Elasticity and equation of state of $MgSiO_3$ perovskite *Geophys. Res. Lett.* **14** 1053–6

[69] Cohen R E and Gong Z 1994 Melting and melt structure of MgO at high pressures *Phys. Rev.* B **50** 12301–11

[70] Cohen R E, Gramsch S, Steinle-Neumann G and Stixrude L 2002 Importance of magnetism in phase stability, equations of state, and elasticity *Proc. Int. School of Physics Enrico Fermi, Varenna, Italy, Volume CXLVII, High Pressure Phenomena* ed R J Hemley *et al* (Washington DC: IOS Press) pp 215–38

[71] Cohen R E, Mehl M J and Papaconstantopoulos D A 1994 Tight-binding total-energy method for transition and noble metals *Phys. Rev.* B **50** 14694

[72] Cohen R E, Stixrude L and Wasserman E 1997 Tight-binding computation of elastic anisotropy of Fe, Xe and Si under compression *Phys. Rev.* B **56** 8575–89

[73] Collins G W, Da Silva L B, Celliers P, Gold D M, Foord M E, Wallace R J, Ng A, Weber S V, Budil K S and Cauble R 1998 Measurements of the equation of state of deuterium at at the fluid insulator-metal transition *Science* **281** 1178–81

[74] Collins L A, Bickham S R, Kress J D, Mazevet S, Lenosky T J, Troullier N J and Windl W 2001 2001 Dynamical and optical properties of warm dense hydrogen *Phys. Rev.* B **63** 184110

[75] D'Arco Ph, Sandrone G, Dovesi R, Aprà E and Saunders V R 1994 A quantum-mechanical study of the relative stability under pressure of $MgSiO_3$-ilmenite, $MgSiO_3$-perovskite and MgO-periclase + SiO_2-stishovite assemblage *Phys. Chem. Miner.* **21** 285–93

[76] D'Arco Ph, Sandrone G, Dovesi R, Orlando R and Saunders V R 1993 A quantum mechanical study of the perovskite structure type of $MgSiO_3$ *Phys. Chem. Miner.* **20** 407–14

[77] Daw M and Baskes M 1983 Semi-empirical quantum mechanical calculation of hydrogen embrittlement in metals *Phys. Rev. Lett.* **50** 1285–8

[78] Daw M, Foiles S M and Baskes M 1993 The embedded-atom method—a review of theory and applications *Mater. Sci. Rep.* **9** 251–310

[79] De Pater I and Lissauer J J 2001 *Planetary Sciences* (Cambridge: Cambridge University Press)

[80] De Vita A, Gillan M J, Lin J S, Payne M C, Štich I and Clarke L J 1992 Defect energetics in MgO treated by first-principles methods *Phys. Rev.* B **46** 12964–73

[81] de Wijs G A, Kresse G and Gillan M J 1998 First-order phase transitions by first-principles free energy calculations: the melting of Al *Phys. Rev.* B **57** 8223–34

[82] de Wijs G A, Kresse G, Vočadlo L, Dobson D, Alfè D, Gillan M J and Price G D 1998 The viscosity of liquid iron at the physical conditions of the earth's core *Nature* **392** 805–7

[83] Demontis P, Lesar R and Klein M L 1988 New high-pressure phases of ice *Phys. Rev. Lett.* **60** 2284–7

[84] Dobson D and Brodholt J 2005 Subducted banded iron formations as a source of ultralow-velocity zones at the core-mantle boundary *Nature* **434** 371–4

[85] Dobson D, Vočadlo L and Wood I G 2002 A new high-pressure phase of FeSi *Am. Miner.* **87** 784–7

[86] Drummond N D and Ackland G J 2002 *Ab initio* quasiharmonic equations of state for dynamically stabilized soft-mode materials *Phys. Rev.* B **65** 184104

[87] Duffy T 2005 Synchrotron facilities and the study of the Earth's deep interior *Rep. Prog. Phys.* **68** 1811–59

[88] Duffy T H, Hemley R J and Mao H K 1995 Equation of state and shear strength at multimegabar pressures: magnesium oxide to 227 GPa *Phys. Rev. Lett.* **74** 1371–4

[89] Dziewonski A M and Anderson D L 1981 Preliminary reference earth model *Phys. Earth Planet. Inter.* **25** 297–356

[90] Ercolessi F and Adams J B 1994 Interatomic potentials from first principles: the force-matching method *Europhys. Lett.* **26** 583–8

[91] Errandonea D, Schwager B, Ditz R, Gessmann C, Boehler R and Ross M 2001 Systematics of transition-metal melting *Phys. Rev.* B **63** 132104

[92] Feynman R P 1939 Forces in molecules *Phys. Rev.* **56** 340

[93] Finnis M and Sinclair J 1984 A simple empirical n-body potential for transition metals *Phil. Mag.* A **50** 45–55

[94] Finnis M W 2003 *Interatomic Forces in Condensed Matter* (Oxford: Oxford University Press)

[95] Fiquet G, Dewaele A, Andrault D, Kunz M and Le Bihan T 2000 Thermoelastic properties and crystal structure of $MgSiO_3$ perovskite at lower mantle pressure and temperature conditions *Geophys. Res. Lett.* **27** 21–4

[96] Fortes A D 2004 Computational and experimental studies of solids in the ammonia-water system *PhD Thesis* University of London

[97] Fortes A D, Wood I G, Brodholt J P and Vočadlo L 2003 *Ab initio* simulation of the ice II structure *J. Chem. Phys.* **119** 4567–72

[98] Fortes A D, Wood I G, Brodholt J P and Vočadlo L 2003 The structure, ordereing and equation of state of ammonia dihydrate *Icarus* **162** 59–73

[99] Fortes A D, Wood I G, Knight K S, Brodholt J P, Alfedsson M, McGrady G S and Vočadlo L 2003 A high resolution neutron powder diffraction study of ammonia dihydrate phase I *J. Chem. Phys.* **119** 10806–13

[100] Foulkes W M C, Mitaš L, Needs R J and Rajagopal G 2001 Quantum Monte Carlo simulations of solids *Rev. Mod. Phys.* **73** 33–83

[101] Frenkel D and Smit B 2002 *Understanding Molecular Simulation* (New York: Academic)

[102] Galli G, Hood R Q, Hazi A U and Gygi F 2000 *Ab initio* simulations of compressed liquid deuterium *Phys. Rev.* B **61** 909–12

[103] Gannarelli C, Alfè D and Gillan M J 2005 The axial ratio of h.c.p. iron at the conditions of the earth's inner core *Phys. Earth Planet. Inter.* **152** 67–77

[104] Gannarelli C M S, Alfè D and Gillan M J 2003 The particle-in-cell model for *ab initio* thermodynamics: implications for the elastic anisotropy of the Earth's inner core *Phys. Earth Planet. Inter.* **139** 243–53

[105] Giannozzi P, de Gironcoli S, Pavone P and Baroni S 1991 *Ab initio* calculation of phonon dispersions in semiconductors *Phys. Rev.* B **43** 7231–42

[106] Gibson J B, Goland A N, Milgram M and Vineyard G H 1960 Dynamics of radiation damage *Phys. Rev.* **120** 1229–53

[107] Gillan M J 1997 The virtual matter laboratory *Contemp. Phys.* **38** 115

[108] Gillen K T, Douglass D C and Hoch J R 1972 Self-diffusion in liquid water to −31°C *J. Chem. Phys.* **57** 5117

[109] Goncharov A F, Goldman N, Fried L E, Crowhurst J C, Kuo I-F W, Mundy C J and Zaug J M 2005 Dynamic ionization of water under extreme conditions *Phys. Rev. Lett.* **94** 125508

[110] Goncharov A F, Struzhkin V V, Somayazulu M S, Hemley R J and Mao H K 1996 Compression of ice to 210 GPa: infrared evidence for a symmetric hydrogen-bonded phase *Science* **273** 218–20

[111] Gonze X 1997 First-principles responses of solids to atomic displacements and homogeneous electric fields: Implementation of a conjugate-gradient algorithm *Phys. Rev.* B **55** 10337–54

[112] Gonze X 2005 A brief introduction to the ABINIT software package *Z. Kristallogr.* **220** 558–62

[113] Gonze and Lee C 1997 Dynamical matrices, born effective charges dielectric permittivity tensors, and interatomic force constants from density-functional perturbation theory *Phys. Rev.* B **55** 10355–68

[114] Gonze X and Vigneron J-P 1989 Density-functional approach to nonlinear-response coefficients in solids *Phys. Rev.* B **39** 13120–8

[115] Goringe C M, Bowler D R and Hernández E 1997 Tight-binding modelling of materials *Rep. Prog. Phys.* **60** 1447–97

[116] Greenwood D A 1958 The Boltzmann equation in the theory of electrical conduction in metals *Proc. Phys. Soc.* **71** 585–96

[117] Gubbins D, Alfè D, Masters G, Price G D and Gillan M J 2004 Gross thermodynamics of 2-component core convection *Geophys. J. Int.* **157** 1407

[118] Gubbins D, Alfè D, Masters G, Price G D and Gillan M J 2003 Can the Earth's dynamo run on heat alone? *Geophys. J. Int.* **155** 609–22

[119] Hafner J and Heine V 1986 Theory of atomic interactions in (s,p)-bonded metals *J. Phys. F: Met. Phys.* **16** 1429–58

[120] Harrison W A 1980 *Electronic Structure and the Properties of Solids* (San Francisco: Freeman)

[121] Hart S R and Zindler A 1986 In search of a bulk earth composition *Chem. Geol.* **57** 247–67

[122] Heine V 1970 *Solid State Physics* ed H Ehrenreich *et al* (New York: Academic) p 1

[123] Hohenberg P and Kohn W 1964 Inhomogeneous electron gas *Phys. Rev.* **136** B864–71

[124] Holmes N C, Nellis W J, Graham W B and Walrafen G E 1985 Spontaneous Raman scattering from shocked water *Phys. Rev. Lett.* **55** 2433–6

[125] Holmes N C, Ross M and Nellis W J 1995 Temperatures measurements and dissociation of shock compressed liquid deuterium and hydrogen *Phys. Rev.* B **52** 15835–45

[126] Hood R Q and Galli G 2004 Insulator to metal transition in fluid deuterium *J. Chem. Phys.* **120** 5691–4

[127] Hoover W G 1985 Canonical dynamics: equilibrium phase-space distributions *Phys. Rev.* A **31** 1695–7

[128] Hubbard W B, Nellis W J, Mitchell A C, Holmes N C, Limaye S S and McCandless P C 1991 Interior structure of Neptune—comparison with Uranus *Science* **253** 648–51

[129] Iitaka T, Hirose K, Kawamura K and Murakami M 2004 The elasticity of the $MgSiO_3$ post-perovskite phase in the Earth's lower mantle *Nature* **430** 442–5

[130] Ito E and Katsura T 1989 A temperature profile of the mantle transition zone *Geophys. Res. Lett.* **16** 425–8

[131] Jacobsen K W, Norskov J K and Puska M J 1987 Interatomic interactions in the effective medium theory *Phys. Rev.* B **35** 7423–42

[132] Jaffe J E, Snyder J A, Lin Z and Hess A C 2000 LDA and GGA calculations for high-pressure phase transitions in ZnO and MgO *Phys. Rev.* B **62** 1660–5

[133] Janney D E and Banfield J F 1998 Distribution of cations and vacancies and the structure of defects in oxidized intermediate olivine by atomic-resolution TEM and image simulation *Am. Mineral.* **83** 799–810

[134] Jaoul O, Bertran-Alvarez Y, Liebermann R C and Price G D 1995 Fe-Mg interdiffusion in olivine up to 9 GPa at $T = 600$–$900\,^{\circ}$C experimental data and comparison with defect calculations *Phys. Earth Planet. Inter.* **89** 199–218

[135] Jorgensen W, Chandrasekhar J, Madura J D, Impey R W and Klein M L 1983 Comparison of simple potential functions for simulating liquid water *J. Chem. Phys.* **79** 926–41

[136] Karki B B, Stixrude L, Clark S J, Warren M C, Ackland G J and Crain J 1997 Elastic properties of orthorhombic $MgSiO_3$ at lower mantle pressures *Am. Mineral.* **82** 635–8

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

[138] Karki B B, Warren M C, Stixrude L, Ackland G J and Crain J 1997 *Ab initio* studies of high-pressure structural transformations in silica *Phys. Rev.* B **55** 3465–71

[139] Karki B B and Wentzcovitch R M 2002 First-principles lattice dynamics and thermoelasticity of $MgSiO_3$ ilmenite at high pressure *J. Geophys. Res. Solid Earth* **107** 1–10

[140] Karki B B, Wentzcovitch R M, de Gironcoli S and Baroni S 1999 First-principles determination of elastic anisotropy and wave velocities of MgO at lower mantle conditions *Science* **286** 1705–7

[141] Karki B B, Wentzcovitch R M, de Gironcoli S and Baroni S 2000 *Ab initio* lattice dynamics of $MgSiO_3$ perovskite at high pressure *Phys. Rev.* B **62** 14750–6

[142] Karki B B, Wentzcovitch R M, de Gironcoli S and Baroni S 2000 High pressure lattice dynamics and thermoelasticity of MgO *Phys. Rev.* B **61** 8793–800

[143] Kiefer B, Stixrude L and Wentzcovitch R M 2002 Elasticity of (Mg,Fe)SiO_3 perovskite at high pressures *Geophys. Res. Lett.* **29** 014683

[144] Kingma K J, Cohen R E, Hemley R J and Mao H K 1995 Transformation of stishovite to a denser phase at lower-mantle pressures *Nature* **374** 243–5

[145] Kitamura H and Ichimaru S 1998 Metal-insulator transitions in dense hydrogen: equations of state, phase diagrams and interpretations of shock-compression experiments *J. Phys. Soc. Japan* **67** 950–63

[146] Kitamura H, Tsuneyuki S, Ogitsu T and Miyake T 2000 Quantum distribution of protons in solid molecular hydrogen at megabar pressures *Nature* **404** 259–62

[147] Knudson M D, Hanson D L, Bailey J E, Lemke R W, Hall C A, Deeney C and Assay J R 2003 Equation of state measurements in liquid deuterium to 100 GPa *J. Phys. A: Math. Gen.* **36** 6149–58

[148] Kobayashi H, Kamikura T, Alfè D, Sturhahn W and Zhao J 2004 Phonon density of states and compression behavior in FeS under pressure *Phys. Rev. Lett.* **93** 195503

[149] Kohanoff J 2006 *Electronic Structure Calculations for Solids and Molecules: Theory and Computational Methods*

[150] Kohanoff J, Scandolo S, Chiarotti G L and Tosatti E 1997 Solid molecular hydrogen: the broken symmetry phase *Phys. Rev. Lett.* **78** 2783–6

[151] Kohanoff J, Scandolo S, de Gironcoli S and Tosatti E 1999 Dipole-quadrupole interactions and the nature of phase II of compressed hydrogen *Phys. Rev. Lett.* **83** 4097–100

[152] Kohn W and Sham L J 1965 Self-consistent equations including exchange and correlation effects *Phys. Rev.* **140** A1133–8

[153] Kresse G and Furthmüller J 1996 Efficiency of *ab initio* total energy calculations for metals and semiconductors using a plane-wave basis set *Comput. Mater. Sci.* **6** 15

[154] Kresse G and Furthmüller J 1996 Efficient iterative schemes for *ab initio* total-energy calculations using a plane-wave basis set *Phys. Rev.* B **54** 11169–86

[155] Kresse G, Furthmüller J and Hafner J 1995 *Ab initio* force constant approach to phonon dispersion relations of diamond and graphite *Europhys. Lett.* **32** 729–34

[156] Kresse G and Joubert D 1999 From ultrasoft pseudopotentials to the projector augmented-wave method *Phys. Rev.* B **59** 1758–75

[157] Kubo R 1957 Statistical-mechanical theory of irreversible processes: 1. General theory and simple applications to magnetic and conduction problems *J. Phys. Soc. Japan.* **12** 570–86

[158] Kuwayama Y, Hirose K, Sata N and Ohishi Y 2005 The pyrite-type high-pressure form of silica *Science* **309** 923–5

[159] Laio A, Bernard S, Chiarotti G L, Scandolo S and Tosatti E 2000 Physics of iron at earth's core conditions *Science* **287** 1027

[160] Laio A and Parrinello M 2002 Escaping free-energy minima *Proc. Natl Acad. Sci. USA* **99** 12562–6

[161] Lenosky T J, Bickham S R, Kress J D and Collins L A 2000 Density-functional calculation of the Hugoniot of shocked liquid deuterium *Phys. Rev.* B **61** 1–4

[162] Li L, Brodholt J, Stackhouse S, Weidner D J, Alfresson M and Price G D 2005 Electronic spin state of ferric iron in Al-bearing perovskite in the lower mantle *Geophys. Res. Lett.* **32** L17307

[163] Lin J-F, Heinz D L, Campbell A J, Devine J M and Shen G 2002 Iron-silicon alloy in Earth's core? *Science* **295** 313–5

[164] Lorenz R D and Shandera S E 2001 Physical properties of ammonia rich ice: application to Titan *Geophys. Res. Lett.* **28** 215–8

[165] Loubeyre P, Le Toullec R, Häusermann D, Hanfland M, HemleyR J, Mao H K and Finger L W 1996 X-ray diffraction and equation of state of hydrogen at megabar pressures *Nature* **383** 702–4

[166] Loubeyre P, Occelli F and Le Toullec R 2002 Optical studies of solid hydrogen to 320 GPa and evidence for black hydrogen *Nature* **416** 613–17

[167] Ma Y, Somayazulu G, Shen G, Mao H K, Shu J and Hemley R J 2004 *In situ* x-ray diffraction studies of iron to Earth's core conditions *Phys. Earth Planet. Inter.* **143–144** 455

[168] Mao H K and Hemley R J 1998 New windows on the Earth's deep interior *Rev. Mineral.* **37** 1–32

[169] Mao H K *et al* 2001 Phonon density of states of iron up to 153 GPa. *Science* **292** 914–16

[170] Martin R M 2004 *Electronic Structure* (Cambridge: Cambridge University Press)

[171] Marton F C and Cohen R E 1994 Prediction of a high-pressure phase transition in Al_2O_3 *Am. Mineral.* **79** 789–92

[172] Martoňák R, Laio A and Parrinello M 2003 Predicting crystal structures: the Parrinello–Rahman method revisited *Phys. Rev. Lett.* **90** 075503

[173] Masters T G and Shearer P M 1990 Summary of seismological constraints on the structure of the Earth's core *J. Geophys. Res. Solid Earth Planet* **95** 21691–5

[174] Masters T G and Gubbins D 2003 On the resolution of density within the Earth *Phys. Earth Planet. Inter.* **140** 159–67

[175] McConnell J D C, Lin J S and Heine V 1995 The solubility of [4H](Si) defects in alpha-quartz and their role in the formation of molecular water and related weakening on heating *Phys. Chem. Miner.* **22** 357–66

[176] McDonough W F and Sun S S 1995 The composition of the earth *Chem. Geol.* **120** 223–53

[177] Mehl M J, Cohen R E and Krakauer H 1988 Linearized augmented plane-wave electronic structure calculations for MgO and CaO *J. Geophys. Res.* **93** 8009

[178] Mermin N D 1965 Thermal properties of the inhomogeneous electron gas *Phys. Rev.* **137** A1441–3

[179] Metropolis N, Rosenbluth A W, Rosenbluth M N, Teller A H and Teller E 1953 Equation of state by fast computing machines *J. Chem. Phys.* **21** 1087–92

[180] Monkhorst H J and Pack J D 1976 Special points for Brillouin-zone integrations *Phys. Rev.* B **13** 5188–92

[181] Moriarty J A 1986 *Shock Waves in Condensed Matter* ed Y M Gupta (New York: Plenum) pp 46–9

[182] Moriarty J A 1990 Analytic representation of multi-ion interatomic potentials in transition metals *Phys. Rev.* B **42** 1609–28

[183] Moriarty J A 1994 Angular forces and melting in bcc transition metals—a case study of molybdenum *Phys. Rev.* B **49** 12431–45

[184] Moriarty J A, Belak J F, Rudd R E, Söderlind P, Streitz F H and Yang L H 2002 Quantum-based atomistic simulation of materials properties in transition metals *J. Phys.: Condens. Matter* **14** 2825–57

[185] Moriarty J A, Young D A and Ross M 1984 Theoretical study of the aluminum melting curve to very high pressure *Phys. Rev.* B **30** 578–88

[186] Morris J R, Wang C Z, Ho K M and Chan C T 1994 Melting line of aluminum from simulations of coexisting phases *Phys. Rev.* B **49** 3109

[187] Murakami M, Hirose K, Kawamura K, Sata N and Ohishi Y 2004 Post-perovskite phase transition in $MgSiO_3$ *Science* **304** 855–8

[188] Narayana C, Luo H, Orloff J and Ruoff A L 1998 Solid hydrogen at 342 GPa: no evidence for an alkali metal *Nature* **393** 46–9

[189] Nellis W J, Ross M and Holmes N C 1995 Temperature measurements of shock compressed liquid hydrogen: implications for the interior of Jupiter *Science* **269** 1249–52

[190] Nelmes R J and Loveday J S 1998 Diffraction studies of ammonia hydrates *Experimental Report* ISIS

[191] Nelmes R J and Loveday J S 1999 The ambient pressure structure of ammonia dihydrate *Experimental Report* ISIS

[192] Nelmes R J, Loveday J S and Guthrie M 1999 Structural changes under pressure in ammonia dihydrate and ammonia hemihydrate *Experimental Report* ISIS

[193] Nguyen J H and Holmes N C 2004 Melting of iron at the physical conditions of the Earth's core *Nature* **427** 339–42

[194] Nielsen O H and Martin R M 1985 Quantum-mechanical theory of stress and force *Phys. Rev.* B **32** 3780–91

[195] Nielsen O H and Martin R M 1985 Stresses in semiconductors: *ab initio* calculations on Si, Ge and GaAs *Phys. Rev.* B **32** 3792–805

[196] Nosé S 1984 A molecular dynamics method for simulations in the canonical ensemble *Mol. Phys.* **52** 255–68

[197] Nye J F 1998 *Physical Properties of Crystals. Their Representation by Tensors and Matrices* (Oxford: Oxford University Press)

[198] Oganov A R 2002 Computer simulation studies of minerals *PhD Thesis* (University College, London)

[199] Oganov A R, Brodholt J and Price G D 2002 Ab initio theory of phase transitions and thermoelasticity of minerals *EMU Notes Mineral.* **4** 83–170

[200] Oganov A R, Brodholt J P and Price G D 2001 *Ab initio* elasticity and thermal equation of state of $MgSiO_3$ perovskite *Earth Planet. Sci. Lett.* **184** 555–60

[201] Oganov A R, Brodholt J P and Price G D 2001 The elastic constants of $MgSiO_3$ perovskite at pressures and temperatures of the earth's mantle *Nature* **411** 934–7

[202] Oganov A R and Dogokupets P I 2003 All-electron and pseudopotential study of MgO: equation of state, anharmonicity and stability *Phys. Rev.* B **67** 224110

[203] Oganov A R, Gillan M J and Price G D 2003 *Ab initio* lattice dynamics and the structural stability of MgO *J. Chem. Phys.* **118** 10174–84

[204] Oganov A R, Gillan M J and Price G D 2005 Structural stability of silica at high pressures and temperatures *Phys. Rev.* B **71** 064104

[205] Oganov A R, Glass C W and Ono S 2006 High-pressure phases of $CaCO_3$: crystal structure prediction and experiment *Earth Planet. Sci. Lett.* **241** 95–103

[206] Oganov A R, Martoňák R, Laio A, Raiteri P and Parrinello M 2005 Anisotropy of Earth's D″ layer and stacking faults in the $MgSiO_3$ post-perovskite phase *Nature* **438** 1142–4

[207] Oganov A R and Ono S 2004 Theoretical and experimental evidence for a post-perovskite phase of $MgSiO_3$ in earth's D″ layer *Nature* **430** 445–7

[208] Oganov A R and Ono S 2005 The high-pressure phase of alumina and implications for the Earth's D″ layer *Proc. Natl Acad. Sci.* **102** 10828–31

[209] Ogitsu T, Schwegler E, Gygi F and Galli G 2003 Melting of lithium hydride under pressure *Phys. Rev. Lett.* **91** 175502

[210] Ono S and Oganov A R 2005 *In situ* observations of phase transition between perovskite and $CaIrO_3$-type phase in $MgSiO_3$ and pyrolytic mantle composition *Earth Planet. Sci. Lett.* **236** 914–32

[211] Panning M and Romanowicz B 2004 Inference on flow at the base of the earth's mantle based on seismic anisotropy *Science* **303** 351–3

[212] Papaconstantopoulos D A and Mehl M J 2003 The Slater–Koster tight-binding model: a computationally efficient and accurate approach *J. Phys. Condens. Matter* **15** R413–40

[213] Parr R G and Yang W 1989 *Density-Functional Theory of Atoms and Molecules* (Oxford: Oxford University Press)

[214] Parrinello M and Rahman A 1980 Crystal structure and pair potentials: A molecular-dynamics study *Phys. Rev. Lett.* **45** 1196–9

[215] Perdew J P, Burke K and Ernzerhof M 1996 Generalized gradient approximation made simple *Phys. Rev. Lett.* **77** 3865–8

[216] Perdew J P, Chevary J A, Vosko S H, Jackson K A, Pederson M R, Singh D J and Fiolhais C 1992 Atoms, molecules, solids and surfaces—applications of the generalized gradient approximation for exchange and correlation *Phys. Rev.* B **46** 6671

[217] Perdew J P and Wang Y 1992 Accurate and simple analytic representation of the electron-gas correlation energy *Phys. Rev.* B **45** 13244–9

[218] Pfaffenzeller O, Hohl D and Ballone P 1995 Miscibility of hydrogen and helium under astrophysical conditions *Phys. Rev. Lett.* **74** 2599–602

[219] Poirier J P 1988 Transport properties of liquid metals and viscosity of the Earth's core *Geophys. J.* **92** 99–105

[220] Poirier J P 2000 *Introduction to the Physics of the Earth's Interior* (Cambridge: Cambridge University Press)

[221] Price G D, Parker S C and Leslie M 1987 The lattice dynamics and thermodynamics of the Mg_2SiO_4 polymorphs *Phys. Chem. Miner.* **15** 181–90

[222] Pulay P 1969 *Ab initio* calculation of force constants and equilibrium geometries in polyatomic molecules. I. Theory *Mol. Phys.* **17** 197–204

[223] Rahman A and Stillinger F H 1971 Molecular dynamics study of liquid water *J. Chem. Phys.* **55** 3336–59

[224] Rapaport D C 2004 *The Art of Molecular Dynamics Simulation* (Cambridge: Cambridge University Press)

[225] Ren P and Ponder J W 2004 Temperature and pressure dependence of the AMOEBA water model *J. Phys. Chem.* **108** 13427–37

[226] Ross M 1981 The ice layer in Uranus and Neptune—diamonds in the sky *Nature* **292** 435–6

[227] Ross M and Ree F H 1980 Repulsive forces of simple molecules and mixtures at high density and temperature *J. Chem. Phys.* **73** 6146–52

[228] Ross N L and Hazen R M 1989 Single-crystal x-ray-diffraction study of $MgSiO_3$ perovskite from 77 K to 400 K *Phys. Chem. Miner.* **16** 415–20

[229] Sanchez-Friera P and Godby R W 2000 Efficient total energy calculations from self-energy models *Phys. Rev. Lett.* **85** 5611–14

[230] Savrasov P Y and Kotliar G 2003 Linear response calculations of lattice dynamics in strongly correlated systems *Phys. Rev. Lett.* **90** 056401

[231] Saxena S K and Dubrovinski L S 2000 Iron phases at high pressure and temperature: phase transition and melting *Am. Mineral.* **85** 372–5

[232] Saxena S K, Dubrovinsky L S and Haggqvist P 1996 X-ray evidence for the new phase of b-iron at high pressure and high temperature *Geophys. Res. Lett.* **23** 2441–4

[233] Scandolo S 2003 Liquid–liquid phase transition in compressed hydrogen from first-principles calculations *Proc. Natl Acad. Sci.* **100** 3051–3

[234] Scandolo S, Giannozzi P, Cavazzoni C, de Gironcoli S, Pasquarello A and Baroni S 2005 First-principles codes for computational, crystallography in the ESPRESSO package *Z. Kristallog.* **220** 574–9

[235] Schwager B, Chudinovskikh L, Gavriliuk A and Boehler R 2004 Melting curve of H_2O to 90 GPa measured in a laser-heated diamond cell *J. Phys.: Condens. Matter* **16** S1177–9

[236] Segall M D, Lindan P J D, Probert M J, Pickard C J, Hasnip P J, Clark S J and Payne M C 2002 First-principles simulation: ideas, illustrations and the CASTEP code *J. Phys. Condens. Matter* **14** 2717–44

[237] Seto M, Yoda Y, Kikuta S and Zhang X W 1995 Observation of nuclear resonant scattering accompanied by phonon excitation using synchrotron radiation *Phys. Rev. Lett.* **74** 3828–31

[238] Shen G, Mao H, Hemley R J, Duffy T S and Rivers M L 1998 Melting and crystal structure of iron at high pressures and temperatures *Geophys. Res. Lett.* **25** 373–6

[239] Sinogeikin S V, Zhang J Z and Bass J D 2004 Elasticity of single crystal and polycrystalline $MgSiO_3$ perovskite by Brillouin spectroscopy *Geophys. Res. Lett.* **31** L06620

[240] Soderlind P, Moriarty J A and Wills J M 1996 First principles theory of iron up to Earth core pressures: structural, vibrational and elastic properties *Phys. Rev.* B **53** 14063–72

[241] Soler J M, Artacho E, Gale J D, García A, Junquera J, Ordejón and Sanchez-Portal D 2002 The SIESTA method for *ab initio* order-n materials simulation *J. Phys. Condens. Matter* **14** 2745–79

[242] Song X D and Helmberger D V 1993 Anisotropy of the earth's inner core *Geophys. Res. Lett.* **20** 2591–4

[243] Speziale S, Zha C-S, Duffy T S, Hemley R J and Mao H-K 2001 Quasi-static compression of magnesium oxide to 52 GPa: implications of the pressure–volume–temperature equation of state *J. Geophys. Res.* **B106** 515–28

[244] Stackhouse S, Brodholt J P, Wookey J, Kendall J M and Price G D 2005 The effect of temperature on the seismic anisotropy of the perovskite and post-perovskite polymorphs of $MgSiO_3$ *Earth Planet. Sci. Lett* **230** 1–10

[245] Stadler R, Alfè D, Kresse G, de Wijs G A and Gillan M J 1999 Transport coefficients in liquids from first principles *J. Non-Cryst. Solids* **250** 82–8

[246] Steinle-Neumann G, Stixrude L and Cohen R E 1999 First-principles elastic constants for the hcp transition metals Fe, Co and Re at high pressure *Phys. Rev.* B **60** 791–9

[247] Steinle-Neumann G, Stixrude L, Cohen R E and Gülseren O 2001 Elasticity of iron at the temperature of the earth's inner core *Nature* **413** 57–60

[248] Stevenson D J 1998 States of matter in massive planets *J. Phys.: Condens. Matter* **10** 11227–34

[249] Stillinger F H and Weber T A 1985 Computer simulation of local order in condensed phases of silicon *Phys. Rev.* B **31** 5262–71

[250] Stixrude L and Cohen R E 1993 Stability of orthorhombic $MgSiO_3$ perovskite in the earth's lower mantle *Nature* **364** 613–16

[251] Stixrude L and Cohen R E 1995 Constraints on the crystalline structure of the inner core—mechanical instability of bcc iron at high pressure *Geophys. Res. Lett.* **22** 125–8

[252] Stixrude L and Cohen R E 1995 High-pressure elasticity of iron and anisotropy of the Earth's inner core *Science* **267** 1972–5

[253] Stixrude L, Cohen R E and Singh D J 1994 Iron at high pressure—linearized-augmented-plane-wave computations in the generalized gradient approximation *Phys. Rev.* B **50** 6442–5

[254] Stixrude L, Wasserman E and Cohen R E 1997 Composition and temperature of the Earth's inner core *J. Geophys. Res. Solid Earth* **102** 24729–39

[255] Struzhkin V V, Goncharov A F, Hemley R J and Mao H K 1997 Cascading Fermi resonances and the soft mode in dense ice *Phys. Rev. Lett.* **78** 4446–9

[256] Sturhahn W, Toellner T S, Alp E E, Zhang X, Ando M, Yoda Y, Kikuta S, Seto M, Kimball C W and Dabrowski B 1995 Phonon density of states measured by inelastic nuclear resonant scattering *Phys. Rev. Lett.* **74** 3832–5

[257] Sugino O and Car R 1995 *Ab initio* molecular dynamics study of first-order phase transitions: melting of silicon *Phys. Rev. Lett.* **74** 1823–6

[258] Sutton A P, Finnis M W, Pettifor D G and Ohta Y 1988 The tight-binding bond model *J. Phys. C.: Solid State Phys.* **21** 35–66

[259] Teter M D, Hemley J, Kresse G and Hafner J 1998 High-pressure polymorphism in silica *Phys. Rev. Lett.* **80** 2145–7

[260] Tromp J 1993 Support for anisotropy of the earth's inner core from free oscillations *Nature* **366** 678–81

[261] Tsuchiya T, Tsuchiya J, Umemoto K and Wentzcovitch R 2004 Phase transition in $MgSiO_3$ perovskite in the Earth's lower mantle *Earth Planet. Sci. Lett.* **224** 241–8

[262] Tsuchiya T, Tsuchiya J, Umemoto K and Wentzcovitch R M 2004 Elasticity of post-perovskite $MgSiO_3$ *Geophys. Res. Lett.* **31** L14603

[263] Vanderbilt D 1990 Soft self-consistent pseudopotentials in a generalized eigenvalue formalism *Phys. Rev.* B **41** 7892–5

[264] Verhoogen J 1980 *Energetics of the Earth* (Washington, DC: National Academy Press)

[265] Verlet L 1967 Computer 'experiments' on classical fluids. I. Thermodynamical properties of Lennard–Jones molecules *Phys. Rev.* **159** 98–103

[266] Vočadlo L and Alfè D 2002 The *ab initio* melting curve of aluminium *Phys. Rev.* B **65** 214105

[267] Vočadlo L, Alfè D, Brodholt J, Gillan M J and Price G D 2000 *Ab initio* free energy calculations on iron and its stability in the inner core *Phys. Earth Planet. Inter.* **117** 123–7

[268] Vočadlo L, Alfè D, Gillan M J and Price G D 2003 The properties of iron under core conditions from first-principles calculations *Phys. Earth Planet. Inter.* **140** 101–25

[269] Vočadlo L, Alfè D, Gillan M J, Wood I G and Price G D 2003 Possible thermal and chemical stabilization of body-centred-cubic iron in the earth's core *Nature* **424** 536–8

[270] Vočadlo L, Alfè D, Price G D and Gillan M J 2000 First-principles calculations on the diffusivity and viscosity of liquid Fe–S at experimentally accessible conditions *Phys. Earth Planet. Inter.* **120** 145–52

[271] Vočadlo L, Alfè D, Price G D and Gillan M J 2004 *Ab initio* melting curve of copper by the phase-coexistence approach *J. Chem. Phys.* **120** 2872–80

[272] Vočadlo L, Brodholt J, Alfè D, Price G D and Gillan M J 1999 The structure of iron under the conditions of the earth's inner core *Geophys. Res. Lett.* **26** 1231–4

[273] Vočadlo L and Price G D 1996 The melting of MgO—computer calculations via molecular dynamics *Phys. Chem. Mater.* **23** 42–9

[274] Vočadlo L, Price G D and Wood I G 1999 Crystal structure, compressibility and possible phase transitions in epsilon-FeSi studied by first-principles pseudopotential calculations *Acta Cryst.* B **55** 484–93

[275] Vočadlo N L, Wall A, Parker S C and Price G D 1995 Absolute ionic diffusion in MgO—computer calculations via lattice dynamics *Phys. Earth Planet. Inter.* **89** 199–218

[276] Wallace D C 1998 *Thermodynamics of Crystals* (New York: Dover)

[277] Walter M J, Kubo A, Yoshino T, Brodholt J, Koga K T and Ohishi Y 2004 Phase relations and equation of state of aluminous Mg-silicate perovskite and implications for the Earth's lower mantle *Earth Planet. Sci. Lett.* **222** 510–16

[278] Wang C S, Klein B M and Krakauer H 1985 Theory of magnetic and structural ordering in iron *Phys. Rev. Lett* **54** 1852

[279] Wasserman E, Stixrude L and Cohen R E 1996 Thermal properties of iron at high pressures and temperatures *Phys. Rev.* B **53** 8296–309

[280] Weir S T, Mitchell A C and Nellis W J 1996 Metallization of fluid molecular hydrogen at 140 GPa (1.4 Mbar) *Phys. Rev. Lett.* **76** 1860–3

[281] Wentzcovitch R, Karki B B, Cococcioni M and de Gironcoli S 2004 Thermoelastic properties of $MgSiO_3$ perovskite: insights on the nature of the earth's lower mantle *Phys. Rev. Lett.* **92** 018501

[282] Wentzcovitch R, Ross N L and Price G D 1995 *Ab initio* study of $MgSiO_3$ and $MgSiO_3$ perovskites at lower-mantle pressures *Phys. Earth Planet. Inter.* **90** 101–12

[283] Wentzcovitch R M, Martins J L and Price G D 1993 *Ab initio* molecular dynamics with variable cell shape: application to $MgSiO_3$ *Phys. Rev. Lett.* **70** 3947–50

[284] Wigner E and Huntington H B 1935 On the possibility of a metallic modification of hydrogen *J. Chem. Phys.* **3** 764–70

[285] Williams Q, Jeanloz R, Bass J D, Svendesen B and Ahrens T J 1987 The melting curve of iron to 250 gigapascals—a constraint on the temperature at the Earth's center *Science* **236** 181–2

[286] Wood B 1993 Carbon in the core *Earth Planet. Sci.* **117** 593–607

[287] Woodcock L V and Singer K 1971 Thermodynamic and structural properties of liquid ionic salts obtained by Monte Carlo computation *Trans. Faraday Soc.* **67** 12–32

[288] Yamamoto T, Yuen D A and Ebisuzaki T 2003 Substitution mechanism of Al ions in $MgSiO_3$ perovskite under high pressure conditions from first-principles calculations *Earth Planet. Sci. Lett.* **206** 617–25

[289] Yeganeh-Haeri A 1994 Synthesis and re-investigation of the elastic properties of single-crystal magnesium silicate perovskite *Phys. Earth Planet. Inter.* **87** 111–21

[290] Yoo C S, Holmes N C, Ross M, Webb D J and Pike C 1993 Shock temperatures and melting of iron at Earth core conditions *Phys. Rev. Lett.* **70** 3931–4

[291] Zerr A and Boehler R 1994 Constraints on the melting temperature of the lower mantle from high-pressure experiments on MgO and magnesiowustite *Nature* **371** 506–8

[292] Zerr A, Diegler A and Boehler R 1998 Solidus of earth's deep mantle *Science* **281** 243–6

[293] Zhang J and Weidner D J 1999 Thermal equation of state of aluminium-enriched silicate perovskite *Science* **284** 782–4