Calculation of free energies from *ab initio* calculation

2002 J. Phys.: Condens. Matter 14 2975

(http://iopscience.iop.org/0953-8984/14/11/311)

INSTITUTE OF PHYSICS PUBLISHING JOURNAL OF PHYSICS: CONDENSED MATTER

J. Phys.: Condens. Matter **14** (2002) 2975–3000 PII: S0953-8984(02)27809-1

Calculation of free energies from *ab initio* calculation

G J Ackland

Centre for Science at Extreme Conditions, Department of Physics and Astronomy, The University of Edinburgh, Edinburgh EH9 3JZ, UK

Received 13 August 2001, in final form 21 January 2002
Published 8 March 2002
Online at stacks.iop.org/JPhysCM/14/2975

Abstract

The calculation of total energy from electronic structure is now well established, and recent interest has moved to evaluation of free energies and equations of state. This paper discusses various methods for evaluating free energies, for equilibrium phases, for reaction pathways and for phase transformations.

1. Introduction

The ability to calculate crystal energies from *ab initio* simulation has been long established, and is now a mature field with widely available software (Gaussian, Castep, Vasp) and a proven record of complementing and reproducing experimental results.

There are basically four types of calculation. In *static lattice* calculations, the total energy is evaluated for a single set of ionic coordinates $\{\boldsymbol{r}_i\}$, no calculation of forces is required. In *static relaxation* the calculation of forces is used to find a local energy minimum[1] of ionic coordinates $\partial U(\{\boldsymbol{r}_i\})/\partial \boldsymbol{r}_i = 0$. This defines the energy of an 'ideal' crystal structure. Thirdly, using forces, *molecular dynamics* (MD) calculations may be performed which follow an energy-conserving deterministic trajectory through $\{\boldsymbol{r}_i\}$ space, from which thermodynamic quantities can be derived. Finally, without requiring forces, Monte Carlo (MC) dynamics follows a stochastic trajectory though $\{\boldsymbol{r}_i\}$ space.

Twenty years ago, some of the earliest *static lattice* calculations were applied to high-pressure phase transitions between polymorphic crystal structures. Using forces calculated with the Hellmann–Feynman theorem [1] and Pulay [2] corrections a huge number of *static relaxation* calculations have been carried out by many groups worldwide, proving the impressive accuracy of the method. A characteristic of these calculations, which calculate energy as a function of pressure, is that they are carried out at zero temperature[2]. In 1985, Car and Parrinello introduced an efficient MD algorithm [3] involving fictitious dynamics of the electronic degrees of freedom. Other MD methods involving direct minimization of the electronic wavefunctions have been introduced since [4], with similar efficiency.

[1] Strictly, if zero force is obtained by symmetry while any second derivative of $U(\{\boldsymbol{r}_i\})$ is negative, this may only be a saddle point representing a structure mechanically unstable with respect to a soft phonon mode.

[2] Strictly, even this is not true as zero-point vibrational energy is neglected.

Practical interest lies in finite temperature properties, where the quantities of interest are the gradients of *free energy* with thermodynamic variables, and the *free energy difference* between structures. For a constant number of particles, in a constant pressure ensemble the relevant quantity is the Gibbs free energy which is defined by

$$G = U + PV - TS \tag{1}$$

where U is the internal energy of ions and electrons. At $T = 0$, U and PV can be readily calculated from a static relaxation calculation. For finite temperature G includes an entropic contribution, so there is a need to sample from a thermodynamic ensemble to evaluate the free energy. Within statistical mechanics, we can relate free energy to the many particle partition function:

$$Z = \int \exp[-H(\{\boldsymbol{r}_i\})/k_B T]\, \mathrm{d}\{\boldsymbol{r}_i\} \tag{2}$$

where the integral runs over all possible atomic arrangements of the system compatible with the phase in question, the notation $\mathrm{d}\{\boldsymbol{r}_i\}$ denoting integration over all members of the set $\{\boldsymbol{r}_i\}$.

It is worth pausing to consider what this actually means, since many of the methods to be discussed involve evaluating Z. The integral over all possible arrangements of atoms clearly cannot be done by evaluating them all. One must therefore construct a process which will sample a representative set of states, visiting all regions of phase space with appropriate probability. The *principle of detailed balance*, that the transition probability between two states is proportional to the ratio of their Boltzmann weights, is helpful here, but in fact is neither sufficient nor necessary[3]. Any valid method must actually visit all the relevant inequivalent pockets of phase space within the duration of the simulation. It also must not sample from those regions of phase space which are *not compatible with the phase in question*. The practical problem of free energy calculation is how to sample most efficiently the phase space described by $\{\boldsymbol{r}_i\}$: in the specific case of calculations based on electronic structure, 'most efficient' means 'with the minimum number of electronic structure calculations', since the limiting factor is available computer time.

If the Hamiltonian $H(\{\boldsymbol{r}_i\})$ is a function of pressure (i.e. it is actually an enthalpy) then the Gibbs free energy is obtained:

$$G = -k_B T \log Z. \tag{3}$$

Otherwise, if $H(\{\boldsymbol{r}_i\})$ is a function of volume, which is often the easiest way to carry out *ab initio* simulations, the derived quantity is Helmholtz free energy:

$$F = G - PV = -k_B T \log Z. \tag{4}$$

The full equation of states or phase diagram can be extracted from calculations of either F or G as a 2D surface in PT or VT space. Although G is generally more relevant to experimental conditions, computational simplicity leads most workers to evaluate F. More complex equations of state for isotropic compression, chemical composition or order parameter may add additional dimensionality to the free energy surface. In other cases one is interested in following a line along a PVT surface, such as the 'geotherm' which follows conditions of temperature and pressure in the deep Earth or the 'Hugoniot' which is the set of states observed in a shock wave.

The alternative to evaluating the partition function at every pressure and temperature where one wishes to know the free energy is to use thermodynamics. The combined first and second laws give

$$\mathrm{d}U = T\,\mathrm{d}S - P\mathrm{d}V. \tag{5}$$

[3] It is not *necessary*, because the order of sampling is irrelevant, e.g. picking states consecutively from three groups may be valid. It is not *sufficient* because it offers no guarantee that all regions of phase space will be sampled.

We can calculate entropy by integrating the heat capacity along isochores (lines of constant volume: the most convenient ensemble for electronic structure calculation)

$$S(T, V) = \int_{U(0,V)}^{U(T,V)} \frac{\mathrm{d}U}{T} = \int_0^T \frac{1}{T}\left(\frac{\partial U}{\partial T}\right)_V \mathrm{d}T. \tag{6}$$

Approaches to evaluating free energies fall into two camps, attempts to evaluate equation (2) directly, and integration along a reversible path from a state of known free energy. The central problem is to sample sufficient uncorrelated microstates to obtain an accurate measure of thermodynamic quantities at each point in PT space. This is tractable because the averaged quantities of thermodynamic interest converge without requiring fully ergodic sampling of phase space [5][4].

For single-phase equation of state application, one only needs gradients of the free energy, so it is sufficient to evaluate directly the relationship between V and P and T from simulation without calculating the partition function or entropy directly. This can be done along a line of points in PT space by ensemble averaging from an MD or MC simulation. It works whenever the PVT surface is continuous and differentiable. In the presence of a phase transition there are discontinuities which will invalidate this approach.

For the free energy of a particular crystalline phase, the partition function can be written as in equation (2) where $\{\boldsymbol{r}_i\}$ is the set of particle coordinates and $H(\{\boldsymbol{r}_i\})$ the total energy, and where the integral runs over all microstates associated with the macrostate (phase at a particular volume) whose free energy is being evaluated. Given a sufficiently large, unbiassed and uncorrelated sampling of microstates representing a particular phase, the partition function integral can be evaluated. In the context of *ab initio* simulation obtaining such a sampling requires long simulations, or some simplifying assumption such as the quasiharmonic approximation which models U_i as a multidimensional parabola.

In this paper, the state of the art in *ab initio* free energy calculations is reviewed, starting from simple parametrized models where *ab initio* data is used in the parametrization, moving through quasiharmonic lattice dynamics to full phase space integration via MD or MC. New methods, lattice switch MC and temporal embedding, which have the capability to be applied to *ab initio* free energy calculation are also discussed. Many of the techniques used were derived in the context of hard spheres or classical potential simulation, and are now being applied (with necessary adjustments) in conjunction with *ab initio* energies and forces, so in general we concentrate on the issues raised by the extremely high computational cost of electronic structure determination compared with other aspects of the calculation. The review ends with case studies of application of the various methods to silicon, iron and perovskite.

2. Molecular dynamics and Monte Carlo simulation

Measuring the change in pressure with temperature and volume in constant volume MD allows the free energy change to be monitored across a range of volumes and pressures [6]. MD cannot give an absolute free energy, and can be problematic across phase boundaries, but for practical purposes such as evaluating the equation of state for a single-phase material it is a reliable method.

[4] This can be contrasted with defect properties such as equilibrium vacancy concentration which converge unreasonably slowly with both time and system size to be simulated directly by awaiting vacancy production. Sampling is also a problem when phase transitions occur—in general a single simulation will not visit regions of phase space corresponding to both phases.

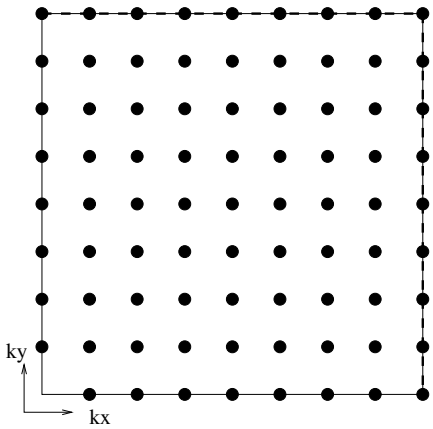

Figure 1. Phase space sampling achieved by MD, shown schematically for a 9 × 9 particle simulation on a 2D square lattice (one particle per unit cell). The region around Γ (bottom left) is under-represented, while the zone boundary (dashed line) is oversampled.

As with all free energy calculation, the drawback is in the statistics gathered. MD does a good job of capturing anharmonic phonon effects, but the absence of any quantization means that the low-temperature limits for the entropy are incorrect. Only $3N$-3 vibrational modes are sampled (rigid translations being ignored), and these do not evenly sample the phase space, omitting the low-k acoustic branch which makes the largest contribution to phonon free energy, while over-representing the zone boundary modes (figure 1). Also, the differential change of atomic positions in MD means that each microstate is heavily correlated with its predecessor. In the limit of a harmonic crystal, strict microcanonical MD can never achieve equipartition between modes. Constant temperature ensembles [40] are similarly slow to equilibrate, so if MD is being used to sample phase space it is preferable to re-randomize the velocities occasionally.

In principle, MC simulation offers a more efficient way obtain uncorrelated samples from the systems' phase space in order to evaluate the free energy. However, large changes ionic positions are problematic in Car–Parrinello [3] *ab initio* simulation, where the plane wave coefficients follow an equation of motion, and in any method where the wavefunction is extrapolated from one step to the next. Moreover, the cost of energy evaluation is such that high move acceptance probabilities are essential. In going from the smooth trajectories of MD to the random walk of MC much of the efficiency is lost, and so most *ab initio* free energy calculations have used MD to sample phase space.

A hybrid method for avoiding this correlation by evolving the atomic coordinates via an analytic potential fitted to exactly reproduce the *ab initio* forces has recently been suggested [7]. There uncorrelated trial configurations are produced quickly with analytic potential MD or MC, and accepted/rejected using Metropolis sampling—only *ab initio* energies are used in the partition function integral, and the lack of correlation between microstates leads to better phase space sampling.

3. Parametrizing a model

A simple approach for evaluating phase stability of free energy differences is to use the *ab initio* calculation to parametrize a model Hamiltonian and evaluate free energies from that.

This may have the advantage that the model Hamiltonian is exactly solvable, or so simple to compute that the sampling error can be reduced. A further advantage is that general principles can be established which allow for straightforward comparison between related systems. The disadvantage is that some of the physics is lost in converting from one model to the other, so quantitative accuracy is lost.

3.1. The Landau-model approach

Soft mode materials lend themselves to a parametrized Landau-model approach. Here, a particular normal vibrational mode (phonon frequency) of the system tends to zero frequency, at which point the structure becomes unstable. All other degrees of freedom can be integrated out of the problem (renormalized), and this instability leads to a phase transition. In the simplest theory, the transition is second order, but in practice the phonon[5] distortion invariably couples to a strain causing a first-order transformation.

The methods described below involve static lattice calculations on small unit cells, which are then used to parametrize models of masses and springs moving in potential wells. They are well suited to extracting qualitative behaviour, but quantitative predictions of, for example, phase transition temperatures are unreliable.

To extract a free energy representation for a soft mode transition the free energy is expanded in terms of a single order parameter ξ

$$G(\xi, V, T) = -A(T, V)\xi^2 + B(V)\xi^4 \tag{7}$$

where $\xi = 0$ represents the high-temperature phase (stable at negative A) and $\xi_0 = \sqrt{A/2B} \neq 0$ is the order parameter for the low-temperature phase (positive A).

The Landau theory can be parametrized in a simple way, if zero-K calculations of energy against phonon amplitude has quartic shape: two equivalent minima separated by a free energy barrier[6] of height $\Delta G(V)$:

$$\Delta G(V) = A(0, V)/4B(V) \tag{8}$$

$$\xi_0(V) = \sqrt{A(0, V)/2B(V)} \tag{9}$$

which gives

$$A(0, V) = 2\Delta G(V)/\xi_0(V)^2 \tag{10}$$

$$B(V) = \Delta G(V)/\xi_0(V)^4. \tag{11}$$

In the simplest, classical, picture we consider the soft mode to be an isolated oscillator with energy $k_B T$, in which case the barrier disappears at $k_B T = \Delta G$ and we can assign the temperature dependence of the Landau parameter

$$A(T, V) = A(0, V)(1 - k_B T/\Delta G). \tag{12}$$

Using this model[7], we can build up a complete P–T phase diagram for the two-phase system, based only on *static relaxation* calculations across a range of $u(V)$ and V.

The main assumption is that the phase transition can be represented by a single order parameter, which in turn is a single phonon coupled to a strain. In general, several phonons will contribute. Other assumptions required for use with *ab initio* simulation are that the free

[5] For the purposes of total energy calculation, and most experimental probes, a phonon is a quantized strain-preserving normal mode excitation of the system—it involves motion of the ions uncoupled to changes in the unit cell.

[6] This is a free energy because, in theory, the contributions of all other modes should be integrated over: in practice this is seldom done, since their contribution to free energy *differences* along $\xi_0(V)$ is small, becoming zero in the harmonic limit.

[7] Provided we ignore the coupling to other modes, as described in the previous footnote.

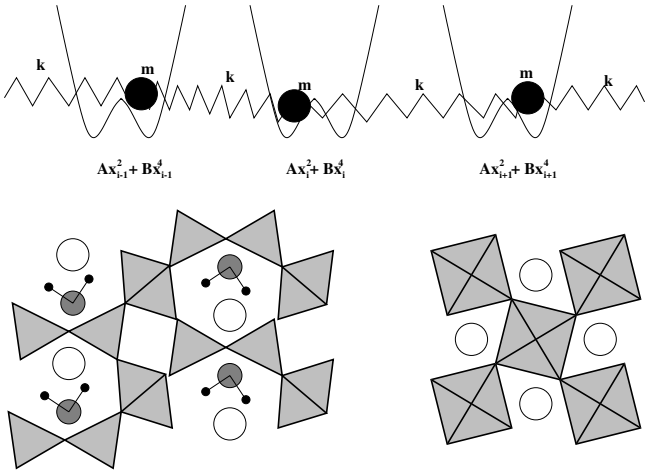

Figure 2. 1D representation of a simplified soft phonon model: particles reside in double-well potentials connected by springs. Below left, the structure of lawsonite can be treated this way, the water molecules and calcium ions (circles) are held in aluminosilicate cages (grey triangles), the orientation of the water has a double-well potential while the spring models coupling between adjacent water molecules (through the framework). Below right, $MgSiO_3$ in the perovskite structure has magnesium ions (circles) in silicate cages (grey triangles). Here the phase transition sees the SiO_6 octahedra rotate (the double well) and distort (the spring).

energy can be calculated and will have a linear temperature dependence for A. Finally one assumes that the quartic oscillator can be treated classically, with has energy $k_B T$, and the phase transition occurs when the system has enough energy on average to cross the barrier. The quantum interpretation is more problematic, as one has to relate a tunnelling probability to the timescale of the transition [8].

3.2. The interacting double-well approach

A more sophisticated version of the Landau model comes from mapping the system onto an interacting double-well model [9] (figure 2). Although still far from an *ab initio* approach, it enables one to capture some of the universality class behaviour of a certain phase transition, which in turn provides a bridge from the lengthscales tractible in *ab initio* simulation to those relevant to the critical behaviour near a phase transition. Now one envisages localized quartic oscillators coupled by harmonic springs. This model can be solved by numerical simulation [10, 11], and maps onto any system where the atom (or group of atoms) which disorder are weakly coupled linked. It is appropriate for describing instabilities in materials where the high-temperature structure has an entire band on unstable phonons (figure 2). The Hamiltonian is

$$H = \sum_i \frac{m}{2}\left(\frac{\mathrm{d}\xi_i}{\mathrm{d}t}\right)^2 - A(V)\xi_i^2 + B(V)\xi_i^4 + \frac{1}{2}\sum_{i,j_i} J\xi_i\xi_{j_i} \tag{13}$$

where j_i represent the neighbours of atom i. The apparently strange form of the interaction term arises from incorporating the on-site parts ($k\xi^2/2$) of the spring interaction into $A(V)$.

It is straightforward to calculate the dispersion relation $\omega(\xi_i)$ for this model, which gives the general result that the mean squared frequency is A (typically negative), while the dispersion is controlled by J. This suggests that an uncoupled branch of phonons with imaginary frequencies can be mapped onto this model [12].

The interesting feature about this model is that the site disorder depends on the spring strength k and the well depth. In the *order–disorder* limit of $k = 0$ the structure is always in the high-symmetry disordered structure, since the occupation of each minimum is uncorrelated and therefore equal. In the *displacive* limit $k \to \infty$ we revert to the Landau model. In general, if an entire branch of instability exists the phase transition temperature corresponds to the energy dispersion of that branch [11].

A good example of a system described by this method is lawsonite (figure 2), a mineral in which water molecules are trapped within an aluminosilicate framework. The phase transitions are primarily due to the orientation of the water molecules which are more stable in a tilted configuration (the double well) and are weakly coupled through strain in the framework (the spring) [13]. On heating correlation between neighbouring tilts is lost, and the crystal structure adopts a higher symmetry structure.

A conceptually more complex example is perovskite ($MgSiO_3$) (figure 2), a mineral which makes up much of the Earth's core. Here the 'double-well' potential represents the orientation of SiO_6 octahedra, and the spring represents the coupling between octahedra (which actually arises as a distortion of the octahedra themselves). Warren [11] has done extensive simulations of this system, as discussed below.

In some cases, it is useful to go beyond the one-dimensional local potential and build a local energy landscape, as we now discuss.

4. Energy landscapes and free energy landscapes

A generalization of the above approaches which is useful in examining transition mechanisms or reaction pathways is the energy or free energy *landscape*.

In modelling processes such as molecule-surface reactions [14] the $3N$-dimensional coordinate space $\{\boldsymbol{r}_i\}$ is reduced to a low-dimensional reaction space $\{\tilde{\boldsymbol{r}}_i\}$, the assumption being that oscillations therein are uncoupled from[8] the other vibrations of the system. For fast dynamic processes, this reaction space defines an energy with other coordinates orthogonal[9] to $\{\tilde{\boldsymbol{r}}_i\}$: being fixed by static relaxation. For slower processes, such as phase transformations, the reaction space should define a free energy $F(\{\tilde{\boldsymbol{r}}_i\})$ (cf equation (2) with the orthogonal coordinates being integrated across and $\{\tilde{\boldsymbol{r}}_i\}$ held fixed.

The reduced dimension energy landscape can be used as an *effective Hamiltonian*, in which the reaction is treated by motion of a quantised wavepacket. Reactions of diatomic molecules with surfaces [14–16] are done with this method, wherein $\tilde{\boldsymbol{r}}_i$ represent the bond length and the distance between the molecular centre of mass to the surface, other members of $\{\boldsymbol{r}_i\}$ being relaxed. The Schrödinger equation is solved in this 2D potential: in some cases it has proved important to extend the potential surface to 6D to account for molecular rotation and surface relaxation [16, 17].

In the extreme case, one may wish to plot a free energy a single coordinate, the order parameter (the reaction path). This involves an integration over all but one coordinate $\xi = \tilde{r}$ in the phase space. This may be achieved either by predetermining ξ and evaluating equation (2) with fixed ξ and all other coordinates integrated over, or systematically by the 'nudged elastic band' (NEB) [18] method.

In NEB, a series of coupled *ab initio* simulations are run starting at positions 'images' along a guessed reaction path between states A and B. The Hamiltonian consists of the energy of each of the images, plus an elastic interaction between them (figure 4), which increases with separation in $\{\boldsymbol{r}_i\}$ space.

[8] or, in the free energy case, renormalized by.

[9] In the multidimensional space of position and i, $\{\tilde{\boldsymbol{r}}_i\}$ can be any linear combinations of components of $\boldsymbol{r}_i$.

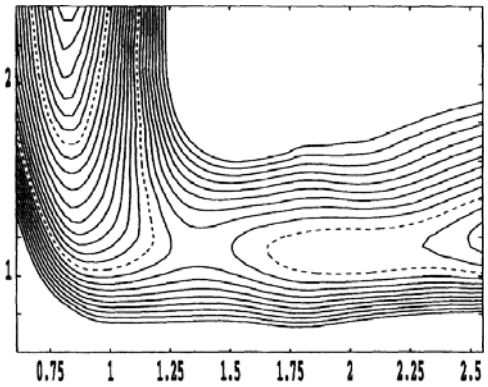

Figure 3. Potential energy surface for reduced set of coordinates $\{\tilde{r}_i\}$, where $\tilde{y}$ is the height of a hydrogen molecule above a copper surface and $\tilde{x}$ is the hydrogen bond length, from [15].

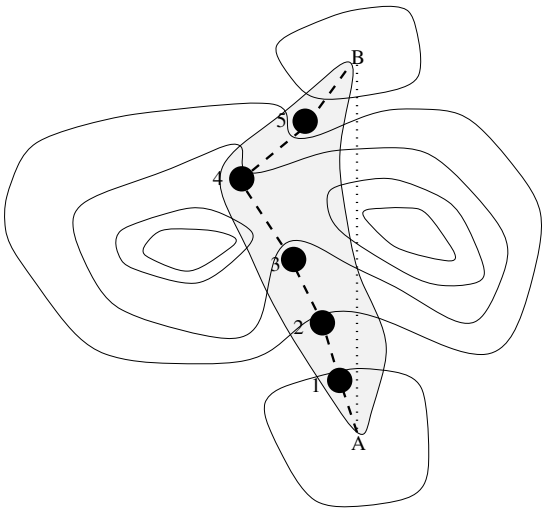

Figure 4. Schematic representation of NEB method. Contours represent $E(\{\boldsymbol{r}_i\})$. The dotted line represents a simple order parameter between states A and B, which can be seen to traverse unnecessarily high-energy states. The dashed curve represents the NEB which follows lower energy states, with the dots representing the images at which *ab initio* simulations are done. The grey shading represents possible trajectories at finite temperature, which should be integrated over to obtain the *free* energy.

$$E_A + E_B + \sum_{j=1,5} E_j + k(\boldsymbol{r}_i^{\boldsymbol{A}} - \boldsymbol{r}_i^{\boldsymbol{1}})^2 + (\boldsymbol{r}_i^{j} - \boldsymbol{r}_i^{\boldsymbol{j+1}})^2 + (\boldsymbol{r}_i^{B} - \boldsymbol{r}_i^{\boldsymbol{5}})^2 \tag{14}$$

where each E_i represents an energy calculation at an image (given by coordinates $\boldsymbol{r}_i$) along the reaction path between A and B while the spring terms introduce an energy penalty against large changes in the coordinates between images.

What is actually being attempted is the minimisation of the line integral of energy between A and B. In the original version, this energy is simply minimized. A technical improvement [19] to this is to consider only the spring force in directions perpendicular to the chain (this suppresses a tendency for the path to develop kinks) and to set the true forces along the path to zero. This increases the density of images near the transition state by preventing the images from sliding away from the highest energy points.

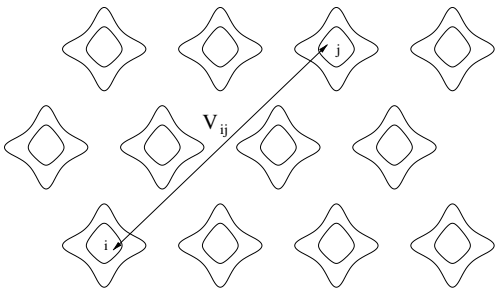

Figure 5. In the effective Hamiltonian approach, long-range electrostatic forces V_{ij} can be incorporated with short-ranged *ab initio* interaction data which determine onsite dipoles (represented here by contours). The resulting model can be used for MC free energy calculations. This type of model can be used to predict ferroelectric transitions and domain formation at length scales well beyond what is practical in pure *ab initio* simulation.

To date, the NEB method has been used in conjunction with static minimisation, wherein E_i is the total energy and the variables $\{\boldsymbol{r}_i^j\}$ comprise all atomic positions. This essentially gives the zero temperature lowest energy reaction path. At finite temperatures the free energy associated with each image of the Hamiltonian can be extracted from an MD simulation of the dynamical system (equation (14)), and hence a free energy barrier computed.

In soft-phonon materials such as ferroelectrics, the *effective Hamiltonian* approach can be used to model explicitly a subgroup of phonons, while integrating out the uninteresting degrees of freedom. If the soft modes can be effectively localized, then a model such as equation (13) can be used with long-range electrostatic terms (external electric fields) replacing the springs J (figure 5 [20, 21]). A promising approach for localization is to expand the phonon modes in Wannier functions [22]

In phase transitions we are interested in a discontinuous 'free energy landscape', evaluated only at particular crystal phases. Since free energy encompasses entropy, which cannot be defined for a microstate, any calculation of the free energy of a phase must first address the question of the definition of the set of microstates which make up the phase. For a crystalline phase, α, this can be defined as 'all microstates lying within a basin in energy space from which the system cannot escape on the timescale of the experiment'[10]. This defines the limits of the integral for the free energy (equation (2)) which is then

$$F_\alpha = -k_B T \log \int_\alpha \mathrm{d}\{\boldsymbol{r}_i\} \exp[-E(\{\boldsymbol{r}_i\})/k_B T]. \tag{15}$$

Under this definition, a free energy landscape cannot be plotted in atomic position space[11]: for determining phase stability we are only interested in the free energy differences of phases.

5. Parametrizing a potential

The accuracy of free energy calculation using analytic interatomic potentials is limited by the accuracy of the potential, while *ab initio* calculation is limited by finite size and sampling time effects. Some work has been done to develop hybrid methods which take the advantages of both. A simple approach is to fit directly an analytic interatomic potential using data generated

[10] The reference to a timescale may appear worrying, but recall that for zero pressure the equilibrium state of any system is an infinitely dilute gas, and even at finite pressure all but one of the states we consider are metastable.
[11] The free energy does, of course, vary with P and T.

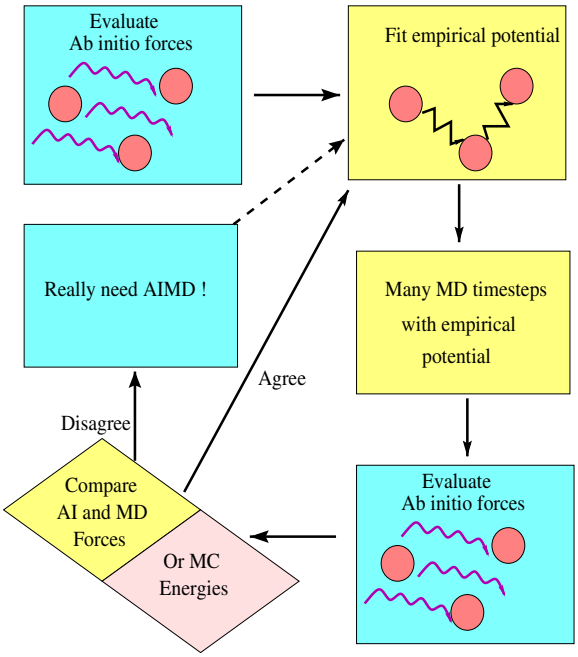

Figure 6. Flowchart depicting the temporal-embedding scheme.
(This figure is in colour only in the electronic version)

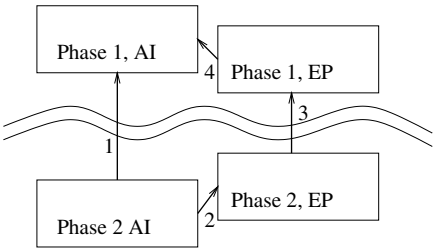

Figure 7. Schematic diagram of free energy calculation by direct *ab initio* LSMC (route 1), or split into two HSMC and one classical LSMC (2, 3 and 4) steps. AI labels the *ab initio* Hamiltonian, while EP represents an analytic potential chosen such that the EP energy is as close as possible to the AI one. In contrast, thermodynamic integration requires a continuous set of calculations along 2 and 4, together with either a further integration to zero temperature with the EP, or a restricted choice of EP such that the entropy becomes analytic.

from *ab initio* calculation. This technique is not discussed here, as it is not significantly different from work using empirical potentials: where equivalent experimental data is available, such an approach only serves to incorporate errors in the *ab initio* calculation into the potential[12]. A better approach is to incorporate *ab initio* calculation of properties for which experimental data is either unreliable or unavailable. There are three levels of doing this, at each level one trades transferability for accuracy. All depend on the assumption that the functional form chosen for the potential is valid.

[12] And allow authors to claim their work to be *ab initio* rather than empirical, which often adds credibility without adding accuracy.

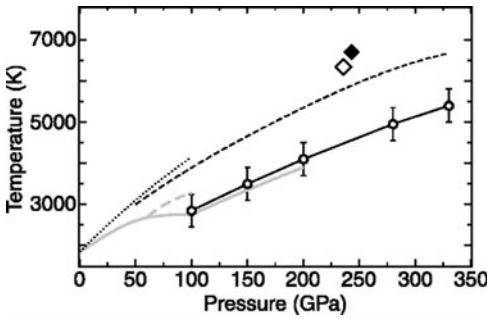

Figure 8. Results of *ab initio* free energy calculations and experimental data for the melting curve of iron. Gray curves are diamond anvil cell experiments by Boehler [65](solid), Shen [66](dashed) and Williams [67](dotted). Diamonds show shock wave data by Yoo [68] (open) and Williams [67] (filled). Open circles with error bars are calculations from Laio [23] while black dashed curve shows calculations by Alfe [43] reproduced from the paper by Laio *et al* [23].

(1) Parametrizing a potential for use across all thermodynamic and phase space. The same potential is then used throughout.
(2) Parametrizing a potential for use in a specific region of thermodynamic space (e.g. volume and temperature dependence). Different parametrizations of the same potential are then used at different P and T [23][13].
(3) Parametrizing a potential 'on the fly' for moving between adjacent microstates in phase space.

The first option is well established and equivalent to any other empirical potential approach and will not be discussed further.

The second has recently been implemented by Laio *et al* [23] to investigate the problem of iron under Earth-core conditions. Here an MEAM [25] potential V_0 is fitted to forces from a single 'typical' microstate calculated *ab initio* for 64 iron atoms. This potential is used in an isothermal–isobaric MD run to evaluate the density. A typical microstate from the V_0 run is then recalculated *ab initio*, and a new potential V_1 constructed by a force matching procedure [26]. This procedure is iterated until the density is constant (within 0.5%) between V_{n-1} and V_n at which point V_n is adopted as the optimised potential. Now a standard [6] MD calculation with V_n will elicit thermodynamic quantities free from the finite sampling size[14] and short timescales of *ab initio* simulation. The drawback of this method is that it still requires the potential to have a functional form flexible enough to fit the anharmonic interactions accurately: too few parameters cannot capture the correct physics, while too many can lead to instabilities between V_{i-1} and V_i.

'On the fly' fitting of potentials makes use of the fact that both the fitting procedure and analytic potential MD or MC are many orders of magnitude faster than *ab initio* calculation. The scheme is relatively simple: an *ab initio* timestep is taken, forces $\boldsymbol{F}^{AI}$ calculated and an analytic interatomic potential is parametrized using force matching [26] and an MD simulation continues using forces $\boldsymbol{F}^{EP}$ taken from the analytic potential. For example, for an analytic potential characterised by some site (j) dependent parameters $\{a_{j,k}\}$, at each step one minimizes a target function

[13] This reparametrizing approach should be distinguished from explicitly volume dependent potentials, such as screened effective pair potentials arising from nearly free electron theory [24].
[14] Finite size problems may recur since the potential is fitted to data which does suffer from small system size. In this case the use of a short-ranged MEAM incorporates this error.

$$\chi = \sum_j (\boldsymbol{F}_j^{SE}(\{a_{k,j}\}) - \boldsymbol{F}_j^{AI})^2 + \sum_k W_k \sum_j (a_{j,k} - \langle a \rangle)^2 \tag{16}$$

where W_k and are penalty functions which serve to minimize the variation in potential parameters between sites. In the limit of large W_k this gives the same potential everywhere.

At the end of the MD segment another *ab initio* calculation is done to obtain forces and energies at $t + \Delta t$, which are compared to those from the analytic potential using parameters $\{a_{j,k}(t)\}$. Now one adopts some accept/reject criterion to determine whether the analytic potential has adequately described the segment. There are several ways to perform this test, depending on the application. Energies, mean or maximum forces can be compared to some tolerance, or the potential can be refitted and the analytic potential MD run time-reversed for comparison. If accepted, the analytic potential is refitted to obtain $\{a_{j,k}(t + \Delta t)\}$, otherwise the segment is rerun with *ab initio* forces throughout (in MD) or with different initial velocities (in hybrid MC q.v.).

A major difference with other fitting methods is that each atom has its own potential parameters (within a common functional form). This is at odds with the idea of transferability, but it allows the parametrization to best compensate for local environmental effects which are missing in the functional form of the potential. In an extreme view, the analytic potential can be viewed simply as an advanced integration algorithm.

This 'on the fly' fitting can be combined with an embedding approach where the refitting is done more regularly using *ab initio* clusters in spatial regions which are expected to be of interest, such as a chemical bond in a biomolecule between atoms steered by their environment [27].

For free energy calculation, the embedding of occasional *ab initio* calculations in longer analytic potential segments is known as *temporal embedding* and can be used to quickly obtain uncorrelated microstates, and hence converge the ensemble averages more rapidly, including only the *ab initio* energies. The decorrelation can be further enhanced using a hybrid MC approach in which particle velocities are randomized at each refitting [7][15].

Appropriate choice of functional form is still crucial: in a calculation of high-pressure iron with a pseudopotential plane wave–Lennard-Jones hybrid [7] exact parameter fitting was shown to lead to massive variation in parametrization between atoms, essentially because at high pressure both r^{-6} and r^{-12} terms are fitting the short-range repulsion. This can be alleviated [28] by running a Car–Parrinello-style [3] fictitious dynamics on the analytic potential parameters, with fictitious forces generated from the force-matching steps [26].

6. Quasiharmonic lattice dynamics

Quasiharmonic lattice dynamics involves formally breaking the internal energy of a material into three parts:

$$U(V, T) = U(V, 0) + U_{harm}(V, T) + U_{anharm}(V, T). \tag{17}$$

The 'cold curve' $U(V, 0)$ is the energy of a static lattice calculation, the harmonic contribution U_{harm} is evaluated from the phonons, which are treated as quantized harmonic oscillators,

$$U_{harm}(V, T) = \sum_{i,j,\alpha,\beta} \frac{\partial^2 U(V, 0)}{\partial u_{i,\alpha} \partial u_{j,\beta}} u_{i,\alpha} u_{j,\beta} \tag{18}$$

[15] A typical AIMD calculation converges forces to around 0.1 eV Å^{-1}, providing a natural tolerance for the accept/reject criterion.

where $u_{i,\alpha}$ is the displacement from equilibrium of the ith atom in the α direction, $\partial U(V,0)/\partial u_{i,\alpha}$ being zero at equilibrium.

The quasiharmonic approximation ignores the anharmonic contribution $U_{anharm}(V,T)$ and uses lattice dynamics [33] to evaluate the phonon spectrum at a range of densities. From the associated density of states, the harmonic phonon modes can be populated according to Bose–Einstein statistics.

The canonical partition function for a harmonic oscillator is $Z(\omega) \equiv \sum_{n=0}^{\infty} \mathrm{e}^{-\beta\epsilon_n(\omega)} = 1/(2\sinh(\beta\hbar\omega/2))$, where $\beta = 1/k_BT$, k_B is the Boltzmann's constant. So the mean thermal energy of the mode is [29]

$$\bar{\epsilon}(\omega) = -\frac{\partial \log(Z(\omega))}{\partial\beta} = \frac{\hbar\omega}{2\tanh(\beta\hbar\omega/2)}.$$

If the frequency density-of-states function $g(\omega)$ is known, the harmonic energy may be written as

$$U_{harm}(V,T) = \int_0^\infty g(\omega,V)\bar{\epsilon}(\omega,T)\,\mathrm{d}\omega. \tag{19}$$

Once the internal energy U of each phase is calculated as a function of volume[16] V and temperature T, other thermodynamic quantities can be evaluated using the combined first and second laws of thermodynamics (equations (5), (6)) [30] taking $S = 0$ at $T = 0$ from the third law.

Having found the entropy, we obtain the Gibbs free energy as $G(T,V) = U(T,V) - TS(T,V) - PV$ or Helmholtz free energy $F = U(T,V) - TS(T,V)$.

For simulations of single phase equation of state, this is all that is required. In multiphase systems one has to evaluate the single phase equation of state for each polymorph, and take the one with lowest free energy. In each case the phase boundary is found by constructing the common tangents between the two phases' $F(V,T)$ surfaces along lines of constant T [31].

In determining the phase diagram ensemble becomes important: for static equilibrium the relevant ensemble is *NPT*, and the phase boundary is a line in *PT* space (P being the gradient of the common tangent). Alternately, for applications such as shock waves the relevant ensemble may be *NVT*. In this case the *VT* phase diagram has a coexistence region between the two volumes at which the common tangent touches the free energy curve. Most high-pressure experiments are carried out in diamond anvil cells, in which the pressure is increased by moving two diamonds closer together. This may appear to be a constant volume set up, but the sample is usually immersed in a pressure transmitting medium to ensure hydrostatic pressure. Notwithstanding this, it is common for diamond anvil experiments to report phase coexistence. It is also possible to set up anisotropic stresses across a sample, in which case pressure and volume are replaced by tensors: stress and strain [32].

All of this is standard statistical mechanics and thermodynamics. The problem for *ab initio* calculation becomes evaluating the phonon density of states, which is itself a continuous function of the specific volume. Early workers used Einstein (single frequency) or Debye models, but it is now possible to calculate the harmonic density of states exactly.

Work in determining the volume-dependent density of states has followed two directions: *ab initio* lattice dynamics and linear response theory.

6.1. Ab initio lattice dynamics

A practical method for calculating phonon frequencies is to use classical lattice dynamics. The method involves constructing a matrix of force constants, using this to build a dynamical

[16] Note that the volume is inversely proportional to the mass density.

matrix for a particular wavevector k, and diagonalizing the dynamical matrix to obtain phonon frequencies and hence $g(\omega)$. There is an extensive literature on lattice dynamics [33], so here we concentrate on issues specific to *ab initio* simulation.

The concepts of an interatomic potential or the energy of an atom are not well defined in electronic structure calculation. The force constants are well defined however, being the second derivatives of the total energy with respect to ionic displacements from equilibrium in equation (18): $\partial^2 U/\partial u_{i,\alpha}\partial u_{j,\beta}$. These can be calculated numerically [34] or analytically [35, 36].

Numerical calculation[17] proceeds by carrying out a static lattice calculation for a system in which one atom is displaced slightly from equilibrium, and evaluating the restoring forces $F_{j,\beta}$ on all atoms, then

$$\Phi_{i,\alpha,j,\beta} = \frac{\partial^2 U}{\partial u_{i,\alpha}\partial u_{j,\beta}} = \frac{F_{j\beta}}{u_{i,\alpha}}. \tag{20}$$

One such calculation yields a single row of this matrix of force constants Φ. Symmetry can be used to determine some other rows, but usually more than one calculation is required. The range of the force constants is restricted to less than one half of the simulation cell size, and means that large supercells are required to evaluate long-range force constants. Moreover, the Ewald summation does not properly account for screening of the dipole moment so corrections are required for the LO phonon frequencies near Γ [37].

The matrix of force constants should (a) be symmetric, (b) reflect the crystal symmetry of the system and (c) satisfy Newton's third law:

$$\sum_j \frac{\partial^2 U}{\partial u_{i,\alpha}\partial u_{j,\beta}} = 0 \ \forall \ i, \alpha, \beta. \tag{21}$$

Evaluating the matrix of force constants numerically does not guarantee this: some symmetry-related matrix elements are calculated more than once, and since forces are evaluated by iterative convergence slightly different values may be obtained (e.g. $F_{j\beta}/u_{i,\alpha} \neq F_{i\alpha}/u_{j,\beta}$). The consequence is that the phonon energies might (a) become complex, (b) not reflect crystal symmetry or (c) have non-zero Γ acoustic modes. There are schemes to ameliorate the inequivalent force problem [34], and it may even be turned to advantage as a method of probing anharmonicity [30]

The dynamical matrix $D_{i,\alpha,j,\beta}(\boldsymbol{q})$ can be built from the matrix of force constants:

$$D_{i,\alpha,j,\beta}(\boldsymbol{q}) = \sum_{l=-\infty}^{\infty} \frac{\Phi_{i,\alpha,j_l,\beta}}{\sqrt{m_i m_j}} \exp(\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{x}_l) \tag{22}$$

where i and j now label the atoms in a primitive unit cell, with m_i and m_j being their masses, and j_l being the equivalent of j in the lth supercell while l runs over all the primitive unit cells in the crystal and $\boldsymbol{x}_l$ is the displacement of that cell from the origin. The size of the dynamical matrix $D_{i,\alpha,j,\beta}$ is thus three times the number of atoms in the unit cell, whereas the matrix of force constants Φ is (in principle) infinite. The energies and eigenvectors of modes at arbitrary wavevector $\boldsymbol{q}$ are the eigenvalues and eigenvectors of the dynamical matrix.

Assuming that the force constants have finite range, the dynamical matrix can be constructed using equation (22). In practice, the assumption of short range is implicit in supercell methods, since the restoring forces are only calculated for a finite number of atoms: strictly, the force constant calculated in a supercell simulation is

[17] Sometimes erroneously referred to as the frozen phonon method: strictly a frozen phonon calculation assumes knowledge of the phonon eigenvector to calculate a specific mode compatible with the supercell.

$$\Phi'_{i,\alpha,j,\beta} = \sum_l \frac{\partial^2 U}{\partial u_{i,\alpha} \partial u_{jl,\beta}} \tag{23}$$

where l labels the supercell images of atom[18] j. As a consequence, the frequencies of all phonons which are commensurate with the supercell are calculated exactly. For interpolation of the dispersion relation, one can either assign to the nearest image the calculated value of[19] $\Phi'_{i,\alpha,j,\beta}$, and set longer range interactions to zero (cutting off Φ at half the size of the supercell), or postulate some dependence of Φ on distance. This choice is somewhat arbitrary, and can affect the appearance of the density of states, but generally has little effect on integrated quantities such as free energies [34].

6.2. Linear response theory

The dynamical matrix can also be evaluated analytically. The ionic contributions can be calculated from Ewald sums, as for classical lattice dynamics while the electronic contributions are treated via density functional perturbation theory [35, 36]

$$\Phi_{i\alpha j\beta} = \int \left[\frac{\partial V(\boldsymbol{r})}{\partial u_{i,\alpha}} \frac{\partial \rho(\boldsymbol{r})}{\partial u_{j,\beta}} + \rho(\boldsymbol{r}) \frac{\partial^2 V(\boldsymbol{r})}{\partial u_{i,\alpha} \partial u_{j,\beta}} \right] \mathrm{d}^3 \boldsymbol{r} \tag{24}$$

where $\rho(\boldsymbol{r})$ is the electron density and $V(\boldsymbol{r})$ the potential. The key simplifying point is that the exchange correlation and Hartree parts of the potential are independent of $u_{i,\alpha}$ and do not contribute to first order, so for practical application $V(\boldsymbol{r})$ reduces to the ionic (pseudo)potential. The first term gives the response of the electron density to the displacement of atom j in the β direction.

With periodic boundary conditions on a single unit cell, the $u_{i\alpha}$ refer to each atom and all its periodic images. However, within linear response theory the displacements of the images can be modulated with a phase factor corresponding to an arbitrary phonon wavevector $\boldsymbol{q}$

$$u^l_{j,\beta} = u_{j,\beta} \exp[\mathrm{i}\boldsymbol{q} \cdot \boldsymbol{r}_l] \tag{25}$$

where l labels the periodic image at $\boldsymbol{r}_l$.

This highlights the main practical difference between the linear response and finite displacement methods. In the latter one carries out a small number of calculations on a large supercell to evaluate the matrix of force constants Φ (assumed zero at long range). This matrix is then reduced to the dynamical matrix at a particular $\boldsymbol{q}$. By contrast linear response requires a separate calculation of the primitive unit cell *only*, repeated at each $\boldsymbol{q}$ to obtain the dynamical matrix $D(\boldsymbol{q})$ for that wavevector directly.

In ionic materials, there is a macroscopic electric field associated with the LO phonon in the long-wavelength limit. The energy associated with this is suppressed in the Ewald sum, and must be reintroduced through Born effective charges Z^* and the high-frequency dielectric constant ϵ_∞:

$$\Phi_{i,\alpha,j,\beta} = \Phi^0_{i,\alpha,j,\beta} + \frac{4\pi e^2}{V} \frac{(\boldsymbol{q} \cdot \boldsymbol{Z}^*_i)_\alpha (\boldsymbol{q} \cdot \boldsymbol{Z}^*_j)_\beta}{\boldsymbol{q} \cdot \epsilon_\infty \cdot \boldsymbol{q}} \tag{26}$$

where Φ^0 is the force constant calculated from a standard electronic structure calculation with the long-range electrostatic terms assumed to cancel. This correction to the LO phonon can be implemented in both linear response [36] and finite displacements [37] with additional calculations to evaluate Z^*_i and ϵ_∞.

[18] Φ' is equivalent to the dynamical matrix for $\boldsymbol{q} = 0$ for the supercell.

[19] or $\Phi'_{i,\alpha,j,\beta}/M$ to each of the nearest shell of M images at half the supercell dimension, see [34].

6.3. Treatment of anharmonicity

A single density of states calculation, by whatever method, can be used to evaluate the free energy at one state point. For an equation of states, one needs to evaluate multiple state points. In practice, this involves interpolation. The best quantity to interpolate is the most fundamental, $\Phi(V)$, rather than $g(\omega)$ or ω_i since this will enable crossing of branches and symmetry repulsion to be properly incorporated, and require no additional *ab initio* calculation.

Once the harmonic approximation breaks down the quasiharmonic methodology fails. At high temperatures, this is manifest by the dependence of the phonon frequencies with temperature. The two methods suggest different solutions—in *ab initio* lattice dynamics the atomic displacements from which the restoring forces are calculated can be increased to represent the actual displacements encountered at a given temperature. This requires more than one calculation for each displacement, and results in temperature dependence of the force constants or use of anharmonic oscillators [8] which then propagate through the density of states to the free energy. Within linear response theory the third-order perturbation [38][20] could be used similarly.

In one case, the quasiharmonic approximation fails dramatically. This is for soft phonons, as described in section 3. If the Landau expansion is associated with a phonon mode, the quasiharmonic free energy diverges to $-\infty$ as the quadratic term goes to zero. This is totally at variance with the Landau picture that the phase is becoming *unstable* at that point. Within the quasiharmonic approach this problem can be addressed by treating the soft modes as quartic potential[21] and populating them according to the Schrödinger solution for that potential, but this raises problems of determining which modes to afford special treatment.

7. Thermodynamic integration

The thermodynamic integration approach gives a method of calculating the exact free energy of a system by integrating along a path which slowly changes the description of the forces from the *ab initio* to a Hamiltonian for which the entropy can be calculated exactly (e.g. an Einstein crystal).

For a partition function, Z, we can write

$$F = -k_B T \log Z \tag{27}$$

whence

$$F_2 - F_1 = -k_B T \log Z_2/Z_1. \tag{28}$$

Now, if we have two Hamiltonians acting in the same phase space, the difference in their free energies is

$$Z_2/Z_1 = \frac{\int \exp(-E_2/k_B T)\, \mathrm{d}\{\boldsymbol{r}_i\}\, \mathrm{d}\{\boldsymbol{p}_i\}}{\int \exp(-E_1/k_B T)\, \mathrm{d}\{\boldsymbol{r}_i\}\, \mathrm{d}\{\boldsymbol{p}_i\}} \tag{29}$$

which can be rewritten as

$$Z_2/Z_1 = \int \exp[-(E_2 - E_1)/k_B T] P_1\, \mathrm{d}\{\boldsymbol{r}_i\}\, \mathrm{d}\{\boldsymbol{p}_i\} \tag{30}$$

where $P_1 = \exp(-E_1/k_B T)/Z_1$ is the Boltzmann probability for the state under Hamiltonian 1, and in practice the integral over momenta $\{\boldsymbol{p}_i\}$ will cancel and need not be evaluated.

[20] Which requires no more computation than the second order.

[21] Or any other W shaped function, such as a Gaussian plus a quadratic. Drummond N D and Ackland G J 2002 *Phys. Rev.* B at press.

Thus the change in free energy going from one Hamiltonian to another can be written as the log of the ensemble averaged expectation value of the Boltzmann factor of the energy difference:

$$\Delta F = -k_B T \log\langle \exp(E_1 - E_2/k_B T)\rangle \tag{31}$$

where the ensemble average is over the initial Hamiltonian probabilities.

A practical problem with phase space sampling can arise if the Hamiltonians are very different: the configurations looked in 1 may not be the important ones in 2. To get round this, thermodynamic integration uses a continuum of Hamiltonians, described by a parameter λ. Typically, for an *ab initio* (AI) to Einstein (E) crystal integration we define

$$H(\lambda) = \lambda H_{AI} + (1 - \lambda) H_E. \tag{32}$$

Whence

$$F_{AI} - F_E = \int_0^1 \frac{\partial F(\lambda)}{\partial \lambda}\, \mathrm{d}\lambda. \tag{33}$$

Thermodynamic integration involves evaluating this integral numerically, typically by evaluating

$$\Delta F = -k_B T \sum \frac{\log\langle \exp[-(E(\lambda_i) - E(\lambda_i + \delta\lambda))/k_B T]\rangle_i}{\delta\lambda} \Delta\lambda_i. \tag{34}$$

For *ab initio* MD coupled to an Einstein crystal the MD simulation is carried out with $H(\lambda)$, which gives a correct thermodynamic sampling of that phase space as required for the ensemble average. The value of $E(\lambda_i + \delta\lambda)$ can be trivially evaluated.

The simulation works best when ΔF is as small as possible, so the Einstein frequency should be chosen as close as possible to that of the real material. If a harmonic crystal is used as reference state in place of the Einstein crystal, it is possible to obtain a still more accurate approach since for each phonon eigenstate $[E(\lambda_i) - E(\lambda_i + \delta\lambda)]$ is better reproduced. Note that the method requires numerical integration, so the integral must be continuous, i.e. the change in Hamiltonian should not be accompanied by a phase transition.

Although in broad use for several years with simple energy models, this method was first applied to *ab initio* simulation by Sugino and Car [39] for the melting curve of silicon using isothermal [40]–isobaric [41] MD, pseudopotentials, LDA and DFT with the Stillinger–Weber potential [42] as a reference. In contrast to previous work, they used a dynamic method with a slowly time-varying λ. This meant that the integral (now over enthalpy, H) was evaluated over a single 0.7 ps MD trajectory for each combination of P and T:

$$\Delta G = \int_{t_0}^{t_1} H(\lambda = 1) - H(\lambda = 0) \frac{\mathrm{d}\lambda}{\mathrm{d}t}\, \mathrm{d}t. \tag{35}$$

For silicon, with a 0.06 eV/atom correction for zero-point fluctuations in the crystal, the melting point was calculated to be some 300 K (20%) too low, an error resulting from about 0.1 eV/atom error in the free energy.

A similar calculation by Alfe *et al* [43] evaluated energies for a system in which the melting point is unknown: iron under the extreme conditions of temperature and pressure found in the Earth's core. They claimed accuracy to within 10% in their prediction of a 6670 K melting point at 330 GPa (Earth's inner core boundary). The higher accuracy of the second calculation may arise because silicon exhibits metallization and a significant volume collapse on melting, whereas the density change in iron is small and the character of the material remains as a non-magnetic metal. Thus there will be greater cancellation of errors in the exchange correlation functional for the more homogeneous transformation [44].

8. Dual space Monte Carlo

A class of methods providing an alternative to thermodynamic integration for evaluating free energy differences between distinct macrostates is suggested by lattice switch MC [45]. Rather than integrating from each state separately to a reference state, then comparing free energies, one simply constructs a dual phase space comprising the two macrostates and an unphysical MC step which take the simulation from one to the other. Provided a process can be run which actually samples each macrostate, the finite free energy difference between them can be evaluated directly from their relative statistical weight, with accumulating integration errors entailed in measuring absolute free energies. Three variants are described here, switching crystal structures, Hamiltonians and chemical species.

8.1. Lattice switch Monte Carlo

Lattice switch MC [45] is a method for calculating free energy differences between two crystal structures separated by a large kinetic barrier. The method works by sampling the dual phase space of the two phases, and introducing a MC step which takes the simulation between phases.

Mathematically the method is straightforward: for a system of N particles in the canonical ensemble the spatial coordinates $\boldsymbol{r}$ are written

$$\boldsymbol{r}_i = \boldsymbol{R}_i + \boldsymbol{u}_i \tag{36}$$

where the vectors $\boldsymbol{R}_i$, $i = 1 \ldots N \equiv \{\boldsymbol{R}\}_\alpha$ define the sites of a lattice of type α

The partition function associated with α is

$$Z(N, V, T, \alpha) = \int_{\{\boldsymbol{r}_i\}_\alpha} \exp[-U(\{\boldsymbol{u}\}, \alpha)/k_B T][\mathrm{d}\{\boldsymbol{r}_i\}] \equiv \exp[-F_\alpha(N, V, T)/k_B T] \tag{37}$$

where U represents the configurational energy and the integral is taken over those $\{r_i\}$ which describe phase α.

To facilitate the lattice switch the zero-temperature energies of the two structures are shifted to be equal. Defining the energy difference

$$\Delta U_{12} = U(N, V, T = 0, \{\boldsymbol{u}\} = 0, \alpha = 1) - U(N, V, T = 0, \{\boldsymbol{u}\} = 0, \alpha = 2)$$

between the structures at $T = 0$ as determined by static relaxation, the energies actually used in the simulation are adjusted to

$$\begin{aligned} &\tilde{U}(N, V, T = 0, \{\boldsymbol{u}\}, 1) = U(N, V, T = 0, \{\boldsymbol{u}\}, 1) \\ &\tilde{U}(N, V, T = 0, \{\boldsymbol{u}\}, 2) \equiv U(2, N, V, T = 0, \{\boldsymbol{u}\}, 2) - \Delta U_{12}. \end{aligned} \tag{38}$$

This bias can be removed in the final calculation of free energy differences by a simple shift of $\Delta F_{12}(T = 0) = \Delta U_{12}$.

The probability of the system being in structure α is then

$$P(\alpha|N, V, T) \equiv \int_{\{\boldsymbol{u}\}\in\alpha} P(\{\boldsymbol{u}\}, \alpha|N, V, T)[\mathrm{d}\{\boldsymbol{r}_i\}] = \frac{Z(N, V, T, \alpha)}{Z(N, V, T, 1) + Z(N, V, T, 2)}. \tag{39}$$

The free energy difference between two structures ($\alpha = 1, 2$) can then be written

$$F_1(N, V, T) - F_2(N, V, T) \equiv N\Delta f(V, T) = k_B T \log \mathcal{R} + \Delta U_{12} \tag{40}$$

where,

$$\mathcal{R} = \frac{Z(N, V, T, 2)}{Z(N, V, T, 1)} = \frac{P(2|N, V, T)}{P(1|N, V, T)}. \tag{41}$$

Generalization to the constant pressure ensemble is straightforward, one can include the lattice vectors as MC variables in a Parrinello–Rahman scheme with $\tilde{U}$ becoming the

enthalpy [46], or one can tabulate free energy differences per particle $\Delta f(V, T)$ and construct $\Delta f(P, T)$ using elementary thermodynamics [30] and interpolation. Alternately, a number of strategies exist to follow a line (such as a phase boundary) through phase space.

Practically, the problem is that equation (40) is useful *only* if one can devise a MC procedure that will actually visit the configurations $\{\boldsymbol{u}\}, \alpha$ with the probabilities prescribed by equation (39). In order to achieve adequate sampling, an interphase step (the lattice switch) has to be devised which transforms the $\{\boldsymbol{R}_1\}$ to $\{\boldsymbol{R}_2\}$. In the original scheme [47,48], as applied to hard spheres or Lennard-Jones potentials, a biased sampling was used to move the system towards so-called *gateway* states from which the lattice switch move could be launched with a reasonable probability of success.

In application to *ab initio* systems a different approach is required. The computational cost of evaluating U far outweighs other considerations in the problem. Thus one can contemplate far more complex lattice switch moves than were appropriate with simple potentials.

A promising scheme [49] involves a combination of the lattice switch method with static relaxation and harmonic lattice dynamics. Since the integrals over the phases are independent, we are free to generalize equation (36)

$$\begin{aligned} \boldsymbol{r}_{i,1} &= \boldsymbol{R}_{i,1} + \boldsymbol{u}_i \\ \boldsymbol{r}_{i,2} &= \boldsymbol{R}_{i,2} + \boldsymbol{A}_{ij}\boldsymbol{u}_j \end{aligned} \tag{42}$$

where $\boldsymbol{A}_{ij}$ is any matrix which converts the displacements in one structure to those in the other. $\boldsymbol{A}_{ij}$ is arbitrary, and can be chosen to facilitate the lattice switch. Provided $\boldsymbol{A}_{ij}\boldsymbol{u}_j$ spans the phase space of phase 2, $\boldsymbol{A}_{ij}$ may be non-unitary provided the appropriate correction is made to the free energy difference, $\Delta F_A = -k_B T \log\det \boldsymbol{A}_{ij}$.

An efficient method is to construct $\boldsymbol{A}_{ij}$ such that in the harmonic limit,

$$\tilde{U}(\{\boldsymbol{u}\}, 1) - \tilde{U}(\{\boldsymbol{u}\}, 2) = 0.$$

This is done by first removing the $T = 0$ energy difference (obtaining $\tilde{U}$), then pairing the vibrational modes of each structure according to their energy. In the lattice switch, the amplitude $\xi_i^\alpha = \boldsymbol{e}_{ij,\alpha}\boldsymbol{u}_{j,\alpha}$, of each eigenmode in one structure is transformed to an amplitude $\xi_i^1 = \xi_i^2[\omega_i^2/\omega_i^1]$ of the paired mode in the other structure, such that the harmonic energy in the mode is the same in each lattice.

Thus the lattice switch transformation (the MC trial step taking the simulation between structures) is

$$\begin{aligned} \boldsymbol{R}_{i,1} &\to \boldsymbol{R}_{i,2} \\ \boldsymbol{u}_{i,1} &\to \boldsymbol{A}_{ij}\boldsymbol{u}_{j,2}. \end{aligned} \tag{43}$$

Assuming the two phases have harmonic eigenenergies $\epsilon_{i,\alpha}$ and a matrix of normalized eigenvectors $\boldsymbol{e}_{ij,\alpha}$, then $\boldsymbol{A}$ can be written

$$\boldsymbol{A} = \frac{\boldsymbol{e}_{ij,1}}{\omega_{i,2}}\,\frac{\boldsymbol{e}_{ij,2}^{-1}}{\omega_{i,1}}. \tag{44}$$

The three zero-frequency eigenmodes corresponding to centre of mass motion can be eliminated from the sets of $u_{i,\alpha}$ and the non-unitarity free energy correction becomes the harmonic free energy difference. This can be obtained using the methods of linear response or lattice dynamics outlined above.

Hence the free energy difference can be broken into three terms, each of which can be evaluated separately. The ground state energy difference can be calculated by static relaxation. The harmonic energy (including the zero-point) can be calculated using lattice dynamics to evaluate phonons, and quasiharmonic thermodynamics to evaluate the free energy associated with the (quantised) phonons. Finally, the lattice switch MC method can be used to evaluate directly the anharmonic contribution to the free energy difference.

8.2. Hamiltonian switch Monte Carlo

The formalism of lattice-switch MC can be exactly transformed to performing a switch between Hamiltonians if α (in equation (37)) labels the Hamiltonian, and the switch move becomes a change of Hamiltonian with $\{\boldsymbol{r}_i\}$ remaining fixed[22]. Now the space over which the simulation runs is a dual space of position (continuous) and Hamiltonian (discrete). The free energy difference is determined by how much time the simulation spends in each of the discrete α. In order to evaluate the *ab initio* free energy difference between two structures it is now necessary to perform three simulations, two with Hamiltonian switch and one with the lattice switch (see figure). The advantage is that the second Hamilitonian, typically an interatomic potential, can be chosen arbitrarily to minimize the free energy associated with the Hamiltonian switch (these are the more time consuming calculations, since they involve *ab initio* energy evaluation), such that the major contribution comes from the lattice switch, which does not entail *ab initio* evaluation. There is no need for the second Hamiltonian to be the same for each structure: switch 3 (between non-electronic structure models, see figure) can entail a switch of both lattice sites and Hamiltonian.

HSMC differs from thermodynamic integration in that the accept/reject criterion is based on the energy from both Hamiltonians (for the switch or biassed sampling moves) and one or other for the particle moves. The integration method bases its accept/reject criterion on one Hamilitonian only. It has no equivalent to the switch move. As with LSMC, convergence of the phase space integrals measured by HSMC depends on enacting the switch move a statistically significant number of times.

8.3. Chemical switch Monte Carlo

A similar method can be used for evaluating alloying energies. Here the switch move changes the chemical species of one of the atoms in the simulation. α in equation (37) now labels the atomic species as well as the co-ordinates.

This fits especially well with the pseudopotential method, since all that need be done is to exchange the pseudopotential at a site for one describing a similarly charged atom. Again, to achieve a reasonable accept/reject ratio for the switch one can eliminate harmonic effects by precalculating $\tilde{U}$, and incorporate local relaxation by transforming through normal mode coordinates. In evaluating free energy differences between chemically different materials it is necessary to incorporate the chemical potentials of the switched ions.

Substituting ions with different charges can be done similarly, with different numbers of electrons in each simulation being compensated by $\tilde{U}$. Again, relating results to a real system requires knowledge of chemical potentials. Calculation of vacancy formation energies is also tractible with this method.

Previous work using integration approaches for calculating the free energies of defects [50–52] has had to formulate imaginative methods for continuously decoupling the atom to be switched from the rest of the system, introducing problems with, for example, fractional electrons. This is tractible within the density functional approach but direct switching methods finesse such problems.

9. Applications

9.1. Application to $MgSiO_3$

Magnesium silicate perovskite $MgSiO_3$ is the most studied system with *ab initio* thermodynamics. As the major constituent of the Earth's mantle, where it exists at high

[22] If the two Hamiltonians have different harmonic properties, then equation (43) with $\boldsymbol{R}_{i,1} = \boldsymbol{R}_{i,2}$ can be applied.

pressure and temperature, its properties are crucial to geophysicists and extremely difficult to measure experimentally. The issues are both qualitative and quantitative. Firstly, perovskite materials are known to undergo a series of soft mode phase transitions from orthohombic–tetragonal–cubic: the qualitative question is whether these transitions can occur under mantle conditions[23]. Secondly, quantitative information about the elastic moduli is an essential input to seismological models, and again *ab initio* thermodynamics offers a solution.

Static lattice calculation [57] shows that at 0 K for all compressions the orthorhombic structure is favoured, and thus if other phases exist the higher symmetry modes they must be dynamically stabilized[24]. Since the barrier to rotation of the octahedra in the perovskite structure was much higher than any plausible $k_B T$ found in the mantle, and increases with pressure it was concluded that only the orthorhombic phase could exist [57]. Calculation of phonons across the Brillouin zone [58] allowed parametrization of an interacting double-well mode [11]. This model predicted a significantly lower transition temperature, but still out with the most likely predictions for the geotherm[25], moreover on descending the mantle, the stabilizing effect of pressure on the orthorhombic phase outweighs the destabilizing effect of temperature.

The orthorhombic perovskite and MgO elastic moduli were evaluated for $T = 0$ and high pressure [59, 60] and high temperature using the quasiharmonic approach [61, 81]. The first *ab initio* MD on $MgSiO_3$ showed in detail the soft phonon mechanism in action [11, 53], and suggested that even up to 1720 K (above the melting point at ambient pressure) the orthorhombic phase remained stable.

Single phase equation of state calculations at high pressure and temperature for orthorhombic $MgSiO_3$ followed both with MD [55] and with quasiharmonic lattice dynamics [54, 62, 63]. These can elucidate both the equation of state and the P–T dependence of the elastic moduli: in all cases the orthorhombic–tetragonal phase transition appears to lie at temperatures above those found in the Earth. These results agree well with experiment where experimentation is possible [64], and now represent the best data available to Earth system modellers for comparison with earthquake wave propagation to determine whether other minerals (notably iron-rich perovskite or aluminous phases [56]) might be present in the mantle.

Thus in the past 8 years understanding of the behaviour of perovskite under mantle conditions has gone from a lack of even a qualitative understanding of the phase stability to a detailed knowledge of the full equation of states. A remaining challenge in modelling mantle thermodynamics is to calculate decomposition phase boundaries between forsterite, perovskite and periclase.

9.2. Application to iron

Iron under the conditions of extreme T and P found at the Earth's core is an ideal system for *ab initio* free energy calculation. Experiments on the melting locus are extremely difficult to perform, and the complications of iron calculations arising from ferromagnetism are eliminated by the extreme pressure. The crystalline phase (inner core) is believed to be hexagonally close packed. The crucial region for the data is at the inner-core boundary, where the pressure is 330 GPa.

[23] An issue somewhat clouded by uncertainty in the exact pressure–temperature conditions of the mantle.

[24] i.e. the mean position of the atoms would not represent a minimum of the total energy: the increased symmetry would result from equal sampling of two or more asymmetric configurations.

[25] The *geotherm*, a line in pressure–temperature space reflecting actual conditions as one goes into the mantle, is not precisely known, these calculations did not rule out the existence of the tetragonal phase at the top of the lower mantle for the hottest geotherm.

The results from Alfe *et al* [43] are truly *ab initio*: they used DFT–GGA with the projector augmented wave [69, 70] all-electron technique to include effects of high pressure on the 3s and 3p shells. The free energies of the two phases were derived by thermodynamic integration to a Hamiltonian defined by a repulsive pair potential ($V(r) = A/r^{\alpha}$) with A and α optimized to minimize $H(\lambda = 1) - H(\lambda = 0)$. The same values for A and α were found to be satisfactory throughout. The calculations were restricted to runs of around a picosecond, and cells containing around 100 atoms, but finite size effects between 60 and 120 atoms were undetectable.

The results from Laio *et al* [23] also used DFT–GGA with a pseudopotential, but included the 3s and 3p electrons explicitly in the calculation. They include the uncontrolled approximation of fitting an embedded atom-type potential [25] to 64 atom *ab initio* calculations at each P and T, and using this to evaluate the free energy. The use of the potential allows for elimination of finite size effects[26]. The potential parameters changed significantly as a function of P and T.

The discrepancy between these two calculations is more than 1000 K. Partly, this illustrates the extreme difficulty of comparing two similar free energies: for iron a 1000 K difference can arise from a 0.1 eV error in either phase. For less sensitive quantities such as the bulk moduli, radial distribution function etc, the agreement is much better. It is worth noting that the experimental uncertainty in composition (e.g. sulphur concentration) and viscosity is huge, and even the present level of accuracy found in these calculations is able to provide closer bounds on these quantities [71, 72].

9.3. Application to silicon under pressure and shock waves

Among the earliest total energy calculations was the work of Yin and Cohen [73] predicting successfully the phase transition from diamond to β-tin structure in silicon [73][27]. A huge amount of static lattice 0 K work on other high-pressure and metastable phases followed [74–80], then in 1994 came the first direct *ab initio* MD simulation of a phase transition at constant pressure [81, 82][28]. Latterly, melting in silicon was the subject of the first *ab initio* thermodynamic integration simulation [39] and the implementation of 'on the fly' parametrization of a Stillinger–Weber potential from a tight-binding Hamiltonian [28].

The agreement between all these static free energy calculations and experiment inspired a completely *ab initio* treatment of shock waves in polymorphic silicon [31]. Here the speed u_s of a shock wave with respect to the undisturbed material ahead is given by the Rankine–Hugoniot equations [83]

$$u_s^2 = v_0^2 \frac{P - P_0}{v_0 - v} \tag{45}$$

$$(v - v_0)(P + P_0)/2 = u - u_0 \tag{46}$$

where v is the reciprocal mass density and u the specific internal energy.

If $(P - P_0)/(v_0 - v)$ does not increase monotonically with material density, as at a first-order phase transition, then a shock of some pressure P_1 may split into a pair of shocks.

[26] Excluding finite size effects arising from fitting to a small simulation.

[27] ultimately, it has been found that the first high-pressure phase is not the β-tin investigated there, but the $Imma$ structure. This illustrates a perennial problem with *ab initio* free energy calculation: one can only calculate the electronic energy of the postulated ionic structure: there is no guarantee that this structure is the global minimum of energy.

[28] In a curious echo of Yin and Cohen's work [73], these authors also found the 'wrong' phase, simple hexagonal rather than $Imma$, in this case due to inadequate k-point sampling.

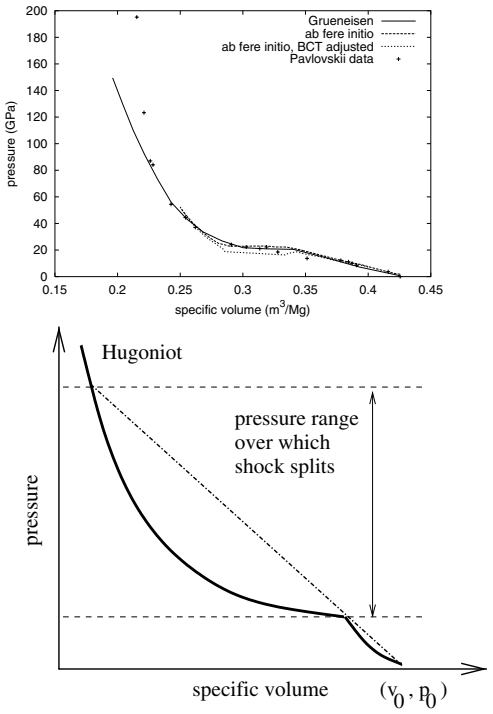

Figure 9. Lower: schematic equation of states with convex region: in the pressure range shown the shock wave splits into rapid preshock in diamond-structure silicon and slower transformation wave in transformed material. Upper: calculated Hugoniot equation of states for silicon, showing convex region leading to shock splitting and compared to shock wave data [84]. Experimental transient x-ray diffraction data [31] shows shock splitting of about 20 GPa, consistent with the simple convexity argument and with finite element shock wave simulation using the *ab initio* Hugoniot.

Ab initio calculation of equation of states, combined with finite element shock wave analysis, has been used to model this process.

The faster shock moves at the maximum speed in the untransformed material for $P < P_1$, and with the corresponding pressure close to the transition pressure. The second shock which completes the compression to P_1 is slower, and moves through the transformed material (figure 9). The onset of shock wave splitting under given pre-shock conditions of P and T can be calculated once the polymorphic equation of states is known. Calculations for silicon have been compared with shockwave data from the Los Alamos TRIDENT laser, and with the crystal structure monitored using nanosecond spectroscopy with an x-ray streak camera. Above 20 GPa, the shock wave splits into two, and the crystal structure is observed to change. This is a significantly higher pressure than the static phase transition, indicating that the shock wave propagates through metastable diamond structure silicon, but is captured by using the *ab initio* equation of states in conjunction with a Lagrangian hydrocode [85] to investigate the shock wave structure, by simulating the propagation of a planar shock wave through a sample of silicon, driven by a constant applied pressure at one side.

10. Conclusions

A large number of methods for evaluating free energies from *ab initio* simulation exist. Many of these are borrowed from classical potential simulations, but some new methods which explicitly take account of the computational expense of evaluating the *ab initio* energy of a microstate are appearing.

Simulation of properties of single phases, where the properties of interest are differentials of a differentiable free energy, can be evaluated directly from MD simulations along a slowly varying trajectory through *PVT* space. Finite size problems exist, but can be addressed by judicious use of analytic potentials.

Phase transformations are more difficult to calculate, since the energy of the stable phase in *PVT* space is discontinuous. In the low-T regime quasiharmonic lattice dynamics allows for quantization of vibrations and hence direct free energy evaluation. For liquids, or anharmonic solids, free energies can be evaluated by thermodynamic integration along a reversible path through phase space leading from one phase to another (possibly via a gradual change in Hamiltonian). Alternately, free energy differences between two phases can be measured directly by constructing a dual phase space which incorporates regions corresponding to each, and devising an MC procedure which samples each region.

Ab initio free energy calculation has found an important niche in determining thermodynamic data for materials in regions of extreme pressure, temperature or strain which are inaccessible to experimental probes. It is likely to become the method of choice for such studies in the future.

Acknowledgments

The author would like to thank A D Bruce, C Verdozzi, M W Finnis, D Alfe, A R Oganov, D C Swift, D Bird, A de Vita, M J Gillan, M C Warren, S J Clark, A Laio and A N Jackson for illuminating discussions.

References

[1] Hellmann H 1937 *Einfuhrung in Die Quantumchemie* (Leipzig: Franz Deutsche)
Feynman R P 1939 *Phys. Rev.* **56** 340
[2] Pulay P 1969 *Mol. Phys.* **17** 197
[3] Car R and Parrinello M 1985 *Phys. Rev. Lett.* **55** 2471
[4] Payne M C, Teter M P, Allan D C, Arias T A and Joannopoulos J D 1992 *Rev. Mod. Phys.* **64** 1045
[5] Cox S G and Ackland G J 2000 *Phys. Rev. Lett.* **84** 2362
[6] Allen M P and Tildesley D J 1987 *Computer Simulation of Liquids* (Oxford: Clarendon)
[7] Ackland G J and Verdozzi C 1999 New advances in materials prediction *Mater. Res. Soc. Symp. (Warrendale, PA: USA)* D **6** 4
[8] Drummond N D and Ackland G J 2001 *Phys. Rev.* B submitted
Drummond N D and Ackland G J 2002 *Phys. Rev.* B at press
[9] Bruce A D 1980 *Adv. Phys.* **29** 111
[10] Padlewski S, Evans A K, Ayling C and Heine V 1992 *J. Phys.: Condens. Matter* **4** 4895
[11] Warren M C 1997 *PhD Thesis* University of Edinburgh
[12] Thomas H 1971 *Structural Phase Transitions and Soft Modes* Universtitets Forlaget (Oslo) p 15
[13] Ackland G J 2000 *RIKEN Rev.* **29** 34
[14] Bird D M and Gravil P A 1997 *Sur. Sci.* **377** 555
[15] White J A, Bird D M, Payne M C and Stich I 1994 *Phys. Rev. Lett.* **73** 1404
[16] Mills G and Jónsson H 1994 *Phys. Rev. Lett.* **72** 1124
[17] Kroes G J 1999 *Prog. Surf. Sci.* **60** 1
[18] Mills G, Jonsson H and Schenter G 1995 *Surf. Sci.* **324** 305

Jónsson H, Mills G and Jacobsen K H 1998 *Classical and Quantum Dynamics in Condensed Phase Simulations* ed B J Berne, G Ciccotti and D F Coker (Singapore: World Scientific) p 385
[19] Henkelman G and Jónsson H 2000 *J. Chem. Phys.* **113** 9978
[20] Zhong W, Vanderbilt D and Rabe K M 1995 *Phys. Rev.* B **52** 6301
[21] Zhong W, Vanderbilt D and Rabe K M 1994 *Phys. Rev. Lett.* **73** 1861
[22] He L X and Vanderbilt D 2001 *Phys. Rev. Lett.* **86** 5341
[23] Laio A, Bernard S, Chiarotti G L, Scandolo S and Tosatti E 2000 *Science* **287** 1027
[24] Cohen M H and Heine V 1961 *Phys. Rev.* **122** 1821
[25] Baskes M I 1992 *Phys. Rev.* B **46** 2727
[26] Ercolessi F and Adams J B 1994 *Europhys. Lett.* **26** 583
[27] de Vita A 2001 private communication
[28] de Vita A 1998 Tight-binding approach to computational materials science *Mater. Res. Soc. Symp. (Warrendale, PA: USA)* xiii+541 473
[29] Mandl F 1988 *Statistical Physics* 2nd edn (New York: Wiley)
[30] Swift D C 2000 *PhD Thesis* The University of Edinburgh
[31] Swift D C, Ackland G J, Hauer A, Wark J S, Kalantar D H, Kyrala G A and Kopp R 1998 *40th APS Plasma Physics Group Meeting (New Orleans, USA, Nov. 1998)*
[32] Dewaele A, Fiquet G and Gillet P 1998 *Rev. Sci. Instrum.* **69** 2421
[33] Born M and Huang K 1956 *Dynamical Theory of Crystal Lattices* (Oxford: Oxford University Press)
[34] Ackland G J, Warren M C and Clark S J 1997 *J. Phys.: Condens. Matter* **9** 7861
[35] Baroni S, Giannozzi P and Testa A 1987 *Phys. Rev. Lett.* **58** 1861
[36] Baroni S, de Gironcoli S, Dal Corso A and Giannozzi P 2001 *Rev. Mod. Phys.* **73** 515
[37] Parlinski K Lazewski J and Kawazoe Y 2000 *Phys. Chem. Solids* **61** 87
[38] Debernardi A, Baroni S and Molinari E 1995 *Phys. Rev. Lett.* **75** 1819
[39] Sugino O and Car R 1995 *Phys. Rev. Lett.* **74** 1823
[40] Nose S 1984 *J. Chem. Phys.* **72** 2384
[41] Parrinello M and Rahman A 1980 *Phys. Rev. Lett.* 45
[42] Stillinger F A and Weber T A 1985 *Phys. Rev.* B **31** 5262
[43] Alfe D, Gillan M J and Price G D 1999 *Nature* **401** 462
[44] Alfe D 2001 private communication
[45] Bruce A D, Wilding N B and Ackland G J 1997 *Phys. Rev. Lett.* **79** 3002
[46] Ackland G J, Wilding N B and Bruce A D 1997 *Mater. Res. Soc. Symp. Proc.* **499** 253
[47] Bruce A D, Jackson A N, Ackland G J and Wilding N B 2000 *Phys. Rev.* E **61** 906
[48] Jackson A N 2001 *PhD Thesis* University of Edinburgh
[49] Acharya A, Bruce A D and Ackland G J, in preparation
[50] Smargiassi E and Madden P A 1995 *Phys. Rev.* B **51** 117
[51] Smargiassi E and Madden P A 1995 *Phys. Rev.* B **51** 129
[52] Smargiassi E and Car R 1996 *Phys. Rev.* B **53** 9760
[53] Warren M C, Ackland G J, Karki B B and Clark S J 1998 *Mineral. Mag.* **62** 585
[54] Oganov A R, Brodholt J P and Price G D 2000 *Phys. Earth Planet Inter.* **122** 277
[55] Oganov A R, Brodholt J P and Price G D 2001 *Earth Planet. Sci. Lett.* **184** 555
[56] Oganov A R and Broadholt J P 2000 *Phys. Chem. Miner.* **27** 430
[57] Stixrude L and Cohen R E 1993 *Nature* **364** 613
[58] Warren M C and Ackland G J 1996 *Phys. Chem. Miner.* **23** 107
[59] Karki B B, Stixrude L, Clark S J, Warren M C, Ackland G J and Crain J 1997 *Am. Mineral.* **82** 635
[60] Karki B B, Stixrude L, Clark S J, Warren M C, Ackland G J and Crain J 1997 *Am. Mineral.* **82** 51
[61] Karki B B, Wentzcovitch R M, de Gironcoli S and Baroni S 2000 *Phys. Rev.* B **61** 8793
[62] Karki B B, Wentzcovitch R M, de Gironcoli S and Baroni S 2000 *Phys. Rev.* B **62** 14 750
[63] Parlinski K and Kawazoe Y 2000 *Eur. Phys. J.* B **16** 49
[64] Fiquet G, Dewaele A, Andrault D, Kunz M and Le Bihan T 2000 *Geophys. Res. Lett.* **27** 21
[65] Boehler R 1993 *Nature* **363** 534
[66] Shen G, Mao H, Hemley R J, Duffy T S and Rivers M I 1998 *Geophys. Res. Lett.* **25** 373
[67] Williams Q, Jeanloz R, Bass J, Svendsen B and Ahrens T J 1987 *Science* **236** 181
[68] Yoo C S, Holmes N C, Ross M, Webb D J and Pike C 1993 *Phys. Rev. Lett.* **70** 3931
[69] Perdew J P, Chevary J A, Vosko S H, Jackson K A, Pederson M R, Singh D J and Fiolhais C 1992 *Phys. Rev.* B **46** 6671
[70] Blöchl P E 1994 *Phys. Rev.* B **50** 17 953
[71] Alfe D, Gillan M J and Price G D 2000 *Nature* **405** 172

[72] deWijs G A, Kresse G, Vocadlo L, Dobson D, Alfe D, Gillan M J and Price G D 1998 *Nature* **392** 805
[73] Yin M T and Cohen M L 1982 *Phys. Rev.* B **26** 5668
[74] Needs R J and Martin R M 1984 *Phys. Rev.* B **30** 5390
[75] Chang K J and Cohen M L 1985 *Phys. Rev.* B **31** 7819
[76] Biswas R, Martin R M, Needs R J and Neilsen O H 1987 *Phys. Rev.* B **35** 9559
[77] Lewis S P and Cohen M L 1993 *Phys. Rev.* B **48** 16 144
[78] Clark S J, Ackland G J and Crain J 1994 *Phys. Rev.* B **49** 5341
[79] Crain J, Ackland G J, Maclean J R, Piltz R O, Hatton P D and Pawley G S 1994 *Phys. Rev.* B **50** 13 043
[80] Christensen N E 1998 *Semicond. Semimet.* **54** 49
[81] Focher P, Chiarotti G L, Bernasconi M, Tosatti E and Parrinello M 1994 *Eur. Lett.* **26** 345
[82] Bernasconi M, Chiarotti G L, Focher P, Scandolo S, Tosatti E and Parrinello M 1995 *J. Phys. Chem. Solids* **56** 501
[83] Skidmore I C 1965 An introduction to shock waves in solids *Applied Materials Research* pp 131–47
[84] Pavlovskii M N 1968 *Sov. Phys.–Solid State* **9** 11
[85] Benson D J 1992 *J. Comput. Methods Appl. Mech. Eng.* **99** 235
[86] Karki B B 2000 *Am. Mineral.* **85** 1447