Studying the energy hypersurface of continuous systems - the threshold algorithm

1996 J. Phys.: Condens. Matter 8 143

(http://iopscience.iop.org/0953-8984/8/2/004)

J. Phys.: Condens. Matter **8** (1996) 143–156. Printed in the UK

Studying the energy hypersurface of continuous systems—the threshold algorithm

J C Schön, H Putz and M Jansen

Institut für Anorganische Chemie, Universität Bonn, Gerhard-Domagk-Strasse 1, D-53121 Bonn, Germany

Received 21 March 1995, in final form 28 September 1995

Abstract. A new method is presented for the study of the structure of the energy hypersurface of continuous systems. This so-called threshold algorithm is an adaptation of the 'lid method' introduced by Sibani and co-workers in 1993 (Sibani P *et al* 1993 *Europhys. Lett.* **22** 479–85) for the investigation of discrete energy landscapes. The algorithm produces an estimate of the local densities of states near deep-lying local minima of the potential energy of the system together with the barrier heights around these minima. This allows the computation of, for example, specific heats, the estimation of the kinetic stability of configurations that represent metastable minima of the potential energy, and the description of the relaxation behaviour of the system. As an example, a two-dimensional neon crystal is studied, where for example the calculated specific heat is found to agree with the experimental values of the three-dimensional case scaled to two dimensions.

1. Introduction

In the fields of physics, chemistry or combinatorial optimization, the energy or cost as a function of the possible states of a system or solutions of some optimization problem, respectively, is a central quantity. Knowledge of this energy function allows the application of equilibrium statistical mechanics to the evaluation of the global equilibrium properties of the system. However, the detailed behaviour of a given system as a function of time depends crucially on the dynamics that describes its trajectories through state space. This dynamics is reflected in the neighbourhood structure of the state space, i.e., the 'list' of possible states that could be reached from a given state. In physical systems, this neighbourhood structure is usually implicit in the laws and assumptions of classical or quantum mechanics, while in optimization problems we actually make a considerable effort to design a good such 'list', often called a 'moveclass', in order to improve the optimization process.

If this energy landscape is simple enough that the typical trajectories of the system can 'reach' a representative sample of the whole state space within a given amount of time, then time averages of thermodynamic quantities along such trajectories will be equal to the ensemble averages over the whole state space, i.e., the system is ergodic. On the other hand, many systems exhibit non-trivial energy landscapes; among the most prominent ones are, e.g., glasses, spin glasses, and hard optimization problems like the 'travelling salesman problem'. The behaviour of such systems has been a major area of research for a long time [2, 3, 4, 5]. They often do not fulfil the ergodic hypothesis, at least on the time-scale of interest, i.e. they are not in or near global equilibrium. For a given instance of the system, only a small (non-representative) region of the total phase space (usually

0953-8984/96/020143+14$19.50

delimited by energetic or entropic barriers) is accessible within this time, and the observed 'equilibrium' properties of the system depend on the free energy restricted to this region. Furthermore, the time evolution of such a system usually cannot be described by a simple (e.g. exponential) relaxation to global equilibrium. Instead, the dynamics (especially at low temperatures) is highly sensitive to the barrier structure of the energy landscape, leading to power-law behaviour of time correlation functions, for example.

In order to understand and predict such a system's properties, both with respect to 'local equilibrium' and dynamics, it is therefore necessary to know the barrier structure of the energy landscape, and also the local density(ies) of states for the accessible region(s) of configuration space. In recognition of this fact, several methods have been introduced in recent years for the study of the energy landscape of complex systems, like the Coulomb glass [6, 7], clusters [8, 9, 10, 11, 12, 13], spin glasses [14, 15, 16], combinatorial optimization problems [1, 17, 18, 19, 20, 21], seismology [22, 23], and the protein folding problem [24].

For those systems exhibiting a discrete configuration space, Sibani *et al* [1] have introduced the so-called 'lid method'. Starting from a deep-lying minimum, the pocket in configuration space that can be reached from the starting point without crossing a prescribed energy lid is searched exhaustively, and all the connections among the states within the pocket are noted. Thus, the local density of states, i.e. the density of states restricted to this pocket in state space, together with the energy barriers between the local minima within the pocket can be determined. On the basis of these results, the statistical mechanical properties of the system can be studied, as long as it remains within the prescribed region of phase space. In addition, this detailed knowledge of the energy landscape near a deep-lying minimum allows the calculation of the relaxation behaviour within this pocket. Originally tested on the example of the 'travelling salesman problem' [1], this procedure has since been applied to the investigation of spin glasses [14, 16].

The details and power of this algorithm depend on the discreteness of the configuration space, with its clear definition of the neighbourhood of a given state. Even in studying large spin glasses using statistical measures for the distribution of local minima [16], such a controllable configuration space and moveclass prove to be quite helpful.

In contrast to these systems, amorphous materials exhibit a continuous energy landscape. Already, many simple models based on assumptions about the barrier structure of the landscape are employed for some aspects of such systems [25, 26]. Clearly, more direct and detailed knowledge about such landscapes would prove to be highly useful. But any straightforward discretizations of their configuration spaces tend to be either prohibitively expensive in computer resources or so rough that the conclusions one would draw about the distribution of the minima and their connectivity would have to be highly tenuous.

In this paper, we address this problem by presenting an adaptation of the 'lid method' to the study of continuous energy landscapes. Again, deep-lying minima are found using global optimization methods. These minima serve as starting points for random walks below a sequence of energy lids, exploring the accessible region of configuration space. Furthermore, keeping track of the energies of the states visited during the random walk yields a sample of the local density of states.

After giving a detailed description of the algorithm in section 2, we demonstrate in section 3 how to proceed in modelling the dynamic behaviour of such a system, on the basis of the knowledge of the barrier structure and the local densities of states. Finally, in section 4, we illustrate the procedure through an example—a study of the region near the global minimum of the energy landscape belonging to a single (two-dimensional) layer of neon containing point defects. In addition to investigating the barrier structure and the

relaxation behaviour in this pocket, we show that the estimate of the density of states leads to a specific heat that agrees well with expectations.

2. The threshold algorithm

The first step of the procedure is the determination of one or more local minima of the energy landscape, x_i. This can be achieved by a large number of methods [27]; since these are well known, we will not discuss them in any detail. Each of these minima serves as a starting point for a sequence of 'threshold' runs. Just like in the 'lid method', a set of energy lids are chosen, which lie above the energy of the minimum under consideration. In the case of a discrete energy landscape, it would now be possible to begin an exhaustive search for the configurations that can be reached from the minimum along paths that do not cross the energy lid. Since in the continuous case the number of states below any lid is infinite for all practical purposes, even if one were to discretize the phase space in boxes of $h^{(N)}$, this exhaustive search procedure could not be applied directly.

Thus, instead of trying for complete information about the energy landscape within the pocket right away, we need to proceed in steps. Focusing on the determination of the energy barriers and the local densities of states, we propose to replace the exhaustive search by a statistical one in the following manner: starting from the local minimum under consideration, x_0, a random walk with a physically reasonable moveclass† is performed, where every step is accepted as long as the prescribed energy threshold, L, is not crossed. During this walk, the energy landscape is sampled, and the number of states with energy E accessible from x_0 using paths below the lid, $n(E; L, x_0)$, can be estimated up to a scale factor. If the number of samples is sufficiently large, $n(E; L, x_0)$ agrees with the density of states within the pocket, $g(E; L, x_0)$, up to the missing normalization factor. Clearly, for $L \to \infty$ this local density of states becomes identical to the global density of states, $g(E; L \to \infty, x_0) \to g(E)$, since then the pocket would encompass the whole state space. The sampling should take place at large intervals in order to exclude correlations as much as possible. Such a random walk would be termed a single 'threshold run'.

In addition, at the end of the run, the system is quenched into the nearest local minimum. Of course, additional quenches can take place during the threshold run, if so desired. For lid values just above the energy of the starting minimum, the quench will always return the system into its starting configuration. This will change upon increasing the threshold, and at some point, a quench run will end in a minimum x_1 different from the starting point x_0. The lowest lid value where this occurs is thus an upper bound for the height of the energy barrier between these two local minima.

We perform such threshold runs for all chosen values of the energy lid, using as starting points all local minima encountered during the original minimization or any of the preceding threshold runs. This is important for the determination of the barrier heights, since in most instances one tends to get better bounds by starting from a higher local minimum, which commonly 'owns' a smaller basin in phase space, i.e. a smaller region around the minimum before the first saddle point is reached. In order to achieve some measure for the statistical distribution of the results, both as far as the density of states is concerned and with respect to the distribution of local minima, we repeat each threshold run for a given energy lid and starting point.

† The moveclass is the set of neighbouring configurations in phase space that can be reached from a given point with one step of the random walk. A physically reasonable moveclass allows only those moves of the system that might occur during the regular time evolution of the system, e.g., small displacements of atoms in a solid.

3. The lumped model and relaxation behaviour

Obviously, the above procedure results in a tree-like barrier structure. Since the number of states available grows very fast with energy, the system will spend most of the random walk within the region of phase space with energies just below the lid†. Thus, for most realistic systems, a random walk of affordable length will not sample enough of the relatively rare low-lying states to allow us to draw satisfactory conclusions from $n(E; L, x_0)$ about the functional form of $g(E; L, x_0)$. A solution to this problem consists in using the overlap of the distributions $n(E; L, x_0)$ for different energy lids to determine $g(E; L, x_0)$ for high values of the threshold L. As long as no additional minima with their concomitant states are added while proceeding from lid L_i to lid $L_i + 1 > L_i$, no special problems arise by using the overlap procedure for boot-strapping, and the density of states can be estimated up to the missing normalization factor.

However, if such a new region becomes suddenly accessible at L_k, a larger number of low-lying states below L_{k-1} are now represented in the sampling $n(E; L_k, x_0)$ than had been available for $n(E; L_{k-1}, x_0)$. Thus, one has to correct for this effect when using the overlap procedure. Since each sample lacks an overall normalization, this step is usually not trivial‡. However, in those regions of phase space that are dominated by one large minimum, e.g. x_0, the correction due to the contributions of the side minima to $g(E; L, x_0)$ will be small. On the other hand, any process, e.g. a relaxation, starting from a side minimum x_1, will only show a dependence on x_1 for the time during which the system remains in the basin of x_1, which is a function of $n(E; L, x_1) \propto g(E; L, x_1)$ for $L < L_k$.

The combination of all the local densities of states, $g(E; L, x_i)$, represents a lumped picture of the phase space of the system. This description is intermediate between the overall density of states for the whole pocket $g(E; L_{max})$ and the exhaustive description of every microscopic detail of the energy landscape within the pocket that one can achieve in the discrete case. With this information, it is possible to construct a transition matrix $\mathbf{M}(T)$ in the lumped configuration space (cf. the appendix). This matrix allows the simulation of the evolution of the system for temperature T, especially its approach to equilibrium. In particular, this information connects the numerical results for a particular system to the growing body of work on hierarchical structures of phase spaces and diffusion processes on tree-like graphs [28–35].

In particular, one can calculate the time it takes for the local minima to equilibrate for different temperatures. In all instances, the system is prepared in a single node, i.e., all probability (=one unit) is initially placed into one of the local minima. The sequence of equilibrations among the minima can be represented by an 'equilibration/merger tree' [1].

While the major features of the equilibration trees, e.g. the sequence of mergers, are usually independent of the starting point and the temperature, the time-scales for equilibration will clearly depend on temperature. In particular, for temperatures much higher than the energy difference between the highest lid and the ground state, the observed equilibration time reflects the time the system needs to reach all parts of the available phase space. Defining an effective diameter of the system, d_{eff}, as the length of the longest path

† One should note that the highest thresholds used are often well above the highest barriers in the system. Thus, the bottlenecks, which for a low-energy lid can occur because of the relatively small number of paths available across a saddle point, are no longer critical, and each individual threshold run tends to be extensive enough to sample the whole of the available piece of the energy landscape. As a consequence, the distribution of local minima found upon quenching will no longer depend on the starting minimum of the threshold run(s).

‡ If it is possible to calculate the matrix of second derivatives for each minimum analytically, one can find the normal modes and use their densities of states to determine the missing normalization factor for each minimum.

within the set of shortest paths between any two states of the system, we can assume that for simple barrier structures the equilibration time τ_{eq} will be proportional to d_{eff}. Note that this allows us to determine a measure of the 'compactness' of the (lumped) phase space, $d_{eff} = f(V)$, by plotting τ_{eq} as a function of phase-space volume V.

4. Example: a single layer of neon (on a two-dimensional surface)

4.1. The configuration space and moveclass

In order to illustrate the method described in the previous section, we have chosen to investigate a single layer of solid neon. Using a standard Lennard-Jones potential for the pairwise interaction between the atoms ($\sigma = 2.75$ Å, $\varepsilon = 0.003\,68$ eV), the potential energy per particle is given by

$$e = \frac{E}{N} = \frac{1}{N}\sum_{\langle i,j\rangle} \varepsilon\left(\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right) + pv. \tag{1}$$

Here, we have included for the sake of completeness a term pv (v = volume per atom), since we allow volume changes during the determination of both the local minima and the threshold runs. All calculations were performed for $p = 0$; thus the total energy $E_{total} = E_{pot} + E_{kin}$ agreed with the enthalpy H. In order to deal with surface effects, we have chosen to apply periodic boundary conditions to the simulation cell. The number of atoms within the simulation cell was $N = 16$. Thus, the configuration space of the system is parametrized by the vectors describing the (repeated) simulation cell, and the coordinates of the atoms within this cell.

In order to determine the global minimum, we employed simulated annealing, followed by a quench and a gradient descent for the final adjustment of the coordinates. The moveclass used during the optimization allowed for random movements of the sixteen individual atoms and the variation of the size and shape of the simulation cell. The ground state was found to consist of a close packing of the neon atoms. Local minima of the energy corresponded to one or more vacancies within such a hexagonal packing. Of course, due to the periodic boundary conditions, the vacancies were present in each copy of the simulation cell. Thus, there existed a finite gap in the energy/atom between the ground state and a given vacancy structure.

4.2. The barrier structure

During the threshold runs, the moveclass employed was the same as during the global optimization runs. The lid energy was varied from $L = -3.07$ meV to $L_{max} = -2.27$ meV in steps of $\Delta E_{grid} = 10^{-4}$ eV. The highest lid value was chosen such as to remain slightly below the estimated melting point for two-dimensional neon ($\Delta(L_{max}) = L_{max} - E_{ground} = 0.83$ meV < 1.1 meV $\simeq T = 13$ K). Six different local minima were found during the runs and subsequently used as starting points (table 1).

For each minimum and each value of the lid energy, 30 threshold runs of length 500 000 steps were performed. From the results of the quenches at the end of these runs [36], we can deduce the barrier structure shown in figure 1.

As has been pointed out in section 2, the algorithm yields upper bounds on the height of the highest point along the lowest path between two minima, with respect to the energy of the starting local minimum. The best bounds on the barrier heights were found during runs starting in the high-lying minima. Since a simple system was used as an example, no

Table 1. A list of the main local minima in the two-dimensional neon crystal with 16 atoms per simulation cell.

Name	Description; d = centre–centre distance of vacancies ($R = r_{Ne} = 1.54$ Å)	Energy per atom (meV)
x_0	Ground state	−3.10
x_1	One vacancy per cell; $d \geqslant 4\sqrt{3}\ R$ (in next cell!)	−2.90
x_{2a}	Two vacancies per cell; $d = 2\ R$	−2.73
x_{2b}	Two vacancies per cell; $d = 2\sqrt{3}\ R$	−2.67
x_{2c}	Two vacancies per cell, $d = 4\ R$	−2.67
x_{2d}	Two vacancies per cell; $d = 4\sqrt{2}\ R$	−2.67

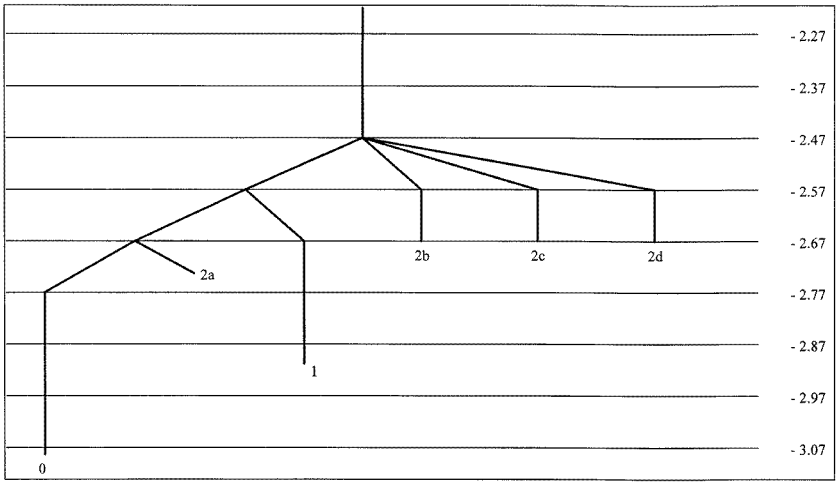

Figure 1. A tree diagram of the energy barrier structure of the two-dimensional neon crystal. Barrier heights are given in meV per atom. For the description of the minima, see table 1.

complicated barrier structure was found. Once the system could leave the local basin, paths were available to reach the main part of the pocket, which was connected to the ground state.

4.3. Local densities of states

In addition, the energy landscape was sampled during the threshold runs. The interval between samples was chosen to be 1000 steps, in order to exclude correlations as far as possible. The results for $n(E; L, x)$ are summarized in figure 2(a) for the most important case, $x = x_0$. Since the number of states grows very fast with energy, the registered energies tend to be concentrated below the lid. In order to derive a density of states from these results, we note that for successive lid values L_k, L_{k+1} there exists a considerable overlap between $n(E; L_k, x_0)$ and $n(E; L_{k+1}, x_0)$. Using this overlap, we can determine $g(E; L, x_0)$ even for high values of L, up to a normalization factor. Figure 2(b) shows $g(E; L, x_i)$ for the highest lid value $L = -2.27$ meV for all starting points x_i. The large jump in the number of accessible states, which occurs when joining the main region of the pocket from a high-lying basin, is clearly visible. The actual size of the jump cannot be

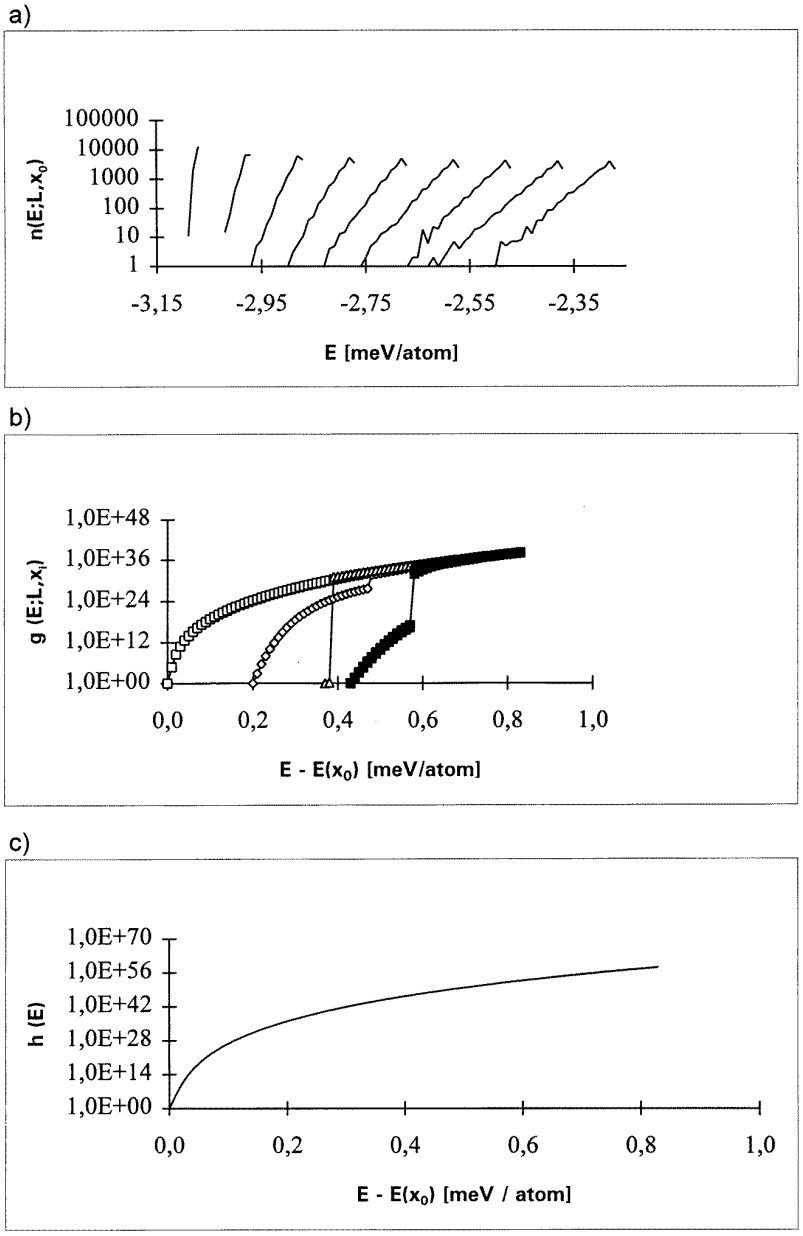

Figure 2. (a) Sampled densities of states $n(E; L, x_0)$ for a sequence of lid values. The starting point was the minimum x_0. (b) The local densities of states $g(E; L, x_i)$. White squares correspond to x_0, white diamonds to x_1, white triangles to x_{2a}, black squares to x_{2b}, black triangles to x_{2c}, and black diamonds to x_{2d}. (c) The density of states $h(E)$ including kinetic degrees of freedom.

determined directly, since the normalization factors of the two basins need not be identical (cf. section 3).

So far, we have only determined the density of states belonging to the potential energy of the system, $g \equiv g_{pot}$. In order to calculate statistical mechanical properties, we need to

include the contribution of the kinetic energy. Within the classical picture we have been using, this is straightforward, since we can split E_{total} into a sum of potential and kinetic energy. The potential energy depends only on the coordinates, and its density of states has been calculated above. The contribution of the kinetic energy depends only on the momenta of the particles, like in the ideal gas, and the 'kinetic' density of states g_{kin} is therefore easily evaluated [37]. The convolution of these contributions determines the required density of states $h(E_{total}; L, x_0)$ as a function of the total energy, up to a normalization factor:

$$h(E) = \sum_{E'=0}^{E} g_{pot}(E')g_{kin}(E - E'). \tag{2}$$

In equation (2), the ground state of the system was used to define the zero of the energy scale. The result is shown in figure 2(c) for the starting minimum x_0.

From this, we can now compute [37], for $p = 0$, the average enthalpy (figure 3(a)) for several values of the lid L ($D(L) = L - E_{ground} = 0.23$ meV, 0.43 meV, 0.83 meV) and

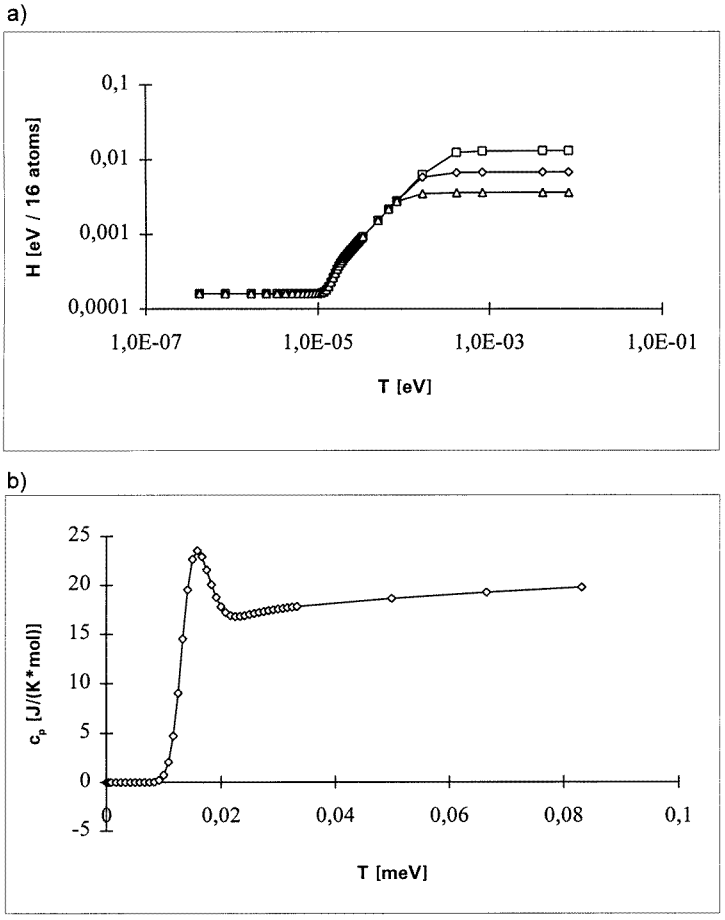

Figure 3. (a) The average enthalpy H as a function of temperature T. White squares correspond to a lid with energy $L - E(x_0) = 0.83$ meV above the ground state, white diamonds to $L - E(x_0) = 0.43$ meV, and white triangles to $L - E(x_0) = 0.23$ meV. (b) Specific heat C_p as a function of temperature T.

the specific heat (figure 3(b)) as a function of temperature ($D(L) = 0.83$ meV). The width of the energy slices used for the construction of the $h(E)$ was $\Delta E = 10^{-5}$ eV.

We note that only those quantities that do not require knowledge of the overall normalization of the density of states can be calculated. Fortunately, this is sufficient for many expectation values of interest. Furthermore, the results (figure 3(a)) depend on the height of the lid above the ground state, $D(L) = L - E_{ground}$, and on the width of the energy slices ΔE used for the binning of the sampled $n(E; L, x)$: if $T \geqslant D(L)$, then $\langle H \rangle \approx D(L) =$ constant. Similarly, if $T \leqslant \Delta E$, $\langle H \rangle \approx 0 =$ constant. From classical arguments one would expect that the specific heat of two-dimensional neon should be two thirds of the value for the three-dimensional case. These classical arguments appear to be quite reasonable, since the average experimental [38] value $C_p^{3d}(\text{exp}) = 25$ J K^{-1} mol^{-1} is in good agreement with the derived value $C_p^{3d}(\text{class}) = 24.94$ J K^{-1} mol^{-1}. Thus, the expected value in the two-dimensional case should be $C_p^{2d}(\text{class}) = 16.67$ J K^{-1} mol^{-1}, which agrees very nicely with our calculated value $C_p^{2d}(\text{calc}) \approx 17$ J K^{-1} mol^{-1}. Finally, we point out that the results presented in figures 3(a) and 3(b) are based on a classical approximation that will no longer hold for very low temperatures. However, the basic approach presented here should be applicable even to quantum systems as long as tunnelling through the barriers may be neglected.

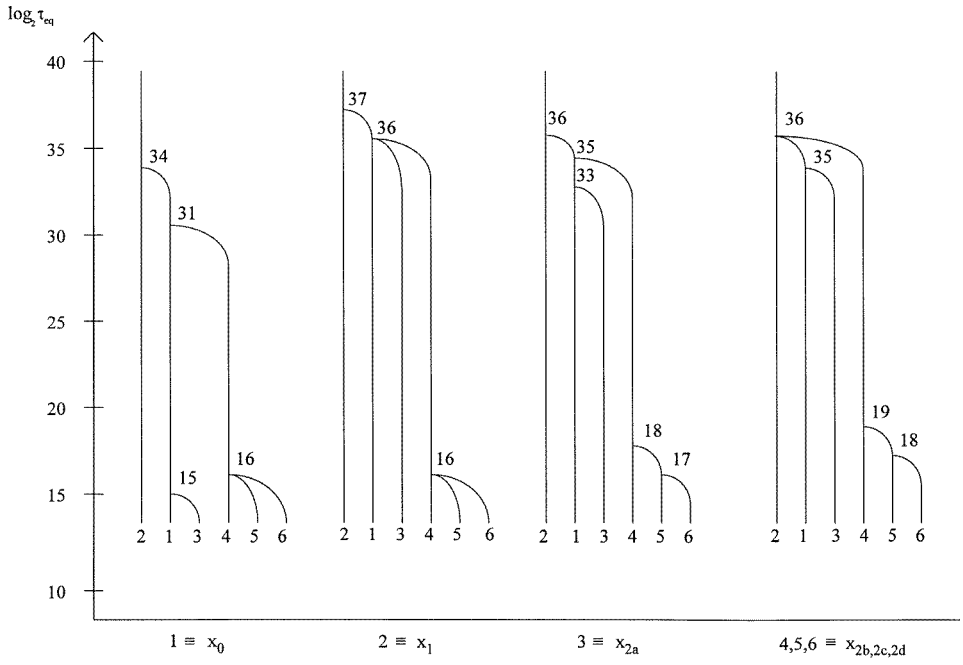

Figure 4. Equilibration trees for different starting points. The equilibration time τ_{eq} is given by the number of multiplications by $\mathbf{M}(T)$ necessary to reach equilibrium with accuracy $a = 10^{-5}$. The tree '1' is found for starting point x_0, tree '2' for starting point x_1, tree '3' for starting point x_{2a}, and tree '4' = tree '5' = tree '6', for starting points x_{2b}, x_{2c}, x_{2d}, respectively. Within each tree, the branches are labelled accordingly. The numbers where they merge indicate the equilibration time.

4.4. Relaxation behaviour

Following the procedure outlined in the appendix and section 3, we have derived a lumped description of the system from the local densities of states, where we have used slices of width $\Delta E = 3 \times 10^{-5}$ eV to form the nodes of the tree. This value was chosen, since the distribution of energy differences per move was sharply peaked at 0 eV, and approximately 90% of the steps showed an energy difference below 3×10^{-5} eV. Thus, this choice allowed us to restrict the connections among nodes to those belonging to successive energy slices. In addition, it was estimated from this step size distribution that for about a third of the time, a move (at $T = \infty$) should have led to a point within the same node, i.e. $f_{ii} = 1/3$ for all nodes i except the top node.

After constructing the transition matrix $\mathbf{M}(T)$ in the lumped configuration space, we have calculated the time it takes for the local minima to equilibrate for different temperatures. In all instances, the system was prepared in a single node, i.e., all probability (=one unit) was initially placed into one of the local minima. The set of the resulting equilibration trees for $T = 5 \times 10^{-6}$ eV is depicted in figure 4.

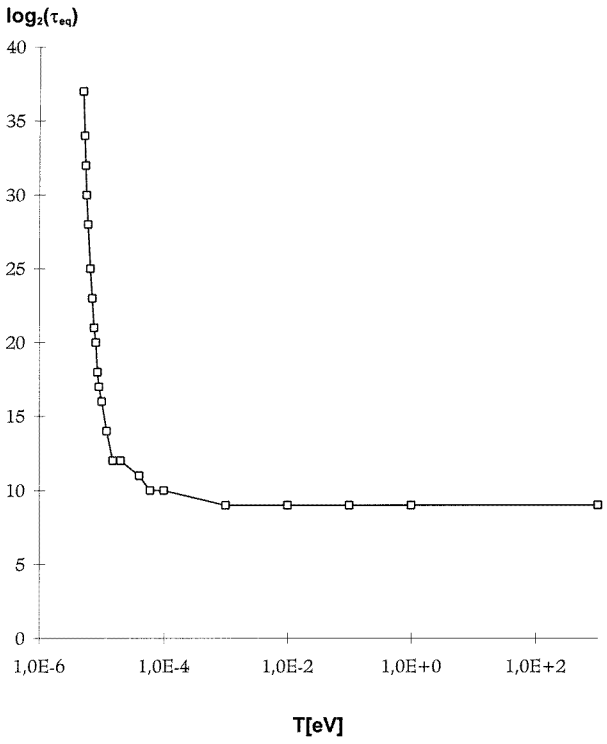

Figure 5. The time necessary for full equilibration of the system τ_{eq} to be achieved as a function of temperature T.

Note that the major features of the equilibration trees are independent of the starting point: minimum x_1 always takes the longest to merge with the rest of the pocket, indicating a relatively high stability of this configuration, while the minima x_{2b}, x_{2c} and x_{2d} equilibrate rather quickly among themselves. In order to understand the detailed structure of such equilibration trees, it is necessary to analyse the individual probability flows among the

minima, taking into account both the energy barriers and the size of the basins, i.e. the phase-space volume 'belonging' to each minimum.

In figure 5, we show the time that the system needs to achieve full equilibration as a function of temperature. We note that there is a plateau for $T > 10^{-3}$ eV. This is no surprise, since for the given maximum lid of 0.83 meV above the ground state, this range of T corresponds to infinity, for all practical purposes. Plotting τ_{eq} as a function of phase-space volume V for $T = 1$ eV, we find $t_{eq} \propto V^{\gamma}$, with $\gamma \approx 0.25$.

5. Summary and outlook

We have presented a method for the study of complex multi-minima systems which allows the determination of barrier heights among the minima and the construction of a simplified 'lumped' model of the configuration space. On the basis of such a model, equilibrium and relaxation behaviour of the system may be investigated. As an example, we have studied neon in two dimensions containing vacancies. By now, this method has already been applied to the determination of the heights of barriers surrounding metastable structures in the Na–Cl system [39].

Clearly, many possibilities for improving and refining the method presented above come to mind. One intriguing option might be to use techniques used by researchers studying the energy hypersurface of continuous systems, e.g. clusters, with the help of molecular dynamics methods [9, 12, 40]. After finding local minima through quenches during molecular dynamics runs, some saddle points between these minima are determined with the use of algorithms like the steepest-ascent method, the method of slowest slides, and 'eigenvector following' in order to identify the saddle points [41, 42]. As long as an analytical form of the energy function and its derivatives are known, these methods can be applied†, and they could be used for the determination of saddle points, in order to associate special configurations with the barriers found by using the threshold algorithm. This would lead to a further, intuitive understanding of the time evolution of a given material.

Acknowledgments

One of us, CS, would like to thank Professors R S Berry, J Nulton, P Salamon and P Sibani for many interesting discussions relating to the topic of the paper. We also gratefully acknowledge funding by the DFG through the Leibniz Programme, and SFB 408.

Appendix. Modelling the relaxation behaviour using local densities of states

The diffusion of probability in configuration/phase space is a possible way of viewing the time evolution of a system in contact with a heat bath at temperature T. If the state space of the system is discrete, and the moveclass (=connections between nodes of the configuration space) is known, this diffusion can be modelled as a discrete Markov process with a certain transition matrix. Repeated application of this matrix to an initial state of the system leads to a probability distribution that corresponds to the expected outcomes of the time evolution

† A note of caution should be sounded, however: in many situations, these analytical energy functions and/or their derivatives are not accessible. In addition, just as gradient methods cannot guarantee the determination of the deepest minima during a global optimization, analytical methods that depend on local information may not discover all of the important saddle points.

of the system. For a continuous state space, this matrix multiplication may be replaced by an integration of the initial distribution with the appropriate transition kernel.

However, in most examples of interest, one does not know this kernel exactly. In our particular example, only an estimate of the local densities of states $g(E; L, x_i)$ together with a distribution of the energy differences δE encountered during the random walks in configuration space below the lid L are available. From this information, one can construct a 'lumped' model of the configuration space, and subsequently a transition matrix for the diffusion on this graph. The suggested procedure outlined below is in some aspects similar to one proposed by Nulton [43] previously, and to the models used for the analysis of simulated annealing [44].

(1) Choose the width of energy slice ΔE for the creation of a lumped model of the configuration space. All the states in the interval $[E, E + \Delta E]$ that belong to the basin of the same minimum are joined into a lumped node. The number of microscopic states within node i, $g_i (i = 1, \ldots, N)$, is then proportional to the local density of states. Once such a node is established, all microscopic states within are supposed to show the same average properties as far as the rest of the system is concerned, especially with respect to connectivity to other nodes. This property allows us to ignore the missing overall normalization factor of the densities of states. The only prerequisite for this procedure is the knowledge of the ratio of the number of states between any two nodes.

(2) Depending on the δE-distribution, non-zero connections c_{ij} are established between all the nodes i and j. Preferably, one would choose ΔE such that only connections between successive energy slices, and from the node back to itself, c_{ii}, are necessary.

(3) Construct the transition matrix $\mathbf{M}(T = \infty)$ with elements f_{ij}, where f_{ij} can only be non-zero if c_{ij} was already non-zero. In general, this transition matrix has to fulfil two sets of conditions: (a) probability must be preserved:

$$\sum_i f_{ij} = 1 \qquad \text{for } j = 1, \ldots, N \tag{A1}$$

and (b) microscopic reversibility (detailed balance) holds, i.e., each microscopic move can be reversed (with equal probability at $T = \infty$). Within the context of this model, this translates into the conditions

$$f_{ij} g_j = f_{ji} g_i \qquad \text{for } i, j = 1, \ldots, N. \tag{A2}$$

This condition is fulfilled trivially by all f_{ij} that belong to unconnected nodes ($f_{ij} = f_{ji} = 0$), and for the case $i = j$. Note that from conditions (a) and (b) it follows that the vector $\boldsymbol{g} = (g_i)$ is an eigenvector of $\mathbf{M}$ with eigenvalue 1:

$$\sum_j f_{ij} g_j = \sum_j f_{ji} g_i = g_i. \tag{A3}$$

This agrees with the physical interpretation of g_i, since the equilibrium distribution at $T = \infty$, which the system should reach for $t \to \infty$, is given by

$$p_i^{eq}(T = \infty) = \pi_i = g_i \Big/ \sum_j g_j.$$

Counting equations and unknowns for a graph with N nodes and F edges ($F \geqslant N - 1$), we find: for each node, one unknown f_{ii}; and for each edge (ij), two unknowns f_{ij} and f_{ji}, i.e. $N + 2F$ variables. For each node we find one column normalization, and for each connected pair of nodes, one equation, i.e. $N + F$ equations. Thus, we may add F equations, i.e. F variables may be chosen freely—within certain limits, of course. Here, one should try to include any additional information available about the system from, e.g., the step

size distribution δE. However, one should keep in mind that for the highest node(s) of the system, f_{ii} has to incorporate all the attempted moves that would lead to states above the lid.

For a tree graph, a common occurrence in lumped models, we have $F = N - 1$. Using an analysis of the step size distribution, it is often reasonable to choose the $N-1$ additional conditions by prescribing the values of the f_{ii} for $i = 1, \ldots, N-1$—the only exception being the top node. As long as the inequality $\sum_i g_i < g_j$, where the sum extends over all nodes i lying below the node j, holds for all nodes j, the remaining $2N-1$ equations can be solved for f_{ij} with physically reasonable results. This condition holds, e.g., for the example considered in section 4.

(4) The matrix (f_{ij}) is the desired transition matrix for $T = \infty$. For finite T, the standard 'Boltzmannization' procedure [45] can be applied: define $x_i = \exp(-E_i/T)$; multiply each entry f_{ij} by $\min(1, x_i/x_j)$; sum each column; add the difference between 1 and the column sum into the diagonal element. The result is the required transition matrix $\mathbf{M}(T)$. Note that the vector $\pi = (g_i x_i)$ is an eigenvector of $\mathbf{M}$ with eigenvalue 1:

$$\begin{aligned}(\mathbf{M}(T))_{ij} g_j x_j &= \sum_{j\neq i} f_{ij} \min(1, x_i/x_j) g_j x_j + \left(1 - \sum_{k\neq i} f_{ki} \min(1, x_k/x_i)\right) g_i x_i \\ &= g_i x_i + \sum_{j\neq i} f_{ij} \min(1, x_i/x_j) g_j x_j - \sum_{k\neq i} f_{ki} \min(1, x_k/x_i) g_i x_i \\ &= g_i x_i + \sum_{j\neq i, E_i > E_j} f_{ij} g_j x_i + \sum_{j\neq i, E_i < E_j} f_{ij} g_j x_j \\ &\quad - \sum_{k\neq i, E_i > E_k} f_{ik} g_k x_i - \sum_{k\neq i, E_i < E_k} f_{ik} g_k x_k \\ &= g_i x_i. \end{aligned} \tag{A4}$$

Thus, the system will converge to the Boltzmann distribution given by

$$\pi_i = g_i x_i \Big/ \sum_j g_j x_j.$$

Having determined $\mathbf{M}(T)$, one can now study the equilibration processes. It is useful to define $z_i(n) = p_i(n)/p_i$. Note that $z_i(n) \to 1$ for $n \to \infty$. If for a given accuracy a there exists a number n_0 such that $|z_i(n) - z_j(n)|/z_j(n) < a$ for all $n > n_0$, then the two nodes i and j are said to be in equilibrium at time step n_0 with accuracy a. The smallest such number n_0 denotes the 'merger' or equilibration time of the two nodes, τ_{eq}. Plotting τ_{eq} against participating nodes yields the so-called equilibration tree [1].

References

[1] Sibani P, Schön J C, Salamon P and Andersson J O 1993 *Europhys. Lett.* **22** 479–85
[2] Palmer R G 1982 *Adv. Phys.* **31** 669
[3] Davis L 1987 *Genetic Algorithms and Simulated Annealing* (London: Pitman)
[4] Elliott S R 1990 *Physics of Amorphous Materials* (Harlow: Longman Scientific & Technical) 186ff
[5] Takayama H 1988 *Cooperative Dynamics in Complex Physical Systems* (New York: Springer)
[6] Tenelsen K and Schreiber M 1994 *Phys. Rev.* B **49** 12 662
[7] Möbius A, Richter M and Drittler B 1992 *Phys. Rev.* B **45** 11 568
[8] Berry R S 1993 *Chem. Rev.* **93** 2379–94
[9] Berry R S 1994 *J. Phys. Chem.* **98** 6910–18
[10] Kunz R E and Berry R S 1994 *Phys. Rev.* E **49** 1895–908
[11] Rose J P and Berry R S 1993 *J. Chem. Phys.* **98** 3246–60
[12] Wales D J and Berry R S 1990 *J. Chem. Phys.* **92** 4473

[13] Wales D J 1990 *Chem. Phys. Lett.* **166** 419–24
[14] Sibani P and Schriver P 1994 *Phys. Rev.* B **49** 6667–71
[15] Klotz T and Kobe S 1994 *Acta Phys. Slovaca* **44** 347–56
[16] Sibani P and Andersson J O 1994 *Physica* A **206** 1–12
[17] Hertz J A, Krogh A and Palmer R G 1988 *Introduction to the Theory of Neural Computation* (New York: Addison-Wesley)
[18] Jepsen D W and Gelatt C D Jr 1983 *ICCD 1983* (New York: IEEE) pp 495–8
[19] Kirkpatrick S and Toulouse G 1985 *J. Physique* **46** 1277–92
[20] Liao W 1987 *J. Phys. A: Math. Gen.* **20** L695–9
[21] Palmer R 1988 *The Economy as an Evolving System* (New York: Addison-Wesley) pp 1–17
[22] Mosegaard K 1987 *PhD Thesis* University of Copenhagen
[23] Mosegaard K and Tarantola A 1995 submitted
[24] Shakhnovich E, Farztdinov G, Gutin A M and Karplus M 1991 *Phys. Rev. Lett.* **67** 1665
[25] Anderson P W, Halperin B I and Varma C M 1972 *Phil. Mag.* **25** 1–9
[26] Buchenau U, Galperin M, Gurevich V L, Parshin D A, Ramos M A and Schober H R 1992 *Phys. Rev.* B **46** 2798–808
[27] Pierre D A 1986 *Optimization Theory with Applications* (New York: Dover)
[28] Sibani P and Hoffmann K H 1991 *Europhys. Lett.* **16** 423
[29] Sibani P and Hoffmann K H 1989 *Phys. Rev. Lett.* **63** 2853
[30] Schultze C, Hoffmann K H and Sibani P 1991 *Europhys. Lett.* **15** 361
[31] Ogielski A T and Stein D L 1985 *Phys. Rev. Lett.* **55** 1634
[32] Hoffmann K H and Sibani P 1990 *Z. Phys.* B **80** 429
[33] Grossmann S, Wegner F and Hoffmann K H 1985 *J. Physique Lett.* **46** 575
[34] Simon H A 1962 *Proc. Am. Phil. Soc.* **106** 467
[35] Hubermann B and Kerszberg M 1985 *J. Phys. A: Math. Gen.* **18** L331
[36] Putz H 1994 *Dipl. Thesis* Universität Bonn
[37] Pathria R K 1978 *Statistical Mechanics* (New York: Pergamon)
[38] *Landolt–Börnstein New Series* 1961 Group 3, vol 4 *Kalorische Zustandsgrößen* (Berlin: Springer)
[39] Schön J C and Jansen M 1995 *Comput. Mater. Sci.* **4** 43–58
[40] Kunz R E and Berry R S 1993 *Phys. Rev. Lett.* **71** 3987–90
[41] Simons J, Jørgensen P, Taylor H and Ozment J 1983 *J. Phys. Chem.* **87** 2745–53
[42] Stillinger F H and Wever T A 1984 *J. Chem. Phys.* **80** 4434–7
[43] Nulton J 1990 personal communication
[44] van Laarhoven P J M and Aarts E H L 1987 *Simulated Annealing* (Dordrecht: Reidel)
[45] Andresen B, Hoffmann K H, Mosegaard K, Nulton J, Pedersen J M and Salamon P 1988 *J. Physique* **49** 1485