Atomistic Modelling of Diffusion

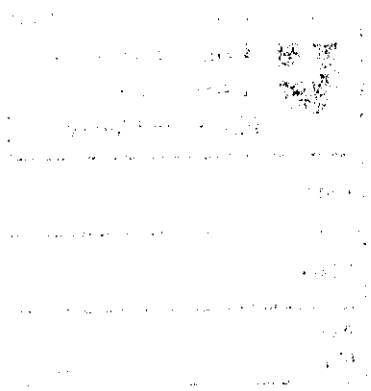
Stewart M. J. Gordon

A Doctoral Thesis

Submitted in partial fulfilment of the requirements for the award of

Doctor of Philosophy of Loughborough University

February 2006



Abstract

This thesis describes the dimer method, which is an algorithm that can be used to find state transitions in an atomistic system, and the application of this method to two different atomistic diffusion problems.

The dimer method is an algorithm that locates the saddle points of a potential field of arbitrary dimensionality. These saddle points correspond to the points of transition between metastable states of an atomistic system. A number of improvements to the algorithm of the dimer method have been described and implemented in this work.

The first atomistic problem to be described is the diffusion of Au adatoms on a face-centred cubic Au(100) surface. By applying the dimer method to this system, a number of state transitions involving varying numbers of atoms are discovered, from the initial configuration of a single adatom on the surface and from configurations of two adatoms close together. The energy barriers are given for the transitions that appeared in the searches. From all of these initial setups, the preferred diffusion mechanism is a simple hop of one adatom to a diagonally adjacent hollow site on the surface. The results are compared and contrasted with previous studies of Al/Al(100) adatom diffusion, which is a similar system, but in which concerted exchanges of adatoms with surface atoms have the lowest energy barriers and are therefore the most probable mechanisms.

The second problem studied in this thesis is the diffusion of point defects in bulk α-Fe with and without P impurities. This problem is important to the study of the embrittlement of steel nuclear pressure vessels. Both vacancy-driven and interstitial-driven diffusion are considered. The dimer method is used to find the transition mechanisms and energy barriers in these defect systems. Transitions are discovered in which a va-

cancy or an Fe–Fe dumbbell moves between positions at different displacements from a substitutional P atom. Kinetic Monte Carlo simulations are then performed at five different temperatures on three α-Fe defect systems: an isolated vacancy, a P–vacancy complex and a P interstitial defect. The results show that the isolated vacancy and P interstitial systems obey the Arrhenius relation, but the P–vacancy complex does not obey this relation. The P–vacancy complex is the least mobile of these defect systems, and the P interstitial defect is the most mobile, and hence an important process by which P impurity atoms can migrate to the grain boundaries of irradiated steel nuclear pressure vessels.

Acknowledgements

I would like to thank my supervisors, Professor Roger Smith and Dr Steven D. Kenny, for their guidance and support for the duration of my programme of research. I would also like to express my gratitude to the former research associate Dr Carlos F. Sanz-Navarro for developing the LBOMD software on which my implementation of the dimer method is based, and to the former postgraduate student Hurryramsingh Hurchand, who has contributed a significant amount of work in the atomistic modelling of P impurities in α-Fe. I also thank all people currently and formerly in my department who have contributed useful information and discussions that have both helped my research to progress and helped me in writing this thesis.

My thanks also go to some of the leading researchers in atomistic modelling of materials, including Graeme J. Ackland for developing interatomic potentials that have been used in this work, Graeme Henkelman for developing the dimer method and for discussions on its implementation, and Arthur F. Voter for his contributions to and discussions on the field of accelerated dynamics techniques.

Thanks to all my family for being generally supportive while I have been working on my PhD, and to the many friends I have made during my years at Loughborough University who have helped me to enjoy my time here.

Finally, I am grateful to Loughborough University for granting me a research studentship and thereby making it possible for me to undertake my research project.

Contents

List of Figures ix

List of Tables xv

1 Introduction 1

1.1 Thesis structure 4

2 Molecular modelling 7

2.1 Interatomic potentials 7

2.1.1 Embedded atom method 9

2.1.2 Ackland potential 9

2.1.3 Morse potential 13

2.1.4 Stillinger-Weber potential 14

2.1.5 Tersoff potential 15

2.1.6 Coulomb potential 16

2.1.7 Ziegler-Biersack-Littmark (ZBL) potential 17

2.1.8 Connecting spline function 19

2.2 Molecular dynamics simulation 22

2.2.1 Cut-off distances and neighbour lists 23

2.2.2 Boundary conditions 25

2.2.3 Parallel processing 26

2.3 Accelerated dynamics methods 29

2.3.1 Parallel replica dynamics 30
2.3.2 Hyperdynamics 32
2.3.3 Temperature-accelerated dynamics 33
2.3.4 Kinetic Monte Carlo 34

3 The dimer method **37**
3.1 Barrier calculation methods 37
3.1.1 Molecular statics 37
3.1.2 The nudged elastic band method 38
3.1.3 Mode-following algorithms and the dimer method 40
3.2 The standard algorithm 41
3.2.1 Initialisation 42
3.2.2 Rotation 42
3.2.3 Translation 45
3.3 Adaptations 47
3.3.1 Generating the initial configuration 48
3.3.2 Fixed and free atoms 48
3.3.3 Generating the initial dimer vector and displacement from the minimum 49
3.3.4 Relaxation of fixed atoms 49
3.3.5 Alternative formulae for the angle of rotation 51
3.3.6 Order of steps and stopping criteria 52
3.3.7 Parallelisation 54
3.4 Benchmarks 55
3.4.1 Basic egg-box potential 55
3.4.2 Distorted egg-box potential 57
3.4.3 Three-dimensional potential 59
3.4.4 Effect of relaxation of fixed atoms 60
3.5 Scaling with system size 61

3.6 Conclusions 62

4 Transitions of gold adatoms on a gold surface 63

4.1 Single adatom 65

4.1.1 Transition rates 69

4.2 Two adatoms 72

4.2.1 Adatoms at orthogonally adjacent sites 73

4.2.2 Adatoms at diagonally adjacent sites 74

4.2.3 Adatoms separated orthogonally by a vacant hollow site 76

4.2.4 Adatoms separated diagonally by a vacant hollow site 76

4.3 Conclusions 80

5 Transitions in Fe and Fe–P systems 83

5.1 Defect structures 84

5.1.1 Neighbour distances 86

5.2 Results from Ackland Fe and Morse potentials 87

5.2.1 Fe self-interstitial defect 87

5.2.2 P interstitial defect 88

5.3 Results from Ackland potential for Fe containing P 92

5.3.1 Isolated vacancy in Fe 92

5.3.2 The P–vacancy complex 93

5.3.3 Isolated Fe self-interstitial 97

5.3.4 P interstitial defect 98

5.4 Conclusions 101

6 Kinetic Monte Carlo 103

6.1 Methodology 103

6.1.1 Dealing with symmetry 104

6.1.2 Dissociation 110

6.2 Calculating diffusion constants 111

6.3 Results for Fe and Fe–P systems 112
6.3.1 Isolated vacancy 112
6.3.2 P–vacancy complex 113
6.3.3 P interstitial defect 117
6.4 Conclusions 119

7 Conclusions and future work **121**
7.1 Future work 124

List of Figures

2.1 The ZBL potential appplied to Fe–Fe interaction (solid line) and the potential energy associated with simple Coulombic repulsion of Fe nuclei (dashed line). 18

2.2 Graph of $V(r)$ for Au–Au interaction. The spline function connects the ZBL potential ($r \leq 1.7$) and the pairwise part of Ackland's potential function ($r \geq 2.85$). 21

2.3 Graph of $-V'(r)$ for Au–Au interaction. The spline function connects the ZBL potential ($r \leq 1.7$) and the Ackland pairwise function ($r \geq 2.85$). . . 21

2.4 Schematic illustration of the spatial decomposition method of building the neighbour list. The box of atoms is partitioned into cells, and the algorithm looks through the atom lists of the shaded cells for neighbours of the atom with a bold border. The result is that the shaded atoms are picked out. In this instance two atoms are in the skin – they do not interact with the central atom at the moment, but are included in case they move to within the cut-off radius in the time between neighbour list updates. 24

2.5 Schematic illustration of periodic boundary conditions applied in a single dimension. The rectangle represents the box within which the simulation takes place, and the shaded circles the atoms within it, which are the only atoms included in the simulation data. This arrangement of atoms is considered to repeat itself either side of the box (unshaded atoms). Consequently, the distance r_{ij} is taken to be the distance between the atom i and the atom j' in the adjacent periodic cell corresponding to j. . 26

3.1 Deriving the rotational force on the dimer. 43

3.2 Unit basis vectors and the rotational force in initial and displaced positions. 44

3.3 Effective force along dimer axis for $C \geq 0$. 46

3.4 Effective force as reflected resultant force for $C < 0$. 46

3.5 Schematic illustration of the arrangement of fixed atoms (white and blue), free atoms (green) and atoms included in the initial dimer vector (yellow background). Also shown is the atoms around the outside (blue) that are fixed even during the periodic conjugate gradient relaxation. 50

3.6 The angle of rotation of the dimer as calculated by the forward difference and forward average approximations of equations (3.12) and (3.13). . . . 51

3.7 Adjustment of the angle of rotation by $\delta\theta/2$ to produce a closer convergence. 52

3.8 Over-rotation occurring with Newton's method. 53

3.9 Constructing a sinusoidal fit on the rotational force curve. 53

3.10 Contour plot of the basic egg-box potential. Light areas are positive and dark areas are negative. A sample path of the centre of the dimer in this potential is shown. 56

3.11 Contour plot of the distorted egg-box potential. Light areas are positive and dark areas are negative. A sample path of the centre of the dimer in this potential is shown. 58

4.1 The three typical sites at which adatoms occur, shown on an fcc lattice. In all diagrams in this chapter, the atoms are shaded by depth, with the lightest shade used for the adatom layer. 64

4.2 Transitions found with 18 free atoms, of which five atoms are included in the initial dimer vector. In each diagram, the atoms that move between sites have been labelled so that the initial and final positions of each atom can readily be seen. 66

4.3 Transitions found with 18 free atoms, all of which are included in the initial dimer vector. 67

4.4 Additional transitions found with numbers of free atoms between 22 and 56. 69

4.5 The rate constants for five energy barriers, plotted on a logarithmic scale against temperature. The ratio between rate constants becomes smaller as the temperature increases, so that higher-energy transitions become more probable. 71

4.6 The four two-adatom initial states from which dimer searches were conducted, in ascending order of potential energy. 72

4.7 Transitions found in the [001] addimer system. 73

4.8 Transitions found in the [011] addimer system. 75

4.9 Transitions found in the system of two adatoms separated by a single hollow site. 77

4.10 Transitions found from the initial state of two adatoms separated diagonally by a single hollow site. 79

5.1 Some defect structures commonly found in bcc materials. The green circles denote lattice sites. The structures shown are a perfect lattice cell (a), two vacancy defect structures (b, c) and four interstitial defect structures (d, e, f, g). The tetrahedron (f) and octahedron (g) defining the respective interstitial positions have been outlined in blue. 85

5.2 Nearest neighbour numbers corresponding to small displacements from a given lattice point, represented by a blue sphere. 86

5.3 A sample sequence of transitions of an Fe self-interstitial defect. 87

5.4 The transitions found in the P interstitial system using the Ackland/Morse model, with energy barriers. 89

5.5 Sample sequence of transitions between ⟨110⟩ dumbbell and tetrahedral configurations. 91

5.6 The rotation of a ⟨110⟩ dumbbell via a ⟨643⟩ dumbbell configuration. . . 91

5.7 The mechanism of an isolated vacancy (green, unfilled circle) exchanging with an adjacent Fe atom (red sphere), via a split vacancy state (interstitial Fe atom shown with a black outline). 93

5.8 The mechanism of a P atom (blue sphere) exchanging with a vacancy (unfilled circle) when they are in first neighbour positions. 94

5.9 The mechanism of the first neighbour P–vacancy reorientation via the second neighbour state, with energy barriers. Fe atoms (red spheres) move to exchange with the vacancy (unfilled circle), while the P atom (blue sphere) does not move. 94

5.10 The energy profiles of the P–vacancy exchange mechanism and three Fe–vacancy exchange mechanisms between different neighbour sites of the P atom. Only the metastable states (•) and the saddle points (+) are plotted. The letters on the profile of the P–vacancy exchange refer to the stages of the exchange illustrated in Figure 5.8. 97

5.11 The mechanism of migration of an isolated Fe–Fe ⟨110⟩ dumbbell by simultaneous rotation and translation. 98

5.12 The mechanism of migration of an isolated Fe–Fe ⟨110⟩ dumbbell via an octahedral configuration. 98

5.13 The mechanism of a P atom moving by changing between the ⟨110⟩ dumbbell and tetrahedral states. 99

5.14 A selection of transitions between ⟨110⟩ dumbbells in the P interstitial defect system, with energy barriers. 100

6.1 Neighbour numbers of bcc lattice sites adjacent to the first (a) and second (b) neighbour positions. 105

6.2 Some examples of vectors used to denote states in the interstitial defect system. Vectors have been given in units of the lattice parameter. In each case, the first vector is given in blue, and the second in purple. 108

6.3 Arrhenius plot of the diffusion constants of an isolated vacancy in Fe (◆), and of a P atom (■) and a vacancy (▲) in a P–vacancy complex with a periodic cell size of $15 \times 15 \times 15$ bcc unit cells. The fine dotted lines show the extrapolation from the diffusion constants in the P–vacancy system calculated at 350 K and 400 K if the Arrhenius relation is assumed. . . . 114

6.4 Arrhenius plot of the diffusion constants of a P interstitial defect in Fe. . 118

List of Tables

2.1 The Ackland potential parameters used for Au–Au pairwise interaction. . 11

2.2 The Ackland potential parameters used for Au–Au many-body interaction. 11

2.3 The Ackland potential parameters used for Fe–Fe pairwise interaction. . 12

2.4 The Ackland potential parameters used for Fe–Fe many-body interaction. 13

2.5 The Tersoff potential parameters for C and Si interaction. 16

3.1 Mean force evaluation counts on the simple egg-box potential for different values of Δt. 57

3.2 Mean force evaluation counts on the distorted egg-box potential. 59

3.3 Mean force evaluation counts on the three-dimensional distorted egg-box potential using steepest descent rotation. 60

3.4 Mean force evaluation counts on the three-dimensional distorted egg-box potential using conjugate gradient rotation. 60

3.5 Comparison of the energy barriers calculated for the third–fourth neighbour transition of Figure 5.14 (e) with different numbers of free atoms, with and without periodic relaxation of the fixed atoms. 61

4.1 The five lowest-energy transitions in Al, with corresponding barriers for Au. 68

4.2 The five lowest-energy transitions in Au. In the Al work [10], a barrier is given only for the first of these transitions, 0.37 eV. 68

4.3 The rate constants for a selection of energy barriers at 300 K, 500 K and 700 K. Four of the lowest energy barriers in the Au/Au(100) system, along with two higher ones, have been chosen to represent the range of barriers in this system. 70

4.4 Comparison of energy barriers between transitions in the [001] addimer system and their equivalent one-adatom transitions given by removing a non-moving adatom. 73

4.5 The rate constants for the three lowest-energy transitions in the [001] addimer system at 300 K, 500 K and 700 K. 74

4.6 The rate constants for the three lowest-energy transitions in the [011] addimer system at 300 K, 500 K and 700 K. 76

4.7 Comparison of energy barriers between transitions from the configuration of two adatoms separated by a vacant hollow site and their equivalent one-adatom transitions. Where a range of values is given, the different barriers are for different orientations of the transition relative to the displacement vector between the initial locations of the two adatoms. 78

5.1 The rate constants for the transitions of an isolated Fe self-interstitial defect found using the Ackland pure Fe potential. 88

5.2 Potential energy levels of states with the vacancy separated at different distances from the P atom in the P–vacancy system. The first neighbour state has been taken as a base level. 95

5.3 Transition barriers found in the P–vacancy system. The transition between the first and third neighbour states was not found to have an intermediate state. Hence for this case only one energy barrier is given in each direction. 96

5.4 Potential energy levels of different states in the P interstitial defect system, relative to the ⟨110⟩ Fe–P dumbbell state. All states except for the first two are Fe–Fe dumbbell positions near a substitutional P atom. For each of these, two direction vectors are given: the orientation of the dumbbell and the direction of the displacement of the dumbbell centre from the P atom. 101

6.1 Vector basis transformation of the impurity and vacancy coordinates of a sample impurity–vacancy complex in a bcc third neighbour position. The vectors are given in units of the lattice parameter. The first line gives the raw coordinates of the impurity atom and vacancy. These coordinates are transformed first by swapping the x and z coordinates (second line) and then by inverting the sign of y (third line) so that the impurity–vacancy displacement vector is in canonical form. 109

6.2 Vector basis transformations of the transition from third to fourth neighbour P–vacancy states that preserve the canonical form of the P–vacancy displacement and hence are possible transitions from the state orientation being considered. 110

6.3 Diffusion constants for an isolated vacancy in Fe. 113

6.4 Diffusion constants and association times for the P–vacancy complex with a PBC cell of 11 × 11 × 11 bcc unit cells, representing a 0.038 at.% P concentration. The time associated is defined here as the percentage of the time for which the vacancy was within 6.8 Å of the P atom, *i.e.* at up to the eighth neighbour position. 115

6.5 Diffusion constants and association times for the P–vacancy complex with a PBC cell of 15 × 15 × 15 bcc unit cells, representing a 0.015 at.% P concentration. The time associated is defined here as the percentage of the time for which the vacancy was within 6.8 Å of the P atom, *i.e.* at up to the eighth neighbour position. 116

6.6 Diffusion constants for the P–vacancy complex at 500 K with different periodic cell sizes, representing different concentrations of P atoms in Fe. 117

6.7 Diffusion constants and association times for the P interstitial defect system with a PBC cell of $11 \times 11 \times 11$ bcc unit cells. The time associated is defined here as the percentage of the time for which the interstitial defect was within 4 Å of the P atom, *i.e.* at up to the sixth neighbour position. 118

Chapter 1

Introduction

An important branch of modern materials science is the investigation of how materials behave on the atomic scale. Computer modelling plays a major part in these studies. A common form of this modelling is molecular dynamics (MD) simulation, in which the movements of the individual atoms that make up the material are simulated over a span of time. Examples of the ways in which MD simulation has been applied include nanoindentation and nanoscratching [1], used to measure material properties such as hardness and friction, and collision cascades used to simulate the damage caused by radiation on a bulk crystal [2].

MD simulation is performed by using a numerical ordinary differential equation (ODE) integration algorithm to integrate Newton's equations of motion, making use of an interatomic potential to model the interactions between atoms. The distances between atoms are of the angstrom order of magnitude, meaning that only microscopic quantities of material can meaningfully be modelled using MD. Atoms vibrate on the timescale of femtoseconds, and this sets an upper limit on the size of the timestep that can be used in an MD simulation. As such, there is a significant limit to the length of time that can be simulated using MD in a reasonable amount of computing time.

The theory of MD has been studied over many years. Various techniques have been described and implemented to optimise the calculations that are involved in an MD simulation. Computing power has also increased by orders of magnitude over the last few

decades, meaning that larger systems and longer timescales can be modelled with MD simulation now than were possible in the past. Parallel processing has also increased the speed at which simulations can be performed on large systems. However, on the fastest machines available today, even the millisecond timescale is still beyond the reach of MD simulation without modifications to the underlying principle.

To overcome this difficulty, several accelerated dynamics methods have been developed [3]. The basic MD technique is sometimes referred to as *classical MD* to differentiate it from these methods. Accelerated dynamics methods mostly work on the MD principle, but incorporate modifications to the principle to increase the length of time that can be simulated in the same amount of computing time. Methods of this kind include parallel replica dynamics [4], temperature accelerated dynamics [5, 6] and hyperdynamics [7, 8]. Another accelerated dynamics method is kinetic Monte Carlo (KMC) [9], which is not based on MD, but is event-driven. To date, most applications of these techniques have been to surface problems, such as island ripening and crystal growth [3]. In this work, KMC is adapted and applied to a problem involving bulk material.

Accelerated dynamics techniques are based on transition state theory (TST), which is based on the treatment of the evolution of an atomistic system as a sequence of transitions between metastable states. Each of these states corresponds to a local minimum in a $3N$-dimensional potential surface describing the atomistic system, where N is the total number of atoms. For a transition to occur, a sufficient amount of energy must be localised. This is the *energy barrier*, and it is equal to the potential energy difference between the saddle point connecting the two potential energy basins and the local minimum corresponding to the initial state of the transition.

In the MD-based accelerated dynamics methods, transitions between states are detected as they occur during the simulation. In KMC, on the other hand, knowledge of the possible transitions is the driving force. This information can be provided to the simulation as a static, predefined event table, or it can be generated as the simulation progresses; this latter approach is known as on-the-fly KMC (OFKMC). In this work, KMC is performed using a predefined event table. To perform KMC, it is necessary to

have a means of discovering the transitions that are possible from each state that is encountered, and the corresponding energy barriers. There are various algorithms that can be used for this, which work by locating the saddle points of the potential energy surface. In this work, the dimer method [10] is used to locate these saddle points, and therefore determine the transitions in the systems being studied and their associated energy barriers. This method relies only on the first derivatives of the potential function, and on the initial state from which transitions are to be found. A significant portion of the work of this thesis is therefore involved in implementing and applying the dimer method.

The main problem to which the methods are applied in this work is the study of point defects involving P impurity atoms in bulk α-Fe. This problem was chosen because it is important to the study of steel nuclear pressure vessels (NPVs). Steel is an alloy composed mainly of Fe and containing a number of elements, some of which are impurities. Of these impurities, P atoms in particular are known to segregate towards the grain boundaries of the steel. When the NPV is subjected to radiation or enhanced temperature, the impurities become more mobile, thereby enhancing the segregation of P atoms to the grain boundaries, and this is thought to be a major cause of embrittlement [11, 12, 13, 14]. This can lead to a serious reduction in the lifetime of NPVs. It is therefore very important to be able to understand the mechanisms by which this segregation takes place, both to predict the mechanical properties of the NPVs as a function of time and to give pointers towards measures that could be taken to reduce the embrittlement problem. To estimate the changes in properties that occur over time at grain boundaries, engineers often use rate theory models [15, 16]. These can give useful information but require the diffusion rates of the various impurity elements as input parameters. These are often not known experimentally.

To understand better the effect of radiation on the impurity elements in NPVs, molecular dynamics (MD) simulations have been carried out [2]. These simulations were performed in α-Fe with a certain concentration of substitutional P impurities, and identified a number of features that require long-timescale analysis. In this work, some of the defect structures that were found in the results of these simulations are studied using the dimer

method and a newer, more sophisticated interatomic potential than that used in the MD simulations.

It has also been proposed that substitutional P atoms could diffuse through the lattice more easily if they were situated close to a vacancy [17]. If this were to be the case, then the production of vacancies by the collision cascades would be an important factor in the enhancement of P segregation. Thus it is of interest to understand the mechanisms by which P atoms diffuse through α-Fe. The problem is addressed in detail in this work, with the intention of determining the diffusion constants for the most common defect systems.

1.1 Thesis structure

Chapter 2 gives a detailed theoretical description of molecular modelling, including some of the most commonly used interatomic potentials describing how the atoms interact with each other. Interatomic potentials are the driving force behind all forms of molecular modelling. Many interatomic potentials exist for different materials; there are also different potentials for the same material of which some may be more accurate than others, or which may be optimised to different phases or arrangements of the materials involved. The chapter then describes how classical MD simulation works, including some ways in which the technique can be optimised. Also given is an overview of some of the accelerated dynamics methods used to perform simulations on timescales longer than those that can be reached using classical MD.

The dimer method is involved in most of the work of this thesis. Chapter 3 explains this algorithm, including modifications that have been implemented to make the algorithm more efficient and how it has been adapted to the problems being studied. Some benchmarks are given to compare the results obtained by varying the parameters on a simple system, and to test some of the modifications to the algorithm to see if they achieve the desired improvements.

In Chapter 4, the dimer method is applied to the diffusion of an Au adatom on the (100) surface of face-centred cubic Au. This was chosen as a simple atomistic system on which our implementation of the dimer method could be tested. Energy barriers are given for transitions from a single-adatom configuration and various two-adatom configurations. The results are compared with those that were previously found for an Al atom on the Al(100) surface.

The dimer method is then applied to the main problem of this work in Chapter 5. This involves looking at common point defect structures in body-centred cubic Fe. For this study, two different potential models have been used: an Ackland Fe–Fe potential [18] together with a Morse potential for the Fe–P and P–P interactions, and a newer Ackland potential [19] designed to model α-Fe containing P impurities. Four different kinds of defects are covered: an isolated Fe vacancy, an isolated Fe self-interstitial, a P–vacancy complex and a P interstitial defect, each of which can occur in various configurations.

By finding the transitions and energy barriers for the Fe and Fe–P defect structures, an event table is constructed to be used for the KMC work of Chapter 6. The KMC method is explained in detail, including the specific details of the implementation that has been used for this work. Three defect systems are studied: an isolated vacancy, a P–vacancy complex and a P interstitial defect. For each of these systems, KMC is performed at five different temperatures, and the diffusion constants are calculated and compared. This makes it possible to determine whether these systems obey the Arrhenius relation, which is a functional form relating diffusion constant to temperature in typical systems.

Finally, Chapter 7 gives conclusions on the work of this thesis. The results of the studies that have been conducted are summarised and discussed, and suggestions for future work in this field are suggested.

Chapter 2

Molecular modelling

2.1 Interatomic potentials

To model the interactions between atoms, many different interatomic potentials have been developed for a variety of materials. These are functions used to calculate the potential energy of a system of atoms, and hence to determine the forces acting upon the atoms that make up the system. Potentials are typically developed by deciding on a functional form and then fitting the parameters to a finite set of experimental results.

A more recent development is that of *ab initio* methods for modelling interatomic interactions. These methods determine the properties of the elements being modelled, and hence how the atoms interact, given only the atomic number and atomic mass of each element. Typically, approaches such as density functional theory are used for this. These methods can in theory produce more accurate results than those obtained by using empirical potentials. However, *ab initio* calculations are more complicated and computationally expensive than the empirical potentials, so are typically used to model no more than a few hundred atoms at a time. Therefore, approximate potentials continue to be developed, variously using experimental or *ab initio* results as fitting data, and used for larger atomistic systems.

Some potentials are *pairwise*, *i.e.* they consider only the interactions between two atoms at a time. A pairwise potential is essentially a function of the distance between

two atoms. When such a potential is applied to a system of more than two atoms, the potential energy calculated for the system is equal to the sum of the potential energies calculated for all pairs of atoms in the system:

$$U = \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} V_{ij}(r_{ij}) \tag{2.1}$$

where V_{ij} is the pairwise potential energy function itself, r_{ij} is the distance between two atoms i and j, and N is the total number of atoms in the system. The function V_{ij} itself depends on the elements to which atoms i and j belong. Usually when there is more than one element, the V_{ij} functions are of the same functional form but take different parameters. For a pairwise potential, it is necessary that the function is symmetric with respect to the order in which two elements are given, *i.e.* V_{ij} and V_{ji} are the same function.

In practice for many systems, while the sum of pairwise interactions may give an accurate indication of the cohesive energy of a system, they often fail to model such things as the elastic properties accurately. As such, pairwise potentials cannot accurately model such systems. For these, there are *many-body* potentials, which do not make this assumption but consider the interactions between atoms in larger groups.

The following are some examples of interatomic potentials:

- The Brenner [20, 21] potential for modelling hydrocarbons and chemical vapour deposition (CVD) diamond growth.
- The embedded atom [22, 23] and Ackland [18, 19, 24] potentials, many-body potentials often used to model metals.
- The Morse [25, 26] potential, a pairwise potential designed for the modelling of diatomic molecules and some face-centred cubic (fcc) metals.
- The Stillinger-Weber [27] and Tersoff [28, 29] potentials, designed for covalently bonded materials.
- The Ziegler-Biersack-Littmark (ZBL) [30] universal potential for modelling close-range interactions in many materials.

2.1.1 Embedded atom method

The embedded atom method (EAM) was developed by Daw and Baskes [22, 23], and originally used to model impurities such as hydrogen in metals. The principle is that each atom is regarded as being embedded in a field of electrons. The potential function has the form

$$U = \sum_i \left[F_i(\rho_i) + \frac{1}{2} \sum_{j \neq i} V_{ij}(r_{ij}) \right] \tag{2.2}$$

where ρ_i is the electron density of the local environment from which the contribution of atom i has been subtracted. This is approximated as the sum of the contributions of atoms other than i to the electron density at the location of i itself, *i.e.*

$$\rho_i = \sum_{j \neq i} \phi_j(r_{ij}). \tag{2.3}$$

In the original EAM potential, the pairwise interaction takes the form

$$V_{ij}(r) = \frac{Z_i(r) Z_j(r)}{r}. \tag{2.4}$$

The F and Z functions are defined by cubic splines.

Other interatomic potentials based on the EAM principle have been developed, including the Finnis-Sinclair potential [31] and the Ackland potential described in the next section.

2.1.2 Ackland potential

The Ackland potential is a many-body potential based on a cubic spline function fitted at separation distances around the lattice parameter distance. Parameterisations of the Ackland potential have previously been developed for pure metals, including Cu, Ag, Au, Ni [24] and Fe [18]. A more recent development is that of a parameterisation suited to the study of P impurities in α-Fe [19], and this is the potential that has been used for most of this work.

The potential energy of a system of atoms modelled by the Ackland potential takes the form

$$U = \sum_{i=1}^{N} \left[\sum_{j=i+1}^{N} V_{ij}(r_{ij}) + F_i(\rho_i) \right] \tag{2.5}$$

where r_{ij} is the distance between two atoms i and j, and

$$\rho_i = \sum_{j \neq i} \phi_{ij}(r_{ij}). \tag{2.6}$$

The $V_{ij}(r_{ij})$ component represents pairwise interactions between atoms, and the $F_i(\rho_i)$ term encapsulates the effect of electronic kinetic energy (a many-body effect) on the behaviour of the atoms. The functions involved take the forms

$$V(r) = \sum_k a_k (r_k - r)^3 \, \mathrm{H}(r_k - r) \tag{2.7}$$

$$F(\rho) = -\sqrt{\rho} + c_2 \rho^2 + c_4 \rho^4 \tag{2.8}$$

$$\phi(r) = \sum_k A_k (R_k - r)^3 \, \mathrm{H}(R_k - r) \tag{2.9}$$

where H is the Heaviside unit-step function, and the parameters are dependent on the elements involved (the i and j subscripts have been omitted for clarity). Of the parameters for a given pair of elements, the greatest of the R_k and r_k values is the value r_c such that $V(r) = \phi(r) = 0$ for all $r \geq r_c$. The value r_c can thus be regarded as a cut-off distance, and evaluation of the potential can be optimised by considering only atom pairs such that $r_{ij} < r_c$, where the correct r_c value is chosen for the species of the atoms i and j. This is achieved by using a neighbour list structure, as will be described in section 2.2.1.

However, there is a serious limitation in the functional form of the Ackland potential. In reality, as the separation distance between atoms decreases, the potential energy increases without limit, guaranteeing that two atoms can never occupy the same point. This effect cannot be correctly modelled by the potential in this form, since it is built from a finite set of cubic polynomial functions and so cannot increase unboundedly as $r \to 0$. To overcome this difficulty, in this work the ZBL potential has been used in place of $V(r)$ for close-range interactions, and a connecting spline has been defined. Section 2.1.8 gives more information on the connecting spline and how it is used.

Parameters for Au

Au has a face-centred cubic (fcc) structure, with a lattice parameter of $r_L = 4.078$ Å. For the Au work of Chapter 4, the Ackland potential was used with the parameters taken from [24]. Tables 2.1 and 2.2 give these parameters. For the embedding function, $F(\rho) = -\sqrt{\rho}$ is used.

k	a_k (eV)	r_k (Å)	$(r_k/r_L)^2$
1	29.059 1	4.995	3/2
2	−153.148	4.709	4/3
3	148.179	4.559	5/4
4	−22.205 1	4.078	1
5	72.714 6	3.532	3/4
6	199.263	2.884	1/2

Table 2.1: The Ackland potential parameters used for Au–Au pairwise interaction.

k	A_k (eV^2)	R_k (Å)	$(R_k/r_L)^2$
1	21.9301	4.559	5/4
2	284.996	3.532	3/2

Table 2.2: The Ackland potential parameters used for Au–Au many-body interaction.

Parameters for Fe and P

The potential for Fe with P impurities is inherently more complicated than that of pure Au. Since there are two elements, three different V functions are defined, namely those for Fe–Fe, Fe–P and P–P interactions. Correspondingly, there are three parameterisations of the ϕ function, and two of F.

Fe has a body-centred cubic (bcc) structure, with a lattice parameter of 2.855 Å. The parameters taken from [19] were used in this work. Table 2.3 gives the parameters of the pairwise Fe–Fe interactions.

k	a_k (eV)	r_k (Å)
1	−0.003 045 88	5.3
2	−0.058 531 8	4.7
3	0.350 186	4.2
4	−1.026 04	3.7
5	2.657 74	3.3
6	−2.319 44	3.0
7	0.806 564	2.8
8	−0.773 613	2.7
9	4.209 97	2.6
10	−2.498 98	2.5
11	2.207 71	2.4
12	15.738 1	2.3
13	−27.444 8	2.2

Table 2.3: The Ackland potential parameters used for Fe–Fe pairwise interaction.

The Fe–P and P–P pairwise interactions do not follow the standard Ackland potential formula. Instead, a single polynomial function is used in place of the cubic spline. The functions are defined as

$$\begin{aligned} V_{\mathrm{Fe-P}}(r) &= 0.0257790(5.3-r)^{11} - 0.396984(5.3-r)^{10} + 2.54943(5.3-r)^{9} \\ &\quad - 8.81207(5.3-r)^{8} + 17.6293(5.3-r)^{7} - 20.3617(5.3-r)^{6} \\ &\quad + 12.6252(5.3-r)^{5} - 3.31366(5.3-r)^{4} \qquad (2.10) \\ V_{\mathrm{P-P}}(r) &= 0.037557(5.3-r)^{5} - 0.0782938(5.3-r)^{4}. \qquad (2.11) \end{aligned}$$

k	A_k (eV^2)	R_k (Å)
1	0.471 935	4.2
2	−0.014 710 7	3.2
3	11.686 9	2.4

Table 2.4: The Ackland potential parameters used for Fe–Fe many-body interaction.

For the many-body interactions, we use the parameters in Table 2.4 for the Fe–Fe interactions, and

$$\phi_{\mathrm{Fe-P}}(r) = \left(\frac{7}{8}\right)^2 \phi_{\mathrm{FeFe}}(r) \tag{2.12}$$

$$\phi_{\mathrm{P-P}}(r) = \left(\frac{7}{8}\right)^4 \phi_{\mathrm{FeFe}}(r). \tag{2.13}$$

for the Fe–P and P–P interactions. The embedding functions are

$$F_{\mathrm{Fe}}(\rho) = -\sqrt{\rho} - 6.73141 \times 10^{-4}\rho^2 + 7.65149 \times 10^{-8}\rho^4 \tag{2.14}$$

$$F_{\mathrm{P}}(\rho) = -\sqrt{\rho} + 0.00119503\rho^2. \tag{2.15}$$

2.1.3 Morse potential

The Morse potential [25, 26] is a simple pairwise potential, typically used to model diatomic molecules. It is also used to model some face-centred cubic metals for which a pairwise potential has been found to give a reasonable result. The Morse potential has also been used for other applications, such as to model Fe–P and P–P interactions, and applied to the modelling of P impurities in Fe [32].

The potential function is

$$V(r) = D(\mathrm{e}^{2a(r_0-r)} - 2\mathrm{e}^{a(r_0-r)}) \tag{2.16}$$

where r_0 and D are respectively the interatomic distance and potential energy depth at equilibrium of the diatomic molecule.

This potential has no built-in cut-off distance. Under the functional form above, $V(r) = 0$ only for a single value $r = r_0 - \frac{\ln 2}{a}$. In order to implement the optimisation

described in section 2.2.1, it is necessary to modify the potential to have a cut-off distance. This is typically done by using a spline function beyond a certain distance at which the potential energy is small, to connect it smoothly to zero at the cut-off distance.

The Morse potential is also unsuitable for modelling close-range interactions. This is because $V(0) = D(\mathrm{e}^{2ar_0} - 2\mathrm{e}^{ar_0})$ is still a finite value. As such, as with the Ackland and other potentials, it is essential to use a close-range potential such as ZBL instead for interactions at very short distances.

2.1.4 Stillinger-Weber potential

The potential developed by Stillinger and Weber [27] can be used to model covalently-bonded semiconductors such as silicon. These materials exhibit the diamond crystal structure. This structure is unsuitable for modelling by means of a pairwise potential, since such a potential cannot determine the diamond structure to be an energetically favourable configuration. Rather, pairwise potentials stabilise materials towards more compact structures in which each atom has many neighbours, such as the face-centred cubic and hexagonal close-packed structures.

The Stillinger-Weber potential overcomes this difficulty by including a three-body term. This is specially designed to stabilise the angle between bonds in the diamond structure, which is $\cos^{-1}(-\frac{1}{3}) \approx 109.5°$.

The Stillinger-Weber potential function takes the form

$$U = \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \left[V(r_{ij}) + \sum_{k \neq i,j} g(r_{ik}, r_{jk}) \left(\cos\theta_{ikj} + \frac{1}{3} \right)^2 \right] \tag{2.17}$$

where θ_{ikj} is the angle between the bonds ik and jk. It is readily seen that $\left(\cos\theta_{ikj} + \frac{1}{3}\right)^2$ is zero when $\cos\theta_{ikj} = -\frac{1}{3}$, and positive otherwise. The functions V and g are defined by

$$V(\sigma\rho) = \begin{cases} A(B\rho^{-p} - \rho^{-q})\, \mathrm{e}^{(\rho-a)^{-1}} & \text{if } \rho < a \\ 0 & \text{if } \rho \geq a \end{cases} \tag{2.18}$$

$$g(\sigma\rho_1, \sigma\rho_2) = \begin{cases} \lambda \mathrm{e}^{\gamma[(\rho_1-a)^{-1}+(\rho_2-a)^{-1}]} & \text{if } \rho_1 < a \text{ and } \rho_2 < a \\ 0 & \text{if } \rho_1 \geq a \text{ or } \rho_2 \geq a \end{cases} \tag{2.19}$$

for $\rho, \rho_1, \rho_2 < a$. It can be seen that these functions and their derivatives approach zero as $\rho, \rho_1, \rho_2 \to a^-$. Thus V and g are continuous and have continuous first derivatives. The form of the function readily gives $a\sigma$ as the cut-off distance of the potential.

For Si, the parameters are $A = 15.2849$ eV, $B = 0.602225$, $p = 4$, $q = 0$, $a = 1.8$, $\lambda = 45.5323$ eV, $\gamma = 1.2$, $\sigma = 2.0951$ Å.

2.1.5 Tersoff potential

The Tersoff potential, originally designed for the covalent C and Si systems [28], has also been adapted for modelling multi-element systems of covalently bonded atoms [29]. The potential is essentially a sum of interactions between pairs of atoms; however, it is not a true pairwise potential, since the interaction between two atoms depends not only on the distance between them but also on the local environment of each atom in the pair.

The potential energy function is

$$U = \frac{1}{2}\sum_i \sum_{j \neq i} f_C(r_{ij})[f_R(r_{ij}) + b_{ij} f_A(r_{ij})] \tag{2.20}$$

where f_R is a repulsive term like the first term in the Morse potential, f_A is an attractive term, f_C is a cut-off function that makes the potential zero smoothly over an interval between the first and second neighbour distances, and b_{ij} is a many-body term. They are defined by

$$f_C(r_{ij}) = \begin{cases} 1 & \text{if } r_{ij} \leq R_{ij} \\ \frac{1}{2} + \frac{1}{2}\cos\left[\frac{\pi(r_{ij}-R_{ij})}{S_{ij}-R_{ij}}\right] & \text{if } R_{ij} < r_{ij} < S_{ij} \\ 0 & \text{if } r_{ij} \geq S_{ij} \end{cases} \tag{2.21}$$

$$f_R(r_{ij}) = A_{ij} e^{-\lambda_{ij} r_{ij}} \tag{2.22}$$

$$f_A(r_{ij}) = -B_{ij} e^{-\mu_{ij} r_{ij}} \tag{2.23}$$

$$b_{ij} = \chi_{ij}(1 + \beta_i^{n_i} \zeta_{ij}^{n_i})^{-1/2n_i} \tag{2.24}$$

$$\zeta_{ij} = \sum_{k \neq i,j} f_C(r_{ik})\, \omega_{ik} \left(1 + \frac{c_i^2}{d_i^2} - \frac{c_i^2}{d_i^2 + (h_i - \cos\theta_{ijk})^2}\right). \tag{2.25}$$

Most of the two-element parameters are defined simply in terms of one-element parameters. Two of these are arithmetic means, $\lambda_{ij} = \frac{1}{2}(\lambda_i + \lambda_j)$ and $\mu_{ij} = \frac{1}{2}(\mu_i + \mu_j)$. A

further four are geometric means, $A_{ij} = \sqrt{A_i A_j}$ and similarly B_{ij}, R_{ij} and S_{ij}. This leaves only two element-pair parameters independent of the single-element parameters, χ_{ij} and ω_{ij}. These parameters are always 1 between atoms of the same element, and defined symmetrically ($\chi_{ij} = \chi_{ji}$) between atoms of different elements.

	C	Si
A (eV)	1.3936×10^3	1.8308×10^3
B (eV)	3.4670×10^2	4.7118×10^2
λ (Å^{-1})	3.4879	2.4799
μ (Å^{-1})	2.2119	1.7322
β	1.5724×10^{-7}	1.1000×10^{-6}
n	7.2751×10^{-1}	7.8734×10^{-1}
c	3.8049×10^4	1.0039×10^5
d	4.3840	1.6217×10^1
h	-5.7058×10^{-1}	-5.9825×10^{-1}
R (Å)	1.8	2.7
S (Å)	2.1	3.0

Table 2.5: The Tersoff potential parameters for C and Si interaction.

Table 2.5 gives the parameters for C–C and Si–Si interaction. Because of the way the Tersoff potential works, to model C–Si interactions one needs to know only these parameters and the element-pair parameters $\chi_{\mathrm{CSi}} = 0.9776$ and $\omega_{\mathrm{CSi}} = 1$.

2.1.6 Coulomb potential

In ionic compounds, the atoms are held together by attractive forces between positively and negatively charged ions. This attractive force is governed by a simple pairwise potential function

$$V_c(r) = \frac{q_1 q_2}{4\pi\epsilon_0 r} \tag{2.26}$$

where q_1 and q_2 are the electrical charges of the respective ions and $\epsilon_0 \approx 8.85 \times 10^{-12}$ C^2 N^{-1} m^{-2} is the electrical permittivity of the vacuum. It can be readily seen that if q_1 and q_2 are both positive or both negative, $V_c(r)$ is positive and diverges to positive infinity as $r \to 0$, creating a repulsive force. On the other hand, if q_1 and q_2 are of opposite sign, $V_c(r)$ is negative, causing the ions to be attracted to each other.

When the separation distance is given in Å, and the ionic charges are given in units of the elementary charge, the energy in eV can be thus calculated by

$$V_c(r) = \frac{14.3998\ q_1 q_2}{r}. \tag{2.27}$$

The Coulomb potential models only the interactions between atoms as the result of ionic charge. As such, this form by itself would cause positive and negative ions to come together to occupy the same point in space, which is impossible. In practice, the overlapping electron clouds of the two atoms will create a repulsive force between the atomic nuclei as the atoms approach each other, counterbalancing the ionic attraction. It is thus necessary to combine the Coulomb potential with another potential, such as the Buckingham potential [33], to take this into account.

The Coulomb potential is long-ranged compared with most interatomic potentials. As such, a simple neighbour list implementation as described in section 2.2.1 is not an efficient and accurate means of evaluating Coulombic interactions. Instead, special algorihtms are used to evaluate the potential energy and forces that account for these interactions. An example is the Distributed Parallel Multipole Tree Algorithm (DPMTA), which is based on the Fast Multipole Method [34]. This reduces the process of evaluating the Coulomb potential for N atoms from $O(N^2)$ to $O(N \log N)$ time complexity.

2.1.7 Ziegler-Biersack-Littmark (ZBL) potential

The potential developed by Ziegler, Biersack and Littmark [30] is a pairwise potential used to model close-range interactions. It is an improvement of the earlier screened Coulomb potential [35]. At small distances, the main interaction between atoms is the Coulombic repulsion of their nuclei. As the distance increases, the surrounding electrons

come into play, gradually screening the positive nuclear charges from each other and thus decreasing the repulsion.

The ZBL potential function is thus obtained by multiplying the Coulombic potential energy between nuclei by a function $\phi(r)$ that represents the electron screening. The potential energy in eV is defined by

$$V(r) = \frac{14.3998\ Z_1 Z_2}{r}\phi(r) \tag{2.28}$$

$$\phi(r) = 0.18175e^{-3.1998r/a_u} + 0.50986e^{-0.94229r/a_u} + 0.28022e^{-0.4029r/a_u} + 0.028171e^{-0.20162r/a_u} \tag{2.29}$$

$$a_u = \frac{0.468377}{\sqrt{{Z_1}^{2/3} + {Z_2}^{2/3}}} \tag{2.30}$$

where Z_1 and Z_2 are the atomic numbers of the elements being considered.

It can be seen that $\phi(r)$ decreases towards zero as r increases, hence the ZBL potential is strictly repulsive. Since $\phi(0) = 1$, as $r \to 0$ the potential energy approaches that of pure Coulombic repulsion between the nuclei of the two atoms.

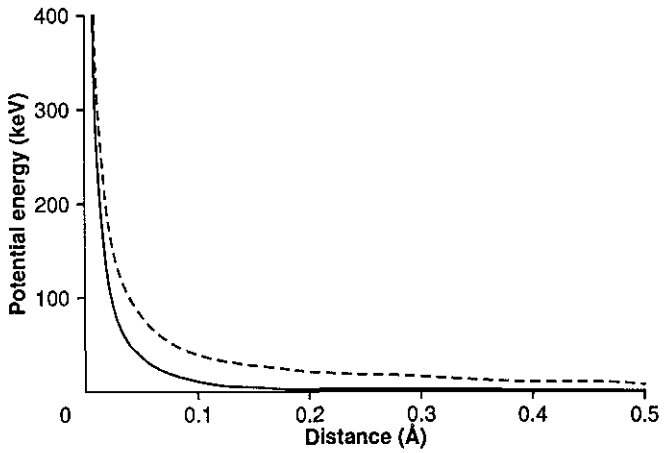


Figure 2.1: The ZBL potential appplied to Fe–Fe interaction (solid line) and the potential energy associated with simple Coulombic repulsion of Fe nuclei (dashed line).

At distances around and beyond the nearest neighbour distance of the atomic lattice, Coulombic repulsion of nuclei practically disappears and other forces predominate. Thus

the ZBL potential is useful only to model interactions at distances smaller than some value typically between 1 Å and 2 Å, as may occur during the simulation of a collision cascade. Therefore, it is common to configure simulations to use the ZBL potential only when the distance between atoms is sufficiently small, with a suitable potential chosen for greater distances and a connecting spline between the two. See section 2.1.8 for more about how this is done.

2.1.8 Connecting spline function

Often it is necessary to define a potential function to connect a potential used for close-range interactions, such as the ZBL potential, to one used for interactions at distances around the bond lengths of the material and beyond. This connecting function is a single-segment spline interpolation between the extremities of the distance ranges to which the two potentials are respectively applied.

For this work, the functional form

$$V_S(r) = \mathrm{e}^{B_0+B_1r+B_2r^2+B_3r^3}. \tag{2.31}$$

was used for the spline function. A four-parameter functional form is used because it must be fitted to four known values in order to connect smoothly with the close-range potential and the potential used around the bonding distance, so that $V(r)$ is continuous and has continuous first derivatives across all $r > 0$. Since the forces acting at the intermediate distance are essentially repulsive, the function should have no local maximum or minimum within the range of distances to which it is applied, as would tend to occur if a simple polynomial function is used. Studies have shown that the exponential form that has been chosen tends to produce a strictly decreasing potential energy curve over the applicable range of distances.

To fit the coefficients in the connecting function, we solve the simultaneous equations

$$V_S(r_L) = V_L(r_L) \tag{2.32}$$

$$V'_S(r_L) = V'_L(r_L) \tag{2.33}$$

$$V_S(r_U) = V_U(r_U) \tag{2.34}$$

$$V_S'(r_U) = V_U'(r_U) \tag{2.35}$$

where V_S is the connecting spline function, r_L and r_U are the lower and upper bounds of r for which V_S is being used, and V_L and V_U are the pairwise components of the potentials chosen for $r < r_L$ and $r > r_U$ respectively.

If a many-body potential is being used for $r > r_U$, then its many-body component (what is left after V_U is subtracted) is applied across all atoms of elements to which it applies, regardless of their separation distance. Thus many-body effects are taken to apply across all separation distances, with pairwise interactions being calculated by a function dependent on the distance between atoms and added to the many-body component to determine the total potential energy.

In this work, a connecting spline was used to connect the Ackland and ZBL potentials in the Au and Fe–P systems. For Au, the spline was used for the range 1.7 Å $< r <$ 2.85 Å. This range was chosen as it was found to give a good approximation of the potential energy and force of the Ackland and ZBL potentials near its respective boundaries, and a good compromise between the predictions of both potentials over its range. Solving the equations of continuity gives

$$V_S(r) = \mathrm{e}^{8.08776 - 2.29308r - 1.13806r^2 - 0.0622234r^3}. \tag{2.36}$$

Figures 2.2 and 2.3 illustrate the continuity of this spline function with the Ackland and ZBL potentials.

For the Fe and P system, the range over which the spline was applied was that used by Ackland. The spline functions are

$$V_{\mathrm{Fe-Fe}}(r) = \mathrm{e}^{7.41227 - 0.641807r - 2.60435r^2 - 0.626254r^3} \quad (1.0 < r < 2.05) \tag{2.37}$$

$$V_{\mathrm{Fe-P}}(r) = \mathrm{e}^{10.7619 - 10.0040r - 4.98543r^2 - 1.25998r^3} \quad (1.0 < r < 2.0) \tag{2.38}$$

$$V_{\mathrm{P-P}}(r) = \mathrm{e}^{9.93828 - 8.56372r - 3.45196r^2 - 0.614538r^3} \quad (0.9 < r < 2.5). \tag{2.39}$$

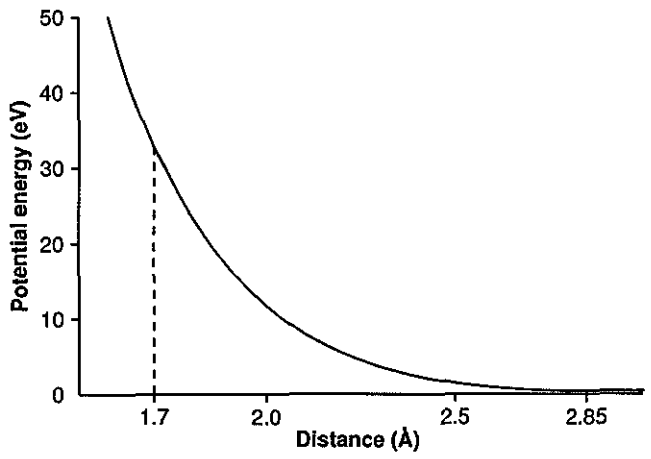


Figure 2.2: Graph of $V(r)$ for Au–Au interaction. The spline function connects the ZBL potential ($r \leq 1.7$) and the pairwise part of Ackland's potential function ($r \geq 2.85$).

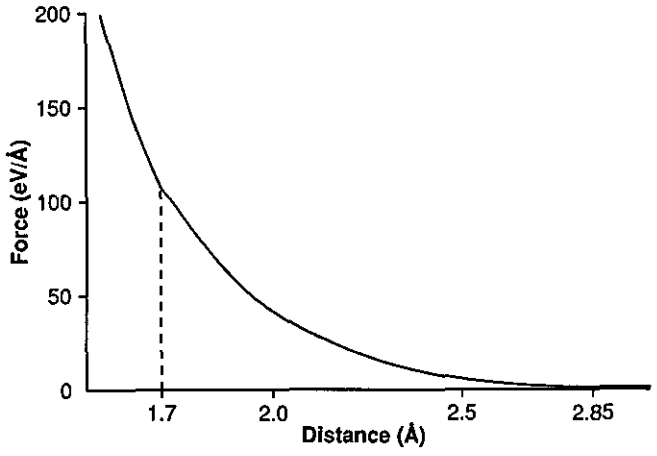


Figure 2.3: Graph of $-V'(r)$ for Au–Au interaction. The spline function connects the ZBL potential ($r \leq 1.7$) and the Ackland pairwise function ($r \geq 2.85$).

2.2 Molecular dynamics simulation

The classical molecular dynamics (MD) approach uses a numerical algorithm to integrate Newton's equations of motion. The analytical first derivative of the potential function, with respect to the coordinates of the atoms in the system, gives the force acting upon the system. By using any of various numerical methods to integrate ordinary differential equations (ODEs), the trajectory of the atoms in the system can be simulated.

This approach is adequate for the simulation of processes taking place over very small timescales. However, because of the high frequency of atomic vibrations, very small timesteps of the order of femtoseconds are needed. As such, simulations are often terminated after tens of picoseconds, because the approach cannot be used for longer timescales in a reasonable amount of computing time.

The LBOMD software, developed at Loughborough University, implements molecular dynamics simulation. It can perform simulations of nanoindentation, collision cascades and simple diffusion.

To advance the system from timestep to timestep, LBOMD uses the Velocity Verlet algorithm. Whilst being only a second-order method, its advantage over other numerical ODE integrators, such as the Runge-Kutta methods, is that it is *symplectic*, *i.e.* it preserves Hamiltonian invariants, and hence is efficient at conserving the total energy of the system over a large number of timesteps [35].

The Velocity Verlet method advances the trajectory of atoms using a fixed finite timestep Δt. The position and velocity of atom i at each timestep are determined by

$$\mathbf{r}_i^{[k+1]} = \mathbf{r}_i^{[k]} + \mathbf{v}_i^{[k]}\Delta t + \frac{\mathbf{F}_i^{[k]}(\Delta t)^2}{2m_i} \tag{2.40}$$

$$\mathbf{v}_i^{[k+1]} = \mathbf{v}_i^{[k]} + \frac{(\mathbf{F}_i^{[k]} + \mathbf{F}_i^{[k+1]})\Delta t}{2m_i} \tag{2.41}$$

where $\mathbf{F}_i^{[k]}$ is the force vector acting upon atom i at the k'th timestep, and m_i is the mass of atom i.

2.2.1 Cut-off distances and neighbour lists

In principle, the formula for the total potential energy of the system considers all possible pairs of atoms in the system. This means that, for both pairwise and many-body potentials, if all interactions are included then calculation of the total potential energy of a system would be an $O(N^2)$ operation.

In practice, there is nearly always a distance beyond which the interaction between two atoms becomes negligible. (The main exception is the Coulomb potential described in section 2.1.6, for which other techniques are typically used.) Interatomic potentials typically make use of this by having a cut-off distance, *i.e.* a distance beyond which the potential energy associated with a pair of atoms is taken to be zero. By excluding such pairs of atoms from the calculations, the efficiency can be increased considerably.

To achieve this improvement in efficiency, a neighbour list structure is used. This lists, for every atom in the system, those atoms that are within the cut-off distance r_c plus a small defined length r_s. This small length creates a 'skin' of atoms just outside a given atom's cut-off boundary. This allows for some movement of the atoms before it is necessary to rebuild the neighbour list.

The simplest way to build a neighbour list is to iterate through all pairs of atoms and pick out those pairs within the appropriate distance. However, this would take $O(N^2)$ time to perform. The process is therefore optimised using a spatial decomposition. This partitions space into a number of equal cuboid cells, with every side length equal to at least the maximum cut-off distance in the system plus the skin thickness. That is, for lattice dimensions $L_x \times L_y \times L_z$, we have $M_x \times M_y \times M_z$ cells, each of dimensions $l_x \times l_y \times l_z$, where $l_x, l_y, l_z >= r_c + r_s$. If there are multiple interactions with different cut-off distances, then the maximum of them must be taken. The algorithm assigns each atom in turn to the correct cell, using a linked list structure so that the atoms of each cell can then be iterated through. This is an $O(N)$ process.

The spatial decomposition is then used to build the neighbour list of every atom in the system. An atom can have neighbours only in the same cell and in the immediately

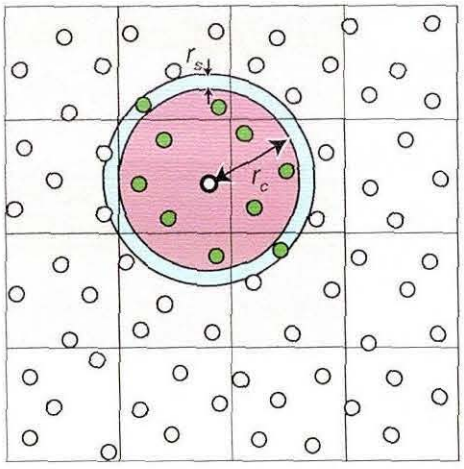


Figure 2.4: Schematic illustration of the spatial decomposition method of building the neighbour list. The box of atoms is partitioned into cells, and the algorithm looks through the atom lists of the shaded cells for neighbours of the atom with a bold border. The result is that the shaded atoms are picked out. In this instance two atoms are in the skin – they do not interact with the central atom at the moment, but are included in case they move to within the cut-off radius in the time between neighbour list updates.

adjacent cells (a maximum of 27 cells); Figure 2.4 illustrates this. The algorithm iterates through the atom lists of these cells to find atoms that are within the required distance. Thus building the neighbour list has $O(DN)$ time complexity, where D is the maximum number of atoms in a cell of the spatial decomposition. Since D generally has an asymptotically linear relationship with the maximum number of an atom's neighbours, it follows that for systems of the same material in the same atomic density, the time taken to build the neighbour list increases linearly with the system size.

To determine when it is time to rebuild the neighbour list, at each step of the MD simulation a measurement is taken of how far the atoms have moved since the neighbour list was last constructed. If the distance between any two atoms not in each other's neighbour lists has decreased by more than the skin thickness, then it is possible that the two atoms have moved to within the cut-off distance of each other, and therefore the pair needs to be included in the potential calculation. A simple $O(N)$ calculation is used

at each timestep to check for this possibility, which is to find the two atoms that are at the greatest distances from where they were when the current neighbour list was built. When the sum of these two greatest distances exceeds the skin thickness, the neighbour list is rebuilt.

To calculate the potential energy and forces, only the interactions between atoms that are within each other's neighbour lists are considered. In equations such as (2.1) and (2.5), the j sums are taken to include only those atoms that are in the neighbour list of atom i. The result is the same as that which would be calculated by including all atom pairs (i, j) in the sum; using the neighbour list simply speeds up the process, reducing the time complexity of the potential calculation from $\mathrm{O}(N^2)$ to $\mathrm{O}(DN)$.

2.2.2 Boundary conditions

A molecular dynamics simulation always takes place over a finite set of atoms over a finite space. Often it is desirable to treat this space as a section of a larger body of material. This is achieved by implementing periodic boundary conditions (PBC) in the simulation. Effectively, the box of atoms is treated as if space is tiled with identical copies of it extending indefinitely.

In a given simulation, PBC may be used in any or all of the three dimensions. A simulation involving a surface, such as adatom diffusion, would be carried out with PBC in two dimensions. Simulations involving bulk material, on the other hand, are generally performed using PBC in all three dimensions.

PBC affect which atoms interact with each other, and hence influences the evaluation of the potential. It also influences how the atom positions are updated, so that the atom coordinates, in those dimensions where PBC are active, are always within those that define the box.

When evaluating the potential energy and forces, it is necessary to consider which atoms interact with each other not only within the box, but also across periodic boundaries. These atoms are identified as part of the process of building the neighbour lists. If the x dimension has PBC, then in the spatial decomposition used in building the neigh-

bour list, cells with $0 \leq x < l_x$ and $L_x - l_x \leq x < L_x$ are treated as adjacent to each other if their ranges of y and z are equal or adjacent.

When building the neighbour list, a note is made of the direction(s) in which a neighbour crosses a periodic boundary. This information is then used when evaluating the potential and forces in order to calculate the interaction distance between the atoms. Figure 2.5 illustrates the principle. For example, if atom j is across the periodic boundary only in the positive x direction relative to atom i, then the distance between the points (x_i, y_i, z_i) and $(x_j + L_x, y_j, z_j)$ must be taken.

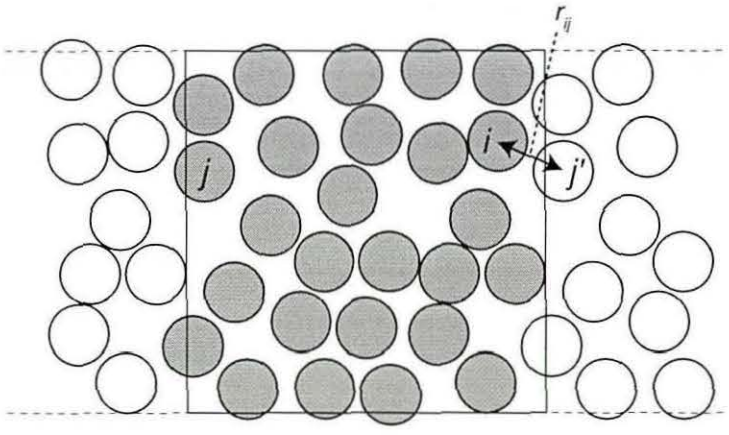


Figure 2.5: Schematic illustration of periodic boundary conditions applied in a single dimension. The rectangle represents the box within which the simulation takes place, and the shaded circles the atoms within it, which are the only atoms included in the simulation data. This arrangement of atoms is considered to repeat itself either side of the box (unshaded atoms). Consequently, the distance r_{ij} is taken to be the distance between the atom i and the atom j' in the adjacent periodic cell corresponding to j.

2.2.3 Parallel processing

Molecular dynamics simulation is inherently a computationally intensive process. One reason is the very small timesteps, typically of the order of femtoseconds, that are required to model the atomic vibrations. The second reason is the large number of atoms in a

typical system. The angstrom scale of the interatomic distances means that billions of atoms are required to model systems of any size approaching even a cubic micron.

This second cause of computational intensiveness can be countered by employing many processors to perform the simulation, dividing the work of evaluating the potential energy and forces at each timestep between the processors.

The simplest form of parallel MD is an atom decomposition (AD) method. In this approach, each atom is assigned to a processor, and remains assigned to the same processor throughout the simulation. Each processor has N/P atoms assigned to it, where P is the number of processors. The processor to which an atom is assigned is responsible for calculating the forces acting on it and for updating its velocity and position vectors. The inefficiency of the AD method is that each processor must know the coordinates of all N atoms in the system at any time. Thus under a message-passing parallel paradigm, each processor has its own copy of the system, occupying a very large amount of memory and taking a significant amount of time to communicate the information between processors. However, using a shared memory parallel machine, in which all processors share a single copy of the atomistic system, AD is an efficient parallelisation technique, and its simplicity can be an advantage.

A variation of the AD technique is a force decomposition (FD) [36]. Under this method, the force matrix is decomposed into rectangular blocks to be assigned to processors. Each block has two sets of atoms associated with it: the exerting atoms and the receiving atoms. Each processor calculates the effect of its block's exerting atoms upon its block's receiving atoms. The forces on each receiving atom are consolidated into a local sum, which is then summed across all blocks having the same receiving atom set in order to update the position and velocity of each atom. The FD technique thus requires less communication between processors than AD, since each processor needs to know only the positions of atoms in its exerting and receiving sets. In an optimal setup with a square number of processors assigned $N/\sqrt{P} \times N/\sqrt{P}$ blocks of the force matrix, the number of atom positions required by each processor is $2N/\sqrt{P} - N/P$. The subtracted

term is achieved by a permutation of the force matrix so that every block has N/P atoms in the intersection of its exerting and receiving sets.

An alternative, implemented in LBOMD, is the spatial decomposition (SD) approach. The box of atoms is subdivided into P cuboid regions. Each processor has assigned to it the atoms that are within its region. This is similar to the atom decomposition approach, except that a rigid principle determines which processor an atom is assigned to, and atoms can move between processors. The assignment of atoms to processors by location can considerably reduce the amount of communication between processors compared with AD and FD.

Under SD, the only information that needs to be constantly passed between processors is the positions of the atoms that are near the boundaries of the atom regions. This enables each processor to calculate the potential energy and force applied to every atom within its region, by having the coordinates of the atoms in neighbouring regions that may be neighbours of its own atoms. It is convenient if each processor maintains a list of atoms that it needs to pass to each of its neighbouring regions. At any time, the neighbour list of any atom is a subset of the set of atoms that the processor to which it is assigned has either in its own region or passed from neighbouring regions. For this reason, the list of atoms whose coordinates are communicated between processors needs to be updated only just before the neighbour list is updated.

Of course, an atom can move from one SD region to another. When this happens, the processor to which it is assigned changes. This guarantees that the atom correctly interacts with other atoms in its new region. Similarly, this needs to be done only when atoms have moved sufficiently that it is time to rebuild the neighbour list. Thus when it is time to build the neighbour list, there are three stages to the process: moving any atoms that need to be moved between processors, building the list of atoms whose coordinates will be passed to each neighbouring cell, and building the neighbour list itself.

Spatial decomposition as an approach to parallel MD is separate from the spatial decomposition used to optimise the process of building the neighbour list as described in

section 2.2.1. When SD parallelisation is used, the regions that are assigned to processors are in turn subdivided in order to build the neighbour list.

2.3 Accelerated dynamics methods

Some atomic-scale processes, such as diffusion, are made up of events that rely on a significant amount of energy being localised in order for them to happen, and are therefore infrequent on the timescale of atomic vibrations. As such, classical MD simulation is very limited in its ability to model these processes. Some studies of diffusion have been successfully carried out using MD simulation, but the accuracy of any results obtained is inherently limited by the accessible MD timescale.

To enable longer timescales to be modelled, several accelerated dynamics methods have been developed, studied and reviewed [3]. This section gives a summary of some of these techniques.

All of the accelerated dynamics methods are based to some extent on transition state theory (TST) [3, 37], a model that describes the dynamics of a system in terms of transitions between metastable states. This theory makes it possible to estimate actual transition times from the times that the transitions take under the accelerated conditions. A useful approximation to TST is harmonic transition state theory (hTST) [38], which provides a simple means of estimating the transition probabilities and escape times if the energy barriers are known. A brief description of hTST will be given in Chapter 6.

Some of the techniques work on the principle that an atomic system can be regarded as a high-dimensional potential field. That is, a system of N atoms has $3N$ degrees of freedom, and can therefore be considered as a $3N$-dimensional potential surface in which each coordinate denotes one of the three position coordinates of a single atom. Transitions then correspond to the saddle points of this surface. The minimum amount of energy required for a transition to occur is then determined by the height of the saddle point.

2.3.1 Parallel replica dynamics

Section 2.2.3 describes methods of parallelising a single MD simulation. Parallel replica dynamics (PRD), developed by Voter [4], is an alternative way of utilising multiple processors, by effectively running a separate MD simulation on each. This is the simplest of all accelerated dynamics techniques.

The system is replicated across all the processors being used. Each processor then carries out a classical MD simulation, with an initial dephasing stage so that each processor follows a different trajectory. During the dephasing stage, the velocities of all atoms that are allowed to move are randomised periodically.

When a transition has occurred in one copy of the system, all processors stop exploring the phase space. In a typical system, the escape time follows an exponential distribution, with the probability density function

$$p(t) = \frac{\mathrm{e}^{-t/\tau}}{\tau}. \tag{2.42}$$

Under this condition, it has been shown that the total accumulated time (excluding the dephasing stage) over all of the replicas is a reasonable estimate of the time that the discovered transition would have taken to occur.

Since a transition occurs when sufficient energy is localised, it is possible that another transition will occur during the excitation period resulting from this localisation of energy. This may be a transition back to the original state, or to a new state. The typical length of time for which this excitation period continues after the initial transition is known as the correlation time, since any transition occurring during this time is correlated with the transition that led to it. To accommodate for this possibility, the trajectory that led to the transition is followed for a length of simulation time equal to at least the correlation time, in order to catch any correlated transition that may occur. Then, in order to find further transitions from the new state, the new configuration is replicated and the process begins again.

The efficiency of PRD depends on the average time taken for a transition to occur compared with the correlation time. There are two reasons for this. Firstly, the simulation

time clock is not incremented during the dephasing stage, since dephasing is a process to generate different initial conditions for each processor and not part of the evolution of the system. Secondly, during the post-transition correlation time only one processor is at work, preparing the system for the next replication. A further cause of slowdown lies in the process of detecting that a transition has occurred. A possible way of implementing this is to perform a steepest descent or conjugate gradient relaxation at intervals, and compare the resulting configuration of atoms with that of the potential energy basin of the state from which it came. The efficiency of the process therefore also depends on choosing an optimal time interval at which to check for a transition.

However, PRD has its benefits compared with classical MD simulation parallelised in the system size domain. Firstly, for systems with a relatively small number of atoms, any of the decomposition-based parallel MD approaches could result in a significant amount of time being spent communicating between processors compared with that spent evaluating the potential energy and forces. Thus PRD is useful as a means of modelling the evolution of small systems over longer timescales rather than as a technique for modelling larger systems. Effectively, it achieves parallelisation in the time domain, rather than that of the number of atoms in the system.

A second respect in which PRD can outperform parallel classical MD is if the processors are unbalanced in speed or load. Load imbalance is especially liable to occur when using an atom-based or region-based decomposition scheme if parts of the system vary in atomic density or in the computational cost of the potentials being used. Under PRD, since each processor evolves a trajectory independently of the others, there is no latency during the parallel parts of the simulation caused by processors waiting for each other to finish their timesteps before they can proceed with the next. Each processor can run at its own speed, meaning that the only time a processor is idle is after a transition has been recorded, while waiting for one processor to follow through the system for correlated events.

Systems studied using PRD have included an Ag(111) island-on-island configuration [3]. On a system of thirty-two 1-GHz Pentium III processors, the method took five

days to reach a simulation time of 1 μs at a temperature of 400 K. The atoms of the upper island were found to sink completely into the lower island after 0.45 μs.

2.3.2 Hyperdynamics

The hyperdynamics technique [7, 8] adds a bias to the potential function defining the system. The bias potential is a function that is non-negative at all points and zero at the dividing surfaces between states. This effectively reduces the energy barrier enabling the system to move from state to state more quickly, while preserving the relative probabilities of the various possible escape paths.

The process of constructing a bias potential is a complex topic. One possibility is to make use of the eigenvalues of the Hessian matrix of the system. Since the lowest eigenvalue is always negative at the dividing surface, the bias potential can be made zero wherever the lowest eigenvalue is zero or negative, and a function of the lowest eigenvalue otherwise. It is also common to make the bias potential a function of the two lowest eigenvalues and g_{1p}, the projection of the force vector onto the lowest eigenvector.

Evaluating and diagonalising the Hessian is a computationally expensive process, but an iterative numerical method has been developed that can be used to calculate the two lowest eigenvalues using only first derivatives of the potential function [8].

The computational speed-up achieved by hyperdynamics depends on two factors: the boost factor achieved by the bias potential and the computational overhead involved in evaluating it. The boost factor is determined by

$$\frac{t_{hyper}}{t_{MD}} = \left\langle \mathrm{e}^{\Delta V(\mathbf{r})/k_B T} \right\rangle \tag{2.43}$$

where ΔV is the bias potential function, k_B is Boltzmann's constant, T is temperature and the average is taken over successive timesteps. It follows that the boost factor decreases exponentially with increasing temperature [3]. For hyperdynamics to achieve the aim of speeding up a simulation, it is important that the bias potential achieves a sufficiently high boost factor to be worth the computational cost of evaluating it. This

temperature dependence of the boost factor means that constructing an efficient bias potential becomes more challenging if higher temperatures are to be modelled.

A weakness of hyperdynamics is that it is necessary to decide in advance the height of the bias potential at the potential energy minima. If this is too low, then the boost factor can be considerably sub-optimal. If the bias is too high, the technique will tend to miss some of the transitions with lower energy barriers.

Voter [8] has tested the hyperdynamics method by applying it to two Ag surface self-diffusion problems. The first of these is an adatom on the Ag(100) surface. At a temperature of 400 K, the bias potential used achieved a boost factor of 1356, with each MD step taking approximately 30 times that of a classical MD step on the same system, giving a net computational boost of 45. The second system studied was an island of 11 atoms on an Ag(111) surface at 300 K. In this system, the island atoms generally moved as a cluster, including many hops between the fcc and hcp stacking positions.

Hyperdynamics has been studied extensively by Sanz-Navarro [39, 40], who also used it to study Ag surface diffusion. In these studies, an approximate method [39] was used to calculate the forces acting on the system as a result of the bias potential, without the need to evaluate g_{1p}. This approximation was found to give satisfactory results in test systems, and for both a single adatom and an 11-atom cluster on the Ag(100) surface.

2.3.3 Temperature-accelerated dynamics

The temperature-accelerated dynamics (TAD) method [5, 6] works by evolving an MD trajectory at a temperature higher than that at which the evolution is to be predicted. A basin-constrained MD (BCMD) is used, in which the trajectory is reflected back into the basin when it tries to escape. This generates a list of escape paths and times, from which it is possible to estimate the expected time at which each transition would occur at the temperature of interest.

In this process, it is essential to filter out transitions that will not occur as a result of the energy constraints at the lower temperature. Once the BCMD simulation has run

for sufficient time, the transition with the lowest escape time is accepted and the TAD procedure is started again in order to find the next transition.

TAD relies on hTST to calculate the escape time at the temperature of interest. This escape time can be determined by

$$t_{low} = t_{high} \mathrm{e}^{E(1/T_{low} - 1/T_{high})/k_B} \tag{2.44}$$

where T_{high} is the temperature at which the TAD simulation is being run, T_{low} is the temperature at which the dynamics are to be predicted, t_{high} is the time the system takes to leave the basin at T_{high} and E is the activation energy of the transition. The BCMD trajectory can be stopped once it has run for a time of

$$t_{stop} = \frac{\ln(1/\delta)}{\nu_{min}} \left(\frac{\nu_{min} t_{low,min}}{\ln(1/\delta)} \right)^{T_{low}/T_{high}} \tag{2.45}$$

where ν_{min} is an assumed lower bound of the attempt frequency, $t_{low,min}$ is the minimum of all escape times determined for T_{low} and δ is a parameter determining the desired accuracy of the simulation. When this time has elapsed, we have a confidence factor of $1 - \delta$ that the transition that gave the escape time of $t_{low,min}$ is indeed the expected escape from the initial state. The boost factor is thus $\frac{t_{low,min}}{t_{stop}}$.

TAD is especially suited to being used in conjunction with classical MD to model crystal growth. By using TAD interspersed with classical MD simulations to model the deposition events, experimental timescales have been reached for Ag/Ag(100) [6] and Cu/Ag(100) [41] systems at temperatures up to 70 K.

2.3.4 Kinetic Monte Carlo

The kinetic Monte Carlo (KMC) approach is different from other accelerated dynamics methods in that MD simulation is not used. Instead, a method is used for finding the possible transitions from state to state, the associated energy barriers and hence the probability and expected escape time of each. From this information, at each step of the KMC process a transition from the current state is chosen at random, weighted according

to the relative probabilities. Chapter 6 explains in detail the KMC method and how it is applied in this project.

However, the list of transitions can be very large, making it difficult to compile an adequate catalogue of transitions for use in KMC. The approach used by Henkelman and Jónsson [9] is to build a catalogue on the fly of transitions from each state encountered during the process. This is known as on-the-fly KMC (OFKMC). To build the catalogue of transitions, the dimer method is used. This method is described in Chapter 3. Systems that have been studied using OFKMC with the dimer method include island ripening of 20 atoms deposited on an Al(100) surface [9]. In this system, the adatoms joined to make a single island in its most compact shape after 65 270 transitions. This system took about a week of computer time to simulate on a single-processor system.

Chapter 3

The dimer method

3.1 Barrier calculation methods

A number of techniques have been used in the past to determine the energy barriers of transitions between states in atomistic systems. These methods generally fall into two categories. The first category consists of methods used to determine the transition barriers between known initial and final states. The second category is those methods that, given only an initial configuration of atoms, can discover transitions that are possible from this state and calculate their energy barriers. Methods of both kinds generally work on the principle of representing a system of N atoms as a $3N$-dimensional potential surface, as described in the introduction to section 2.3.

3.1.1 Molecular statics

Molecular statics (MS) [42] is a means of calculating the energy barrier of a transition given its initial and final states, based on a simple constrained optimisation. Let $\mathbf{R}_i$ and $\mathbf{R}_f$ be the position vectors denoting the initial and final states of the transition path to be examined. The function

$$\mathbf{R}(\alpha) = (1 - \alpha)\,\mathbf{R}_i + \alpha \mathbf{R}_f \tag{3.1}$$

is evaluated for a set of values $\alpha \in (0, 1)$. The system is relaxed using an optimisation algorithm such as a steepest descent or conjugate gradient method, starting from each $\mathbf{R}(\alpha)$ in turn, fixing the vector component parallel to $\mathbf{R}_f - \mathbf{R}_i$ in each case. This produces a sequence of minimum energy levels at successive positions along the transition path. The relaxed configuration having the highest energy level along the path is taken to be the saddle point of the potential surface, and the difference between its energy level and that of the initial state is taken to be the energy barrier. Typically a set of equally-spaced α values is used, and this also makes it possible to plot the energy profile of the transition path. However, the relaxed configurations for such a sample of α values is unlikely to include the actual saddle point and therefore the exact energy barrier. A line optimisation algorithm, such as a golden section search, can be used to find more accurately the value of α that gives the highest relaxation energy and therefore the energy barrier.

An MS calculation uses derivatives of the potential only to the order required by the underlying optimisation algorithm chosen to relax each configuration. The steepest descent and conjugate gradient algorithms both rely only on first derivatives, so using such an algorithm for the relaxations gives a barrier calculation method that does not involve the calculation of the second or higher derivatives.

The MS technique has been used previously to study the behaviour of self-interstitial defects in hcp Zr, bcc Mo [42] and bcc Fe [2, 42]; these Fe studies were carried out using older potentials [18, 43] than that used in Chapter 5 of this work.

3.1.2 The nudged elastic band method

An alternative to molecular statics is the nudged elastic band (NEB) method [44], which is an example of a chain-of-states method that uses only first derivatives of the potential function. In chain-of-states methods, several points on the $3N$-dimensional potential surface, denoting distinct images of the atomistic system, are connected in sequence to trace out a path. Two of these images are fixed; these are the initial and final states of the system, and they serve as the endpoints of the chain. Under NEB, elastic springs of zero natural length connect successive images of the system. An iterative method similar

to an MD simulation is used to converge the sequence of images towards the minimum energy path between the initial and final states. The saddle point is then the maximum energy point along this path, giving the energy barrier of the transition. The effective force acting upon image i of the system is determined by

$$\mathbf{F}_i = \mathbf{F}(\mathbf{R}_i)^{\perp} + (\mathbf{F}_i^s \cdot \hat{\mathbf{T}}_i)\,\hat{\mathbf{T}}_i + f(\phi_i)\,\mathbf{F}_i^{s\perp} \tag{3.2}$$

where $\mathbf{F}(\mathbf{R}_i)$ is the force vector acting on the image as determined by the potential function, $\mathbf{F}_i^s$ is the spring force vector, $\hat{\mathbf{T}}_i$ is the unit vector parallel to the elastic band's tangent at image i, f is a switching function and ϕ_i is the angle between the vectors $\mathbf{R}_i - \mathbf{R}_{i-1}$ and $\mathbf{R}_{i+1} - \mathbf{R}_i$. The vectors $\mathbf{F}(\mathbf{R}_i)^{\perp}$ and $\mathbf{F}_i^{s\perp}$ are defined by

$$\mathbf{F}^{\perp} = \mathbf{F} - (\mathbf{F} \cdot \hat{\mathbf{T}}_i)\,\hat{\mathbf{T}}_i \tag{3.3}$$

and the function f by

$$f(\phi) = \begin{cases} \frac{1}{2}(1 + \cos(\pi \cos\phi)) & \text{if } -\frac{\pi}{2} < \phi < \frac{\pi}{2} \\ 1 & \text{otherwise.} \end{cases} \tag{3.4}$$

The form of $\mathbf{F}_i$ causes the spring forces to act only along the local tangent of the elastic band, while the true forces are projected onto the normal hyperplane before being applied. This means that the spring forces and true forces do not interfere with each other, bringing about an approximately equal spacing between images while allowing the minimum energy path to be determined correctly. This eliminates some of the problems with NEB's predecessor, the plain elastic band method [44], in which the spring forces could prevent the minimum energy path from being correctly determined, or the forces acting on individual images could cause the images to slide down from the saddle point. The term involving $f(\phi_i)$ is to counteract kinks in the path that may otherwise occur in some systems.

A limitation of NEB is that it remains possible that two consecutive images remain either side of the saddle point being sought, meaning that the energy barrier of the transition is not accurately determined. A refinement of NEB, the climbing image NEB method [45], eliminates this problem as follows. After a number of iterations of regular

NEB, the highest-energy link of the chain is chosen to be a 'climbing image'. This image then has its effective force determined not by equation 3.2, but instead by

$$\mathbf{F}_i = \mathbf{F}(\mathbf{R}_i) - 2\left(\mathbf{F}(\mathbf{R}_i) \cdot \hat{\mathbf{T}}_i\right)\hat{\mathbf{T}}_i, \tag{3.5}$$

i.e. the spring force is ignored for this image, and the actual force is reflected in the direction of the tangent vector. By continuing the process under this modification, the climbing image will eventually converge to the saddle point.

3.1.3 Mode-following algorithms and the dimer method

An inherent weakness of the MS and NEB methods is the need to specify the initial and final states between which the transition barrier is to be found. As such, one needs a preconception of what transition mechanisms exist in order to apply the methods. In many systems, intuition will not determine all possible transitions in the system. It is especially likely that co-operative transitions, *i.e.* transitions that rely on two or more atoms moving together, will be missed. We therefore need a method that can find these transitions given only the initial state of the system.

Such methods exist that work by following the modes of the system, *i.e.* the paths on the potential surface determined by the eigenvectors of the potential's Hessian matrix. By climbing 'up the hill' in various directions from the initial state, saddle points on the potential surface, and hence transitions in the atomistic system being studied, can be located without having to know the final states of the transitions in advance. These mode-following algorithms [46, 47] have existed for many years. However, a disadvantage of the traditional mode-following methods is that they rely on the ability to evaluate the Hessian matrix, typically by taking an analytical second derivative of the potential function, and then to calculate its inverse. For atomistic systems of significant size, the evaluation and inversion of the Hessian matrix are very computationally intensive operations.

The dimer method [10] is a modern example of a mode-following algorithm. Like the older methods, it makes no assumptions about the transitions that may occur from

the initial state. The difference is that the dimer method achieves greater efficiency by eliminating the need to evaluate the Hessian matrix. Instead, it uses only the first derivative of the potential energy function.

The method involves the manipulation of a 'dimer', which consists of two nearby points on the potential surface, representing two slightly different configurations of the atoms. These points remain a small constant distance apart throughout the process. First the dimer is rotated to a line of lowest curvature of the potential surface, since any saddle point will lie on such a line. The dimer is then translated 'up the hill' towards the saddle point, and the pattern of rotations and translations is repeated until a saddle is found. By starting the method at a number of different points around the potential energy minimum, and using different initial orientations of the dimer, different transitions can be found.

The dimer method has previously been used to find transitions of an Al adatom on an Al(100) surface. In this work, it is applied to Au adatom diffusion on Au(100) in Chapter 4 and to the diffusion of defects in bulk Fe and Fe–P in Chapter 5.

3.2 The standard algorithm

This section explains the dimer method as described by Henkelman [10]. Section 3.3 describes ways in which the algorithm has been varied and adapted in this work.

The following algorithm parameters are defined:

ΔR	Dimer half-axis length
$\delta\theta$	Small angular increment used for rotation
Δt	Timestep used in translation steps
T_r	Rotational force tolerance
T_t	Translational force tolerance

The state of the dimer search is determined by two vectors in $3N$-dimensional space. These are $\mathbf{R}$, the position vector of the midpoint of the dimer on the potential surface, and $\hat{\mathbf{N}}$, a unit vector that determines the dimer axis. Several variables are calculated

from this information. Of these, the main ones are $\mathbf{R}_1$ and $\mathbf{R}_2$, the position vectors of the two points that constitute the dimer, and $\mathbf{F}_1$ and $\mathbf{F}_2$, the vectors of the forces acting upon the system in the configurations represented by $\mathbf{R}_1$ and $\mathbf{R}_2$ respectively.

3.2.1 Initialisation

A dimer search is always started near a local minimum in the potential surface. This is a point of stable equilibrium in the field, which can be found by a damped classical dynamics simulation or any suitable optimisation algorithm [48]. The dimer vector $\hat{\mathbf{N}}$ is a random unit vector, which is typically generated by assigning to each coordinate a random number conforming to a uniform or normal distribution, and then scaling the whole vector so that it is of unit length. If the system has fixed atoms, the components of the dimer vector corresponding to these must be set to zero. Sections 3.3.2 and 3.3.3 give more information on fixing atoms and generating the initial dimer vector.

The endpoints of the dimer are set by

$$\mathbf{R}_1 = \mathbf{R} + \Delta R \hat{\mathbf{N}} \tag{3.6}$$

$$\mathbf{R}_2 = \mathbf{R} - \Delta R \hat{\mathbf{N}} \tag{3.7}$$

and this relation between $\mathbf{R}_1$, $\mathbf{R}_2$, $\mathbf{R}$ and $\hat{\mathbf{N}}$ is preserved throughout the process.

3.2.2 Rotation

If $\mathbf{R}$ is fixed, then the dimer will be on a lowest curvature mode when its energy is at a minimum. This is achieved by rotating the dimer about $\mathbf{R}$ so that the rotational component $\mathbf{F}^{\perp}$ of the force acting upon the dimer, defined by

$$\mathbf{F}^{\perp} = \mathbf{F} - (\mathbf{F} \cdot \hat{\mathbf{N}})\,\hat{\mathbf{N}} \tag{3.8}$$

tends towards zero, where

$$\mathbf{F} = \mathbf{F}_1 - \mathbf{F}_2. \tag{3.9}$$

Figure 3.1 illustrates how the rotational force is derived. A plane in which to rotate is

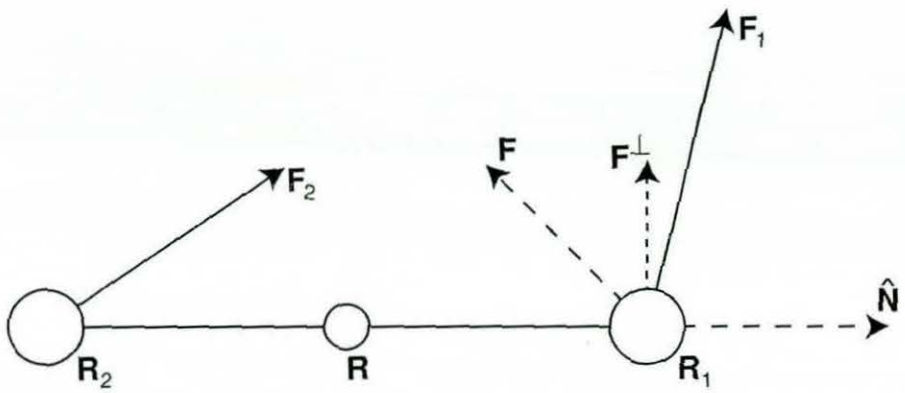


Figure 3.1: Deriving the rotational force on the dimer.

chosen by either a steepest descent method or a conjugate gradient method. This is done by generating a unit vector $\hat{\Theta}$, perpendicular to $\hat{\mathbf{N}}$. The vectors $\hat{\mathbf{N}}$ and $\hat{\Theta}$ then form a basis that spans the rotation plane. In the steepest descent approach to rotation, $\hat{\Theta}$ is simply a unit vector parallel to $\mathbf{F}^{\perp}$.

A second image of the dimer (Figure 3.2) is specified by

$$\mathbf{R}_1^* = \mathbf{R} + (\hat{\mathbf{N}} \cos \delta\theta + \hat{\Theta} \sin \delta\theta)\, \Delta R \tag{3.10}$$

$$\mathbf{R}_2^* = \mathbf{R} - (\hat{\mathbf{N}} \cos \delta\theta + \hat{\Theta} \sin \delta\theta)\, \Delta R, \tag{3.11}$$

and the corresponding forces $\mathbf{F}_1^*$ and $\mathbf{F}_2^*$ and the difference $\mathbf{F}^*$ are calculated.

Calculating the amount of rotation

There are several ways in which the change in rotational force can be used to estimate the amount of rotation needed to bring the dimer to a lowest curvature mode. The most basic of these methods is a simple adaptation of Newton's method. The derivative of F, the scalar magnitude of the rotational force as the dimer rotates, can be approximated by

$$F' = \frac{\mathrm{d}F}{\mathrm{d}\theta} \approx \frac{\mathbf{F}^* \cdot \hat{\Theta}^* - \mathbf{F} \cdot \hat{\Theta}}{\delta\theta}. \tag{3.12}$$

Then the angle by which the dimer will be rotated in this step is

$$\Delta\theta = -\frac{\mathbf{F} \cdot \hat{\Theta} + \mathbf{F}^* \cdot \hat{\Theta}^*}{2F'}. \tag{3.13}$$

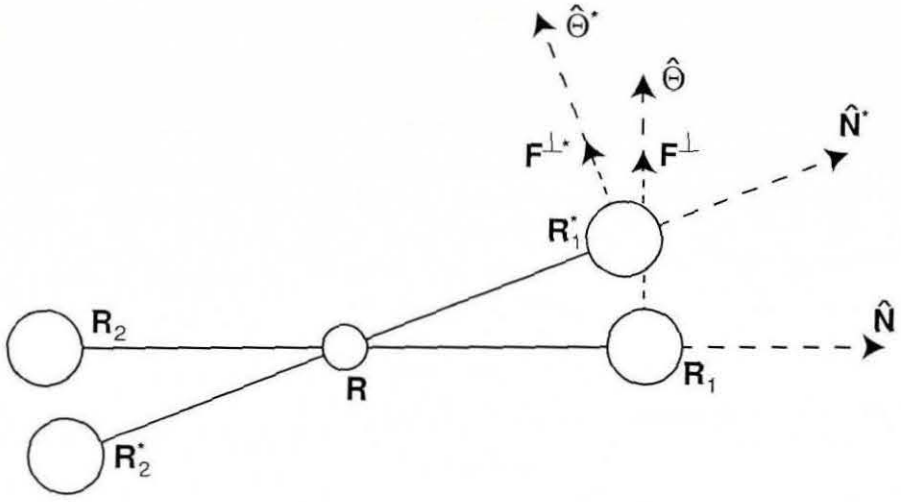


Figure 3.2: Unit basis vectors and the rotational force in initial and displaced positions.

An alternative is to use a sinusoidal fit. When the dimer is rotated through an angle θ, the rotational force acting on it can be approximated by

$$S(\theta) = A \sin 2(\theta - \alpha) \tag{3.14}$$

for some A and α. It is readily seen that $S(\theta) = 0$ when $\theta = \alpha$. The problem is thus that of determining the value of α. Differentiating equation (3.14) gives

$$S'(\theta) = 2A \cos 2(\theta - \alpha). \tag{3.15}$$

Combining equations (3.14) and (3.15) then gives

$$\frac{S(\theta)}{S'(\theta)} = \frac{1}{2} \tan 2(\theta - \alpha). \tag{3.16}$$

At $\theta = 0$, this reduces to

$$\frac{F}{F'} = \frac{1}{2} \tan(-2\alpha) \tag{3.17}$$

$$\Rightarrow \Delta\theta \approx \alpha = -\frac{1}{2} \tan^{-1}\left(\frac{2F}{F'}\right). \tag{3.18}$$

This formula turned out to be variable in its performance – the number of force evaluations was sometimes greater, and sometimes less, than that for the same system using the Newton method.

Conjugate gradient choice of rotation plane

An alternative method of choosing the rotation plane is to use a conjugate gradient approach. Instead of rotating in the plane spanned by $\hat{\mathbf{N}}$ and $\mathbf{F}^{\perp}$, the conjugate vector $\mathbf{G}^{\perp}$ is calculated, and the plane of rotation will then be spanned by $\hat{\mathbf{N}}$ and $\mathbf{G}^{\perp}$.

The formulae involved are a variation of the traditional conjugate gradient algorithm. For the first rotation, $\mathbf{G}^{\perp} = \mathbf{F}^{\perp}$ is taken. Subsequently, for the i'th rotation,

$$\mathbf{G}^{\perp[i]} = \mathbf{F}^{\perp[i]} + \gamma \, |\mathbf{G}^{\perp[i-1]}| \, \hat{\boldsymbol{\Theta}}^{**[i-1]}, \tag{3.19}$$

where

$$\hat{\boldsymbol{\Theta}}^{**[i-1]} = \hat{\boldsymbol{\Theta}}^{[i-1]} \cos \Delta\theta^{[i-1]} - \hat{\mathbf{N}}^{[i-1]} \sin \Delta\theta^{[i-1]} \tag{3.20}$$

$$\gamma = \frac{(\mathbf{F}^{\perp[i]} - \mathbf{F}^{\perp[i-1]}) \cdot \mathbf{F}^{\perp[i]}}{\mathbf{F}^{\perp[i-1]} \cdot \mathbf{F}^{\perp[i-1]}}. \tag{3.21}$$

At each step, $\hat{\boldsymbol{\Theta}}$ is defined as a unit vector parallel to $\mathbf{G}^{\perp}$ instead of $\mathbf{F}^{\perp}$. Rotation is then carried out within the plane spanned by $\hat{\mathbf{N}}$ and $\hat{\boldsymbol{\Theta}}$.

3.2.3 Translation

When the dimer's energy has been minimised by rotation, a translation step is carried out. To do this, the resultant force $\mathbf{F}_R = \mathbf{F}_1 + \mathbf{F}_2$ is calculated. The component of this force along the dimer axis is $\mathbf{F}^{\parallel} = (\mathbf{F}_R \cdot \hat{\mathbf{N}})\,\hat{\mathbf{N}}$. An effective force is defined by

$$\mathbf{F}^{\dagger} = \begin{cases} -\mathbf{F}^{\parallel} & \text{if} \quad C \geq 0 \\ \mathbf{F}_R - 2\mathbf{F}^{\parallel} & \text{if} \quad C < 0 \end{cases} \tag{3.22}$$

where $C = (\mathbf{F}_2 - \mathbf{F}_1) \cdot \hat{\mathbf{N}}$ is a measure of curvature along the dimer axis. C is positive when this cross-section is convex downwards, as when near a minimum in the potential. In this case, the effective force is defined so that it helps the dimer to quickly leave the potential energy basin (Figure 3.3). When the dimer is nearer the saddle point, C becomes negative and then the reflected force is used to converge the dimer towards the saddle point (Figure 3.4).

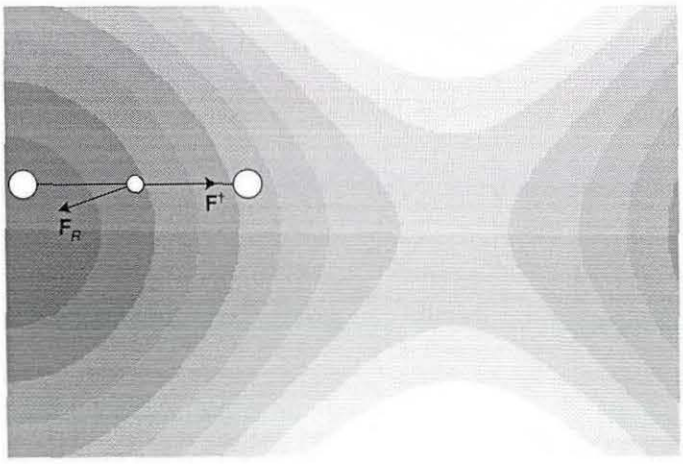


Figure 3.3: Effective force along dimer axis for $C \geq 0$.

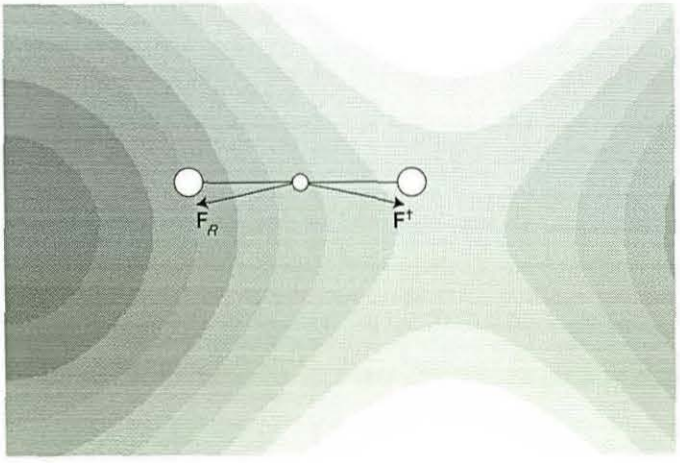


Figure 3.4: Effective force as reflected resultant force for $C < 0$.

This effective force is used to move the dimer 'up the hill' by a simple modification of a classical dynamics simulation, known as the 'quick-min' method. The movement is governed by these equations:

$$\Delta\mathbf{V}^{[i]} = \frac{\mathbf{F}^{\dagger[i]}\Delta t}{m} \tag{3.23}$$

$$\Delta\mathbf{R}^{[i]} = \left(\mathbf{V}^{[i-1]} + \frac{\Delta\mathbf{V}^{[i]}}{2}\right)\Delta t \tag{3.24}$$

$$\mathbf{V}^{[i]} = \Delta\mathbf{V}^{[i]}\left(1 + \frac{\Delta\mathbf{V}^{[i]} \cdot \mathbf{V}^{[i-1]}}{\Delta\mathbf{V}^{[i]} \cdot \Delta\mathbf{V}^{[i]}}\right). \tag{3.25}$$

The symbol m represents mass, but can be considered a mere scaling parameter. This project uses the value $m = 2$, effectively taking the unit of mass to be that of a single image of the system being considered.

The formula for $\mathbf{V}^{[i]}$ streamlines the velocity so that only the component of the velocity parallel to the current effective force is remembered between steps.

If $\mathbf{V}^{[i-1]} \cdot \mathbf{F}^{\dagger[i]} < 0$, the dimer is considered to have overshot the saddle point. In this case, the velocity at the start of the step is reset to zero, by taking the step using the alternative equations

$$\Delta\mathbf{R}^{[i]} = \frac{\Delta\mathbf{V}^{[i]}\Delta t}{2} \tag{3.26}$$

$$\mathbf{V}^{[i]} = \Delta\mathbf{V}^{[i]}. \tag{3.27}$$

3.3 Adaptations

In this project, a number of adaptations to the dimer method were implemented. There are two basic kinds of adaptations involved. The first is to set up the dimer searches in a way that is suited to the problems being studied. The second kind of adaptation, described from section 3.3.5 onwards, consists of modifications to the algorithm of the dimer method itself to make it more efficient.

3.3.1 Generating the initial configuration

The potential energy basin in which a dimer search is started represents the state from which transitions are to be found. Throughout this work, simple defect structures have been constructed and used for this. In the work of Chapter 4, this defect structure is an adatom on a surface. In the work of Chapters 5 and 6, defect structures that were used as starting configurations include an isolated vacancy, a vacancy next to a substitutional impurity atom and an interstitial atom forming a dumbbell with another atom in the lattice.

Since defects often increase or decrease the number of atoms locally, the number of atoms in a system containing an isolated defect generally differs slightly from the number of atoms in a perfect lattice of the same material and volume. This in turn affects the dimensionality of the potential surface. Let N_P be the number of atoms in a perfect lattice of the desired size. To construct the initial configuration, first a perfect lattice is constructed, and then a defect is added. The defect adds an atom, giving $N = N_P + 1$, in the case of an adatom (Chapter 4) or interstitial defect (sections 5.3.3 and 5.3.4). On the other hand, a vacancy defect, such as those of sections 5.3.1 and 5.3.2, subtracts an atom giving $N = N_P - 1$ atoms in the system.

It is necessary that the system is near a potential energy minimum before a search is carried out. To ensure this, after constructing the lattice and placing a defect in it, the system is relaxed to a stable equilibrium by using a conjugate gradient minimisation method.

3.3.2 Fixed and free atoms

The number of degrees of freedom in the system can be reduced by considering only a small subset of the atoms to be allowed to move. This has been achieved by defining a region of free atoms around the defect from which transitions are to be found. For the Au surface work of Chapter 4, a spherical region was used. For the work in Chapter 5 involving Fe and Fe–P bulk systems, a cubic region was used. All atoms outside the

defined region are fixed. In the calculations of the force acting upon the system, the force components on all fixed atoms were taken to be zero. Consequently the components of the dimer vector corresponding to fixed atoms remain zero at all times, and the position coordinates of these atoms does not change during the steps of the dimer method.

The number of free atoms required to give good results varies from system to system, and there is no known systematic way of determining the optimum number for any system. In this work it was found sufficient to use fewer than 100 free atoms for the surface system. To study bulk defects, a much larger number of free atoms was required; generally between 200 and 500 free atoms gave good results once the periodic relaxation of fixed atoms as described in section 3.3.4 was implemented.

3.3.3 Generating the initial dimer vector and displacement from the minimum

A region around the defect was also used to generate the initial dimer vector $\hat{\mathbf{N}}$. The atoms in this region are a subset of the free atoms. Figure 3.5 illustrates the relative size of this region and the region of free atoms. To generate the dimer vector, each component corresponding to an atom in the region was given a random value in the interval $[-1, 1]$, and all other components were set to zero. The vector was then normalised to be a unit vector.

In many applications, the dimer is initialised with a random small displacement from the potential energy minimum, independently of the dimer vector. For this work, the dimer was instead initialised after generating the initial dimer vector, so that $\mathbf{R}_2$ is the potential energy minimum and $\mathbf{R} = \mathbf{R}_2 + \Delta R \hat{\mathbf{N}}$.

3.3.4 Relaxation of fixed atoms

While performing the dimer search steps, only the atoms in a cubic region defined around the adatom or bulk defect were allowed to move. Increasing the size of this region can give

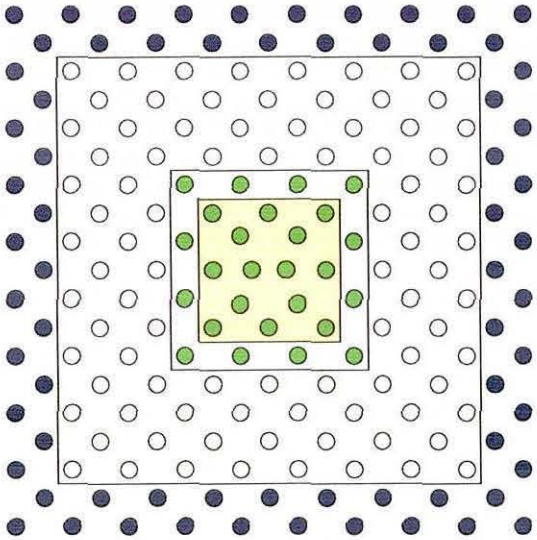


Figure 3.5: Schematic illustration of the arrangement of fixed atoms (white and blue), free atoms (green) and atoms included in the initial dimer vector (yellow background). Also shown is the atoms around the outside (blue) that are fixed even during the periodic conjugate gradient relaxation.

more accurately converged energy barriers, but making it too big was found to reduce the success rate of the dimer searches.

In order to achieve more accurate energy barriers when most of the atoms in the system are fixed, the fixed atoms can be relaxed periodically. This is done by using a conjugate gradient algorithm to minimise the potential energy of the system in only the degrees of freedom corresponding to the fixed atoms. Effectively the status of fixed and free atoms is reversed during the conjugate gradient relaxation, as opposed to the rotation and translation steps of the dimer method. However, the atoms around the outside of the cell (blue in Figure 3.5) are still fixed, preventing translation of the lattice as a whole during the relaxation process. Hence only the atoms that are white in the diagram move while relaxing the system.

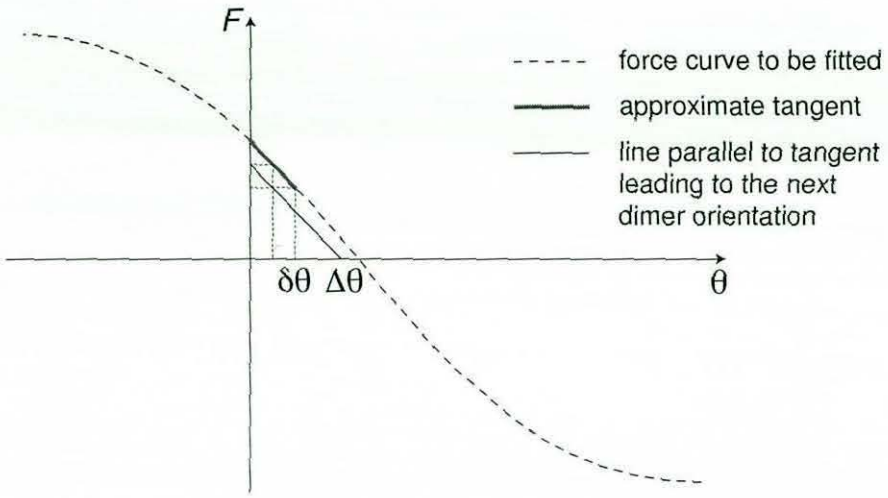


Figure 3.6: The angle of rotation of the dimer as calculated by the forward difference and forward average approximations of equations (3.12) and (3.13).

3.3.5 Alternative formulae for the angle of rotation

Some variations of the rotation formula have been implemented. The first of these is a small adjustment to the value produced by the Newton-based method given in section 3.2.2.

Equation (3.12) given in that section is a forward difference approximation of the rotational force derivative. Similarly, equation (3.13) gives a forward average. As such, the value of $\Delta\theta$ obtained from the equations is an estimate of the amount of rotation needed not from the initial orientation, but from an orientation displaced by $\delta\theta/2$ from it. In some systems, the inaccuracy of equation (3.12) can stop the dimer from properly converging to within the set tolerance (Figure 3.6). In such cases, a closer covergence can be achieved by adding this term to the rotation (Figure 3.7), giving the formula

$$\Delta\theta = -\frac{\mathbf{F}\cdot\hat{\Theta} + \mathbf{F}^*\cdot\hat{\Theta}^*}{2F'} + \frac{\delta\theta}{2}. \tag{3.28}$$

However, there is another limitation inherent in the Newton method of rotation. When the dimer's orientation is near a crest or trough of the sine wave, the derivative of the rotational force with respect to angle will be small, resulting in over-rotation of the dimer

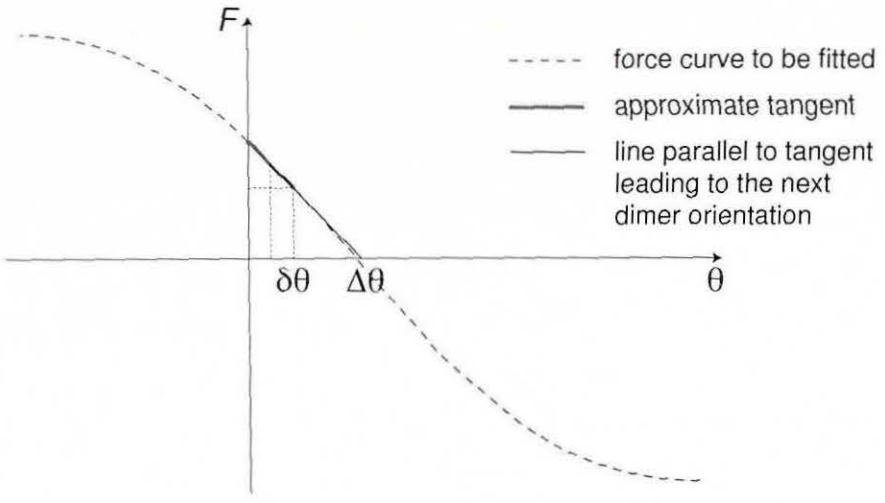


Figure 3.7: Adjustment of the angle of rotation by $\delta\theta/2$ to produce a closer convergence.

(Figure 3.8). This can be countered by constraining the angle of rotation to a maximum at each step.

An alternative sinusoidal fit to the one of section 3.2.2 can be constructed using F and F^* directly, producing the result shown in Figure 3.9. The equations are

$$A\sin(-2\alpha) = F \tag{3.29}$$

$$A\sin 2(\delta\theta - \alpha) = F^*. \tag{3.30}$$

Combining these two equations and solving for α gives

$$\Delta\theta = \alpha = \cos^{-1}\sqrt{\frac{1}{2}\left(1 + \frac{F\cos 2\delta\theta - F^*}{\sqrt{F^2 + F^{*2} - 2FF^*\cos 2\delta\theta}}\right)}. \tag{3.31}$$

In some cases, this formula gave similar evaluation counts to equation (3.18), and in other cases it was found to be an improvement. Compared with the Newton-based formulae, equation (3.31) consistently improved the efficiency of the dimer method.

3.3.6 Order of steps and stopping criteria

The original paper [10] suggests alternating translation and rotation steps, and early experiments in this project used this pattern after a number of initial rotations. However,

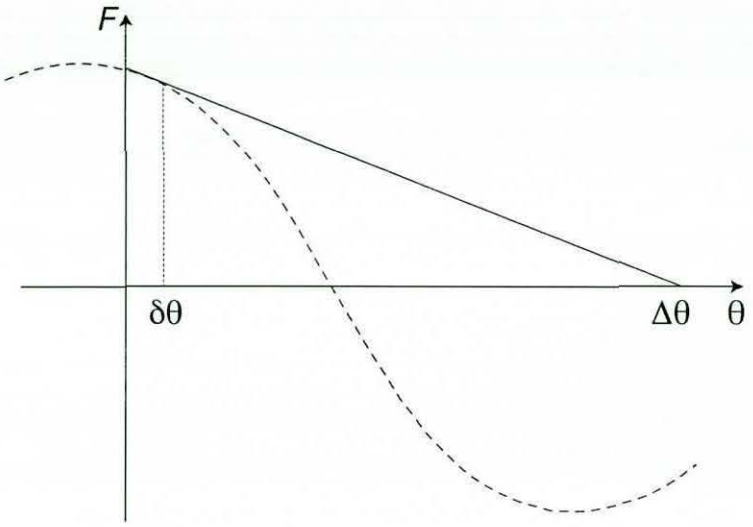


Figure 3.8: Over-rotation occurring with Newton's method.

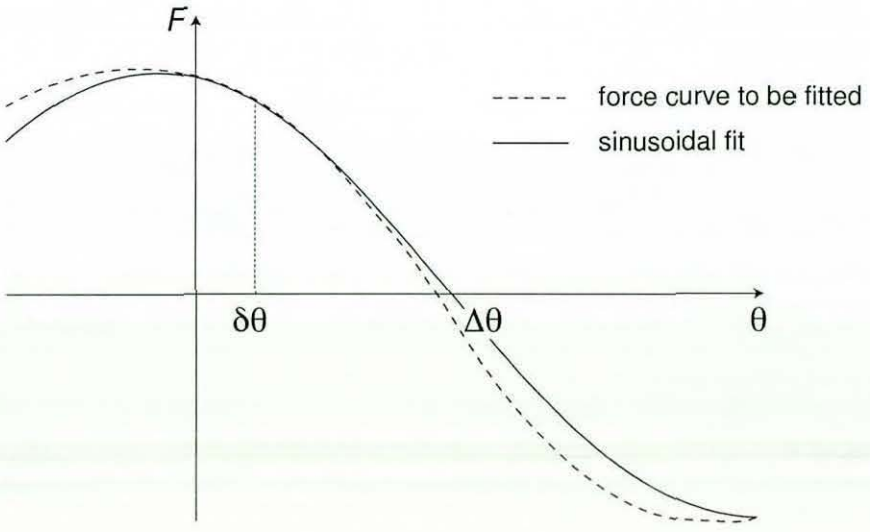


Figure 3.9: Constructing a sinusoidal fit on the rotational force curve.

this was found to be wasteful, since there is often no need to rotate the dimer between translation steps.

The pattern that has been adopted for this work is to take as many rotation steps as necessary to converge the dimer into the lowest curvature mode. The current implementation of the algorithm takes rotation steps until $F/\Delta R < T_r$, at which point a translation is done. After each translation, the rotational force is checked again to see if more rotations are necessary.

Further, it has turned out to be useful to break the rotation into two parts: choosing the rotation plane and actually carrying out the rotation. This means that, after choosing $\hat{\Theta}$ and hence the plane of rotation, we can carry out as many steps of rotation as desired in this plane, rotating $\hat{\Theta}$ with the dimer, before rotating in a different plane. This is similar to the iterative line optimisation often used as part of various optimisation techniques, where straight lines within the objective function's domain are chosen in turn and an optimum value is found along each line before choosing another.

The algorithm terminates when a translation step is taken with $C < 0$ and $||\mathbf{F}^{\dagger}||_{\infty} < T_t$.

3.3.7 Parallelisation

The dimer method can be readily adapted to use parallel processing in order to speed up the search on large atomistic systems. This is because the computational cost of the method derives from the evaluations of the force function and the vector operations, each of which can be parallelised.

To parallelise the vector operations, the components of every vector are divided between the processors. Only four vector operations need to be considered: addition, multiplication by a scalar, dot multiplication and calculating the infinity norm. The first two of these are strictly elementwise, and so each processor simply operates on its own vector components without communicating with the other processors. To calculate the dot product, each processor calculates the dot product of its own sub-vectors and then a

global sum is taken. Similarly, the infinity norm is calculated by the global maximum of the processors' local maxima.

To decide which components are assigned to which processors, the spatial decomposition method is used as described in section 2.2.3, so that each processor's vector components are those corresponding to the x, y and z coordinates of the atoms in its cell. For simplicity, atoms are not moved between processors. We can get away with this because in the course of a dimer search, the atoms move only small distances. In essence it is an atom decomposition rather than a spatial decomposition, but the computational benefit of spatial decomposition is retained. For this to work correctly, a maximum distance that any atom may move from its initial position is defined, and the skin thickness is configured to be at least equal to this distance.

3.4 Benchmarks

To test the dimer method, three sample potentials were used. These are referred to as 'egg-box' potentials because of the shape of the potential surface. The first two potentials are two-dimensional, meaning that all dimer rotations occur in a single plane. For the third example a three-dimensional potential was used, making it possible to compare the results of the steepest descent and conjugate gradient methods of rotation.

3.4.1 Basic egg-box potential

For the first test system, a simple potential (Figure 3.10) with straight transition paths was chosen. This means that rotations will only be needed at the beginining of the search, before translations are carried out. The potential function is

$$E = \sin x + \sin y. \tag{3.32}$$

Dimer searches in this potential were carried out using several different parameter sets. In each set of parameters, 50 searches were performed, using the same list of 50 randomly generated initial dimer vectors each time.

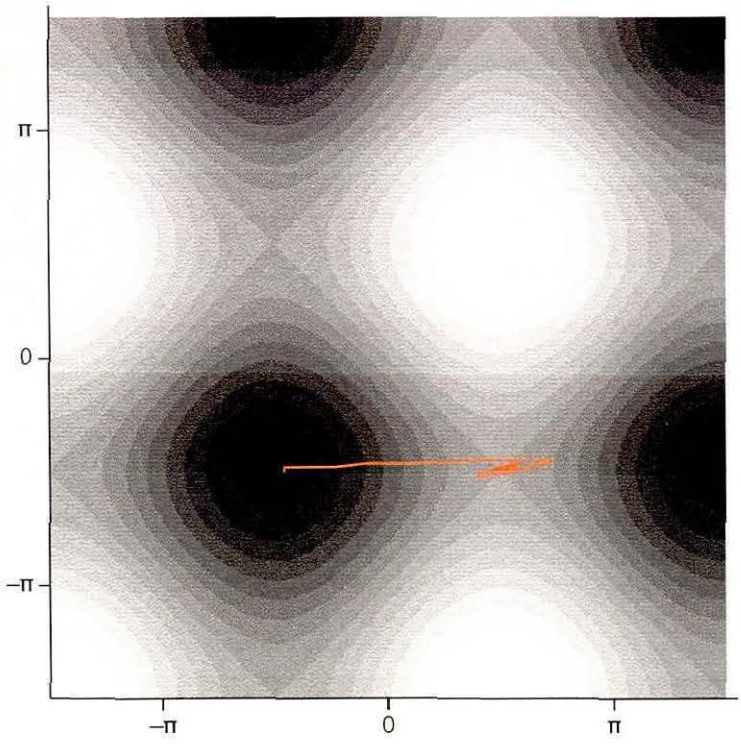


Figure 3.10: Contour plot of the basic egg-box potential. Light areas are positive and dark areas are negative. A sample path of the centre of the dimer in this potential is shown.

The best parameter set found for this system was $\Delta R = 0.1$, $\delta\theta = 0.01$, $\Delta t = 0.5$, $T_r = 0.02$, $T_t = 0.01$.

The value of Δt was found to make a significant difference to the number of evaluations of the force function required for convergence to the saddle point. Evaluation counts for three values of Δt, using the above values for the other parameters, are compared in Table 3.1.

The third formula is the variation of equation (3.28) where if $\Delta\theta > 0.5$, the value 0.5 is used instead. Variations in parameters other than Δt were found to make comparatively little difference to the efficiency of the dimer search in this potential, except that when equation (3.13) was used and $\delta\theta/T_r$ was too large, the dimer search often failed to settle to a saddle point.

Formula	$\Delta t = 0.2$	$\Delta t = 0.5$	$\Delta t = 1$
3.13	85.6	46.6	75.7
3.28	86.6	47.5	74.7
3.28 clipped	82.6	42.1	73.7
3.18	77.7	37.6	67.0
3.31	78.3	38.6	67.7

Table 3.1: Mean force evaluation counts on the simple egg-box potential for different values of Δt.

The relative performances of the rotation formulae varied according to the parameters that were used, but in this instance the sinusoidal fits, equations (3.18) and (3.31), always gave the best results. The average number of rotation steps per search varied from 2.5 in the case of equation (3.28) to 1.5 for equation (3.18); the mean number of translations was between 16 and 17 in all cases.

3.4.2 Distorted egg-box potential

In the last potential, the lowest curvature modes were all straight lines in the potential field, so there were not many rotation steps to be done. In real-world systems, however, the lowest-energy transition paths are not straight. As such, the dimer will need to rotate not only at the beginning, but also during the process of climbing up the line of lowest curvature towards the saddle point. To include this factor in the benchmarks, a modified version of the potential of section 3.4.1 is used, in which the transition paths are curved. This distorted egg-box potential is defined by the formula

$$E = \sin\left(x - \frac{y^2}{40}\right) + \sin\left(y - \frac{x^2}{20}\right). \tag{3.33}$$

Dimer searches in this potential took slightly longer. As Figure 3.11 shows, the path of the dimer towards the saddle point is often not a smooth curve. The best parameters for this search were $\Delta R = 0.1$, $\delta\theta = 0.2$, $\Delta t = 0.5$, $T_r = 0.5$, $T_t = 0.01$.

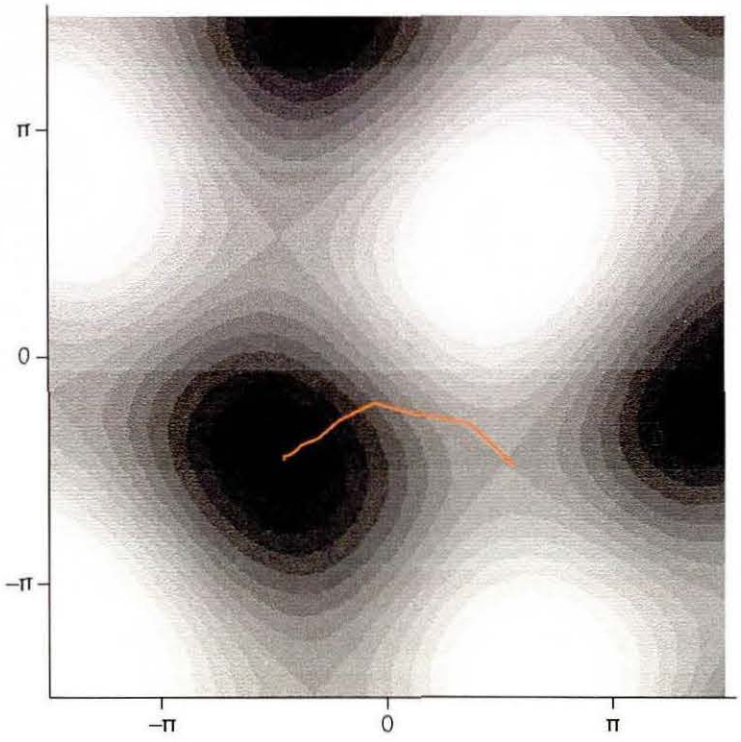


Figure 3.11: Contour plot of the distorted egg-box potential. Light areas are positive and dark areas are negative. A sample path of the centre of the dimer in this potential is shown.

Here, the relative performances of the rotation formulae, given in Table 3.2, were as expected. Smaller values of T_r and Δt were found to produce greater evaluation counts. In this system there was more variation in both the mean rotation counts (ranging from 1.3 to 2.7) and the mean translation counts (from 26 to 31). This shows that choosing an efficient rotation formula can not only minimise the number of rotation steps required; it can also orient the dimer more accurately, leading to more efficient, and hence fewer, translation steps.

Formula	Evaluations
3.13	72.9
3.28	64.2
3.28 clipped	63.0
3.18	60.9
3.31	57.0

Table 3.2: Mean force evaluation counts on the distorted egg-box potential.

3.4.3 Three-dimensional potential

The third test system is a potential similar to the second, but in three dimensions. The potential function is

$$E = \sin\left(x - \frac{y^2}{40}\right) + \sin\left(y - \frac{z^2}{20}\right) + \sin\left(z - \frac{x^2}{10}\right). \tag{3.34}$$

In this instance, the best set of parameters to be found was

$$\begin{aligned} \Delta R &= 0.1 \\ \delta\theta &= 0.05 \\ \Delta t &= 0.5 \\ T_r &= 0.5 \\ T_t &= 0.01. \end{aligned}$$

Table 3.3 compares the numbers of force evaluations per search when different rotation formulae are used, with the steepest descent method used to choose the rotation plane. Table 3.4 gives the corresponding evaluation counts using the conjugate gradient choice of rotation plane. However, as the tables illustrate the conjugate gradient method of choosing a rotation plane was found to be slower than the steepest descent method. With the two-dimensional egg-box potentials, and when steepest descent rotation on this potential, the dimer converged to a saddle point in every search. When conjugate gradient rotation was used, in many searches the dimer failed to settle within 1000 translation

Formula	Evaluations
3.13	89.4
3.28	89.9
3.28 clipped	78.1
3.18	77.7
3.31	74.4

Table 3.3: Mean force evaluation counts on the three-dimensional distorted egg-box potential using steepest descent rotation.

Formula	Success rate	Evaluations
3.13	84%	198.6
3.28	82%	246.0
3.28 clipped	84%	201.8
3.18	90%	115.4
3.31	84%	124.8

Table 3.4: Mean force evaluation counts on the three-dimensional distorted egg-box potential using conjugate gradient rotation.

steps, and even in the cases that did settle, more evaluations of the force function were carried out in the process.

3.4.4 Effect of relaxation of fixed atoms

To test the periodic relaxation of the fixed atoms as described in section 3.3.4, searches have been performed starting from a defect structure in α-Fe consisting of a substitutional P atom separated by a [110] vector (a third neighbour separation) from a [101] Fe–Fe dumbbell, using the Ackland Fe–P interatomic potential. More about the Fe–P defect systems is given in Chapter 5. Energy barriers were compared for the transition to a $[1\bar{1}0]$ dumbbell at a site with a $\frac{1}{2}[311]$ displacement from the P atom (a fourth neighbour

separation), using different numbers of free atoms with and without the relaxation of the fixed atoms. As expected, Table 3.5 shows that when no relaxation is performed, the calculated energy barrier decreases as the number of free atoms is increased. By relaxing the fixed atoms every 20 steps, a lower energy barrier is calculated, and there is negligible variation between the figures calculated with different numbers of free atoms.

Free atoms	No relaxation	Relaxed only at end	Relaxed every 20 translations
181	0.3819	0.3449	0.3346
331	0.3488	0.3400	0.3346
547	0.3401	0.3371	0.3346

Table 3.5: Comparison of the energy barriers calculated for the third–fourth neighbour transition of Figure 5.14 (e) with different numbers of free atoms, with and without periodic relaxation of the fixed atoms.

Also compared is the energy barrier obtained by relaxing the fixed atoms only in the saddle point configuration, not periodically during the dimer search. The results show that this final relaxation by itself lowers the energy barrier, but for best results the fixed atoms should be relaxed during the course of the search.

3.5 Scaling with system size

Evaluating and inverting the Hessian matrix is typically an $O(N^3)$ time complexity operation. As such, this operation is the driving factor in the older mode-following algorithms. It follows that for very large systems, the mode-following algorithms relying on Hessian inversion will take a long time to perform a single step.

Because the dimer method uses only first derivatives, the driving factor is the evaluation of the force function, which is essentially an $O(N)$ operation. The time complexity of a dimer search is therefore a product of two factors: the time taken to evaluate the force function and the number of force evaluations before a saddle point is found.

The number of degrees of freedom in the system influences the number of rotation steps that are likely to be required. This is because as more degrees of freedom are added, the net rotational force on the dimer needs to be converged to zero in more planes. This increases the number of force evaluations required to align the dimer along the lowest curvature mode at the beginning of the search. Experiments [10] showed that there is, in practice, relatively little difference between evaluation counts on the same system as more atoms are allowed to move, and observations in this work have agreed with this. Overall, it appears that the time complexity of the dimer method relative to system size is $O(N^2)$, with a relatively small N^2 coefficient.

3.6 Conclusions

In this chapter, the dimer method and a number of variations of it have been described, implemented and tested. In doing so, it has been possible to compare the performance of the algorithm variations that have been devised.

Five different formulae have been implementd for the rotation of the dimer within a plane. Although tests on various systems did not show either of the rotation formulae in the original paper as being more efficient than the other, a rotation formula has been developed that consistently outperforms the orignal Newton-based formula, and has therefore been used throughout the work of the following chapters.

The work of this chapter has shown that the dimer method works well on simple, low-dimensional potential surfaces. However, it can be challenging to find the optimum set of parameters to find saddle points in a given system. In the following chapters, the method is applied to atomistic systems, which have much greater dimensionalities, and it will be seen that the dimer method can be used to find transitions in these systems.

Chapter 4

Transitions of gold adatoms on a gold surface

Several problems involving the surface of a crystal have been studied using the accelerated dynamics techniques described in section 2.3. These include island ripening and crystal growth of Al atoms on an Al(100) surface [9], the mobility of MgO dimers [49] and small Pd clusters [50] on the MgO(100) surface, and vacancy clusters in Cu(100) [51]. Such problems are interesting because of the rich variety of transitions that occur on surfaces, many of which involve groups of atoms moving together in ways that would not be predicted intuitively.

In this chapter, the system of an *adatom*, *i.e.* an atom placed upon a surface, is studied. Such a system is simple and yet exhibits a considerable variety of transitions. These include simple hops of the adatom between surface sites, concerted exchanges of the adatom with surface atoms, and other mechanisms in which the adatom itself does not move but influences the movement of nearby atoms. For this work, an Au adatom on an Au(100) surface has been chosen. This particular system was chosen for two reasons. First, it is a system for which no previous studies in this manner have been published, and a good test of our modified dimer method on a real system. Second, the choice of a single-element system of an adatom on an fcc (100) surface makes the study very similar to one carried out previously on aluminium [10], meaning that we have results with which

we can compare those of this study and contrast the different transitions between two materials of the same fcc structure.

To model the interactions between Au atoms, the Ackland potential as described in section 2.1.2 has been used. The ZBL potential and a spline function were also used as described in sections 2.1.7 and 2.1.8 to model close-range interactions wherever they occur during the searches.

Au has a face-centred cubic lattice structure, with a lattice parameter of 4.078 Å. Successive (100) layers are therefore 2.039 Å apart. At equilibrium, the surface layer is 2.004 Å above the next layer down.

Adatoms are generally found at three distinct sites on the surface of a material. These are the on-top, bridge and hollow sites shown in Figure 4.1. Of these, only the hollow site was found to be a stable position of an Au adatom on the Au(100) surface. Therefore, configurations in which adatoms are placed at hollow sites were used for the initial states. When the system is relaxed, the adatom assumes a position 2.04 Å above the surface layer.

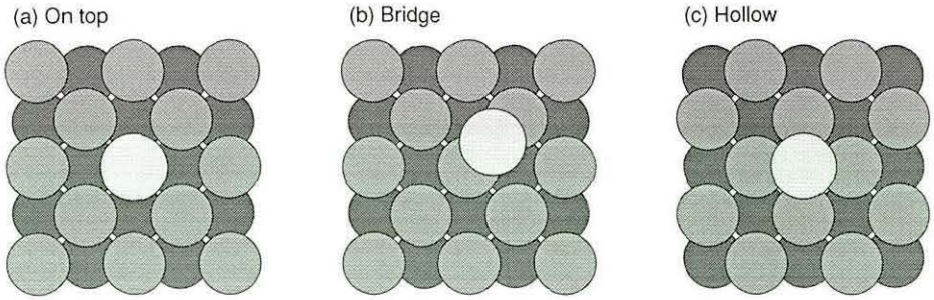


Figure 4.1: The three typical sites at which adatoms occur, shown on an fcc lattice. In all diagrams in this chapter, the atoms are shaded by depth, with the lightest shade used for the adatom layer.

Lattices of around 500 atoms were used for the dimer searches. This enabled ample numbers of free atoms to be used, while leaving enough fixed atoms so that the free atoms do not interact with the boundaries of the lattice. Unlike a defect in bulk material as described in Chapter 5, the strain field created by a adatom is very small, so it was not necessary to increase the size of the lattice significantly to accommodate this.

4.1 Single adatom

The first set of searches was carried out with all atoms fixed except for those within a 5 Å radius of the adatom, giving 18 free atoms, and using initial dimer vectors involving only the adatom and its four immediate neighbours. These searches yielded the eight transitions shown in Figure 4.2. It is immediately apparent that the lowest-energy barriers are those for which only the adatom changes site (a, b). Four of the transitions (c, d, f, g) lead to states in which an atom from the surface layer becomes a second adatom, and a vacancy is left on the surface. Transition (h) is the only transition found in the single adatom system that leads to a configuration with three adatoms and two surface vacancies.

It is already apparent that the lowest-energy transitions differ from those previously found for fcc Al [10]. The lowest energy barrier found in the Au system is that by which the adatom simply hops to an adjacent hollow site in a $\langle 110 \rangle$ direction (a). In Al, this transition has the second lowest barrier, 0.37 eV. The lowest energy barrier of the Al transitions is 0.23 eV for the two-atom concerted displacement, which has a considerably higher barrier of 1.58 eV in Au (d).

As more atoms are included in the initial dimer vector and the number of free atoms is increased, more transitions were found. Figure 4.3 shows those transitions that were found with the same set of 18 free atoms but with all free atoms included in the initial dimer vector. This led to the discovery of more transitions involving three or more atoms or atoms further from the adatom. They include three instances of a three-atom concerted displacement (b, e, h), of which (e) is another low-energy transition (barrier 0.43 eV) in Al. Transition (f) is unusual in that the atoms that move do not finish in lattice sites. This is another transition that has a low energy barrier (0.44 eV) in Al. Mechanisms such as (f) are very difficult to include in traditional KMC simulations, because the methods typically used to build the event table assume that all atoms will be at lattice sites at the beginning and end of every transition. On-the-fly KMC, on the other hand, can more readily be implemented to find transitions without making this assumption [9].

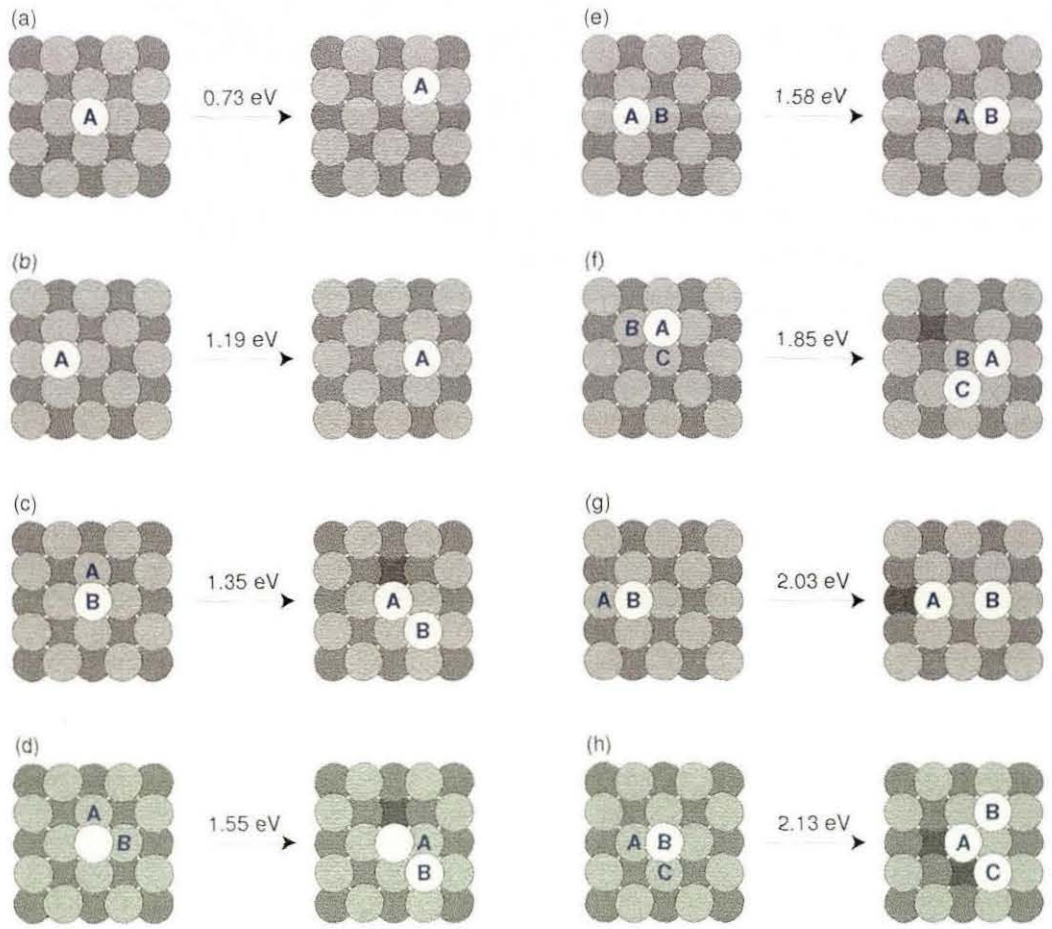


Figure 4.2: Transitions found with 18 free atoms, of which five atoms are included in the initial dimer vector. In each diagram, the atoms that move between sites have been labelled so that the initial and final positions of each atom can readily be seen.

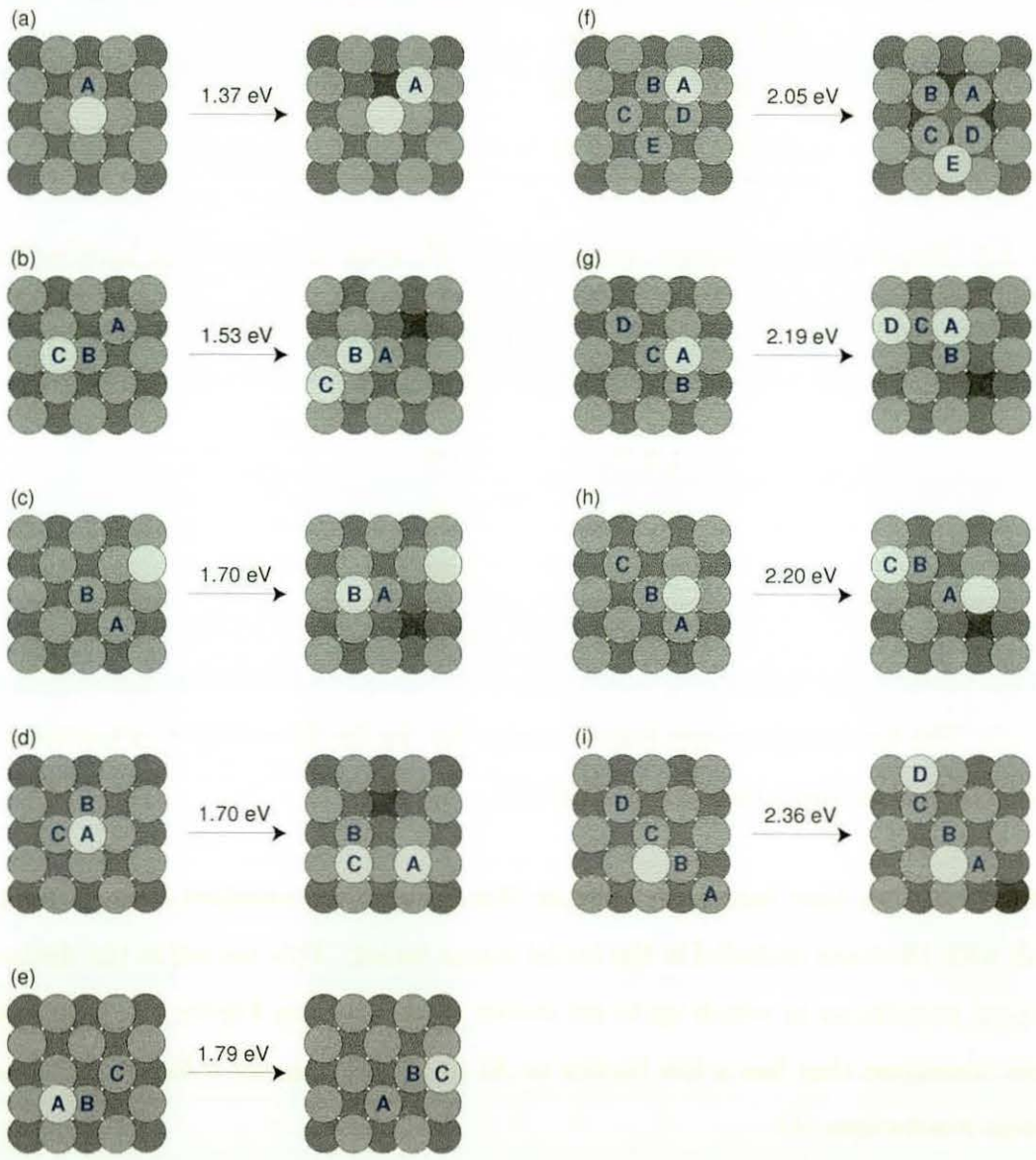


Figure 4.3: Transitions found with 18 free atoms, all of which are included in the initial dimer vector.

Figure	Al barrier (eV)	Au barrier (eV)
4.2 (e)	0.23	1.58
4.2 (a)	0.37	0.73
4.4 (e)	0.41	1.94
4.3 (e)	0.43	1.79
4.3 (f)	0.44	2.05

Table 4.1: The five lowest-energy transitions in Al, with corresponding barriers for Au.

Figure	Au barrier (eV)
4.2 (a)	0.73
4.2 (b)	1.19
4.2 (c)	1.35
4.3 (a)	1.37
4.3 (b)	1.53

Table 4.2: The five lowest-energy transitions in Au. In the Al work [10], a barrier is given only for the first of these transitions, 0.37 eV.

Dimer searches have been performed on this system with numbers of free atoms up to 58, still with 18 atoms included in the initial dimer vector. This has led to the discovery of additional transitions in which up to six atoms move, given in Figure 4.4. These include another transition that has a low barrier in Al (0.41 eV), namely a four-atom concerted exchange mechanism (e).

The lowest energy barriers for the Al and Au systems are summarised in Tables 4.1 and 4.2 respectively. From these tables it can be seen that energy barriers are generally higher in Au than in Al, and that the lowest-energy transitions are very different between the two materials. In general, the lowest energy barriers found on the Au surface are of transitions in which only one or two atoms move. This is in contrast with the Al system, in which concerted displacements constitute most of the low-energy mechanisms.

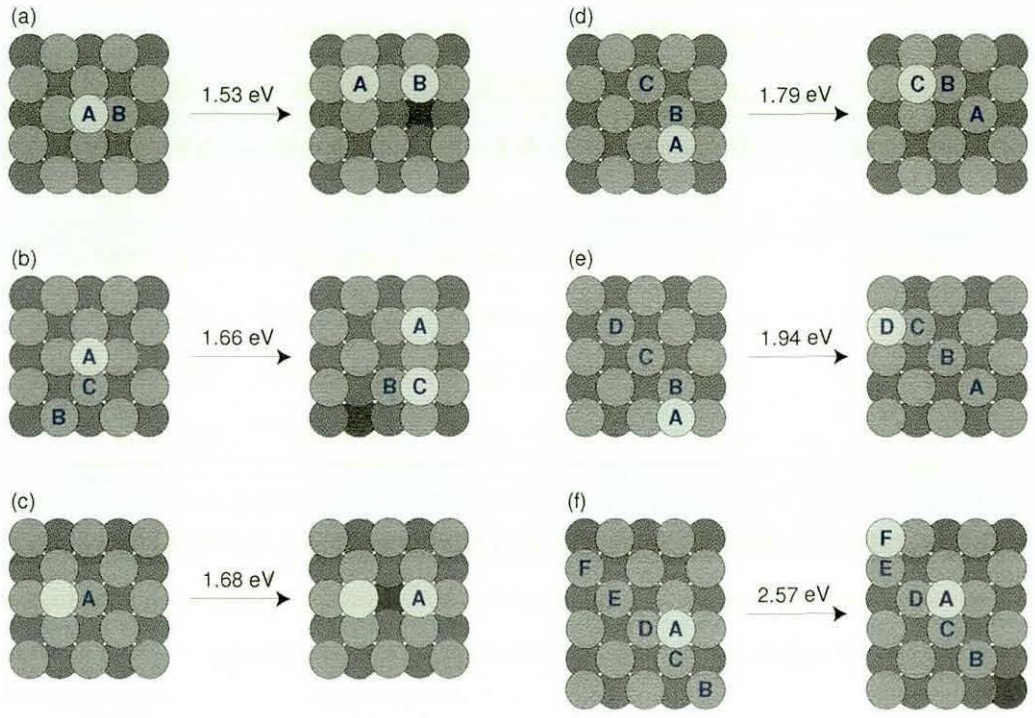


Figure 4.4: Additional transitions found with numbers of free atoms between 22 and 56.

4.1.1 Transition rates

Using harmonic transition state theory (hTST), it is possible to calculate the rate constant of each transition. This is a measure of how frequently the transition will occur, defined as the mean frequency of the transition over the time that the system spends in the transition's initial state. Chapter 6 explains more about transition rates and how they can be used to estimate the relative probabilities of transitions and the time taken for them to occur.

The rate constant of a transition from state i to state j is determined by the Arrhenius equation

$$r_{ij} = \nu \, \mathrm{e}^{-E_{ij}/k_B T} \tag{4.1}$$

where E_{ij} is the transition energy barrier, k_B is Boltzmann's constant, T is the temperature and ν is the attempt frequency. For this work, the value $\nu = 10^{13}$ s^{-1} has been used. This is a typical value assumed for rate theory modelling [15].

	Energy barrier	Rate constant (s^{-1})		
Figure	(eV)	300 K	500 K	700 K
4.2 (a)	0.73	6.1	4.7×10^{5}	5.8×10^{7}
4.2 (b)	1.19	1.0×10^{-7}	1.0×10^{1}	2.7×10^{4}
4.3 (a)	1.37	9.7×10^{-11}	1.6×10^{-1}	1.4×10^{3}
4.3 (b)	1.53	2.0×10^{-13}	3.8×10^{-3}	9.6×10^{1}
4.3 (f)	2.05	3.6×10^{-22}	2.2×10^{-8}	1.7×10^{-2}
4.3 (f)	2.57	8.1×10^{-31}	1.4×10^{-13}	3.5×10^{-6}

Table 4.3: The rate constants for a selection of energy barriers at 300 K, 500 K and 700 K. Four of the lowest energy barriers in the Au/Au(100) system, along with two higher ones, have been chosen to represent the range of barriers in this system.

At a temperature of 300 K, the lowest-energy transition has a rate constant of 6.1 s^{-1}, and the second has a rate of 1.0×10^{-7} s^{-1}. As such, it is to be expected that at this temperature, only adatom hops will occur. Table 4.3 and Figure 4.5 show that at higher temperatures, not only do the rate constants increase in order of magnitude, but the ratio between the rates of lower-energy and higher-energy transitions becomes smaller. It follows that a greater variety of transitions will occur as the temperature is increased. At 500 K, the transition of Figure 4.2 (a) is no longer the only transition that will be observed. When the temperature is increased to 700 K, the variety of transitions increases further to include several transitions in which a surface atom moves to become an adatom. The highest-energy transition found in this system is that of Figure 4.4 (f) with a barrier of 2.57 eV; even at 700 K this transition will very rarely, if ever, occur.

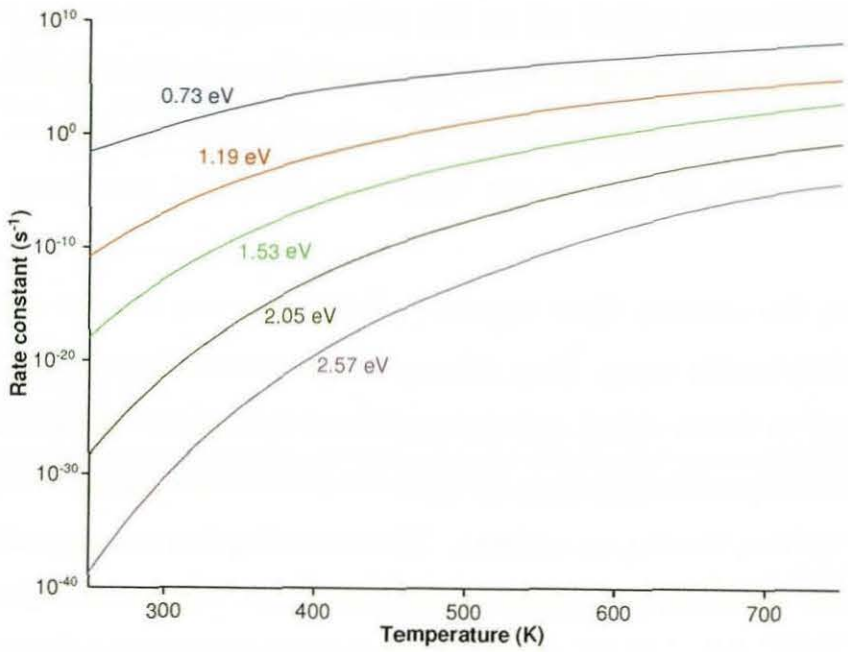


Figure 4.5: The rate constants for five energy barriers, plotted on a logarithmic scale against temperature. The ratio between rate constants becomes smaller as the temperature increases, so that higher-energy transitions become more probable.

4.2 Two adatoms

Some experiments were also carried out on a surface having two adatoms. The first of these was a configuration in which the adatoms are separated by a distance that is significantly greater than the cut-off distance of the interatomic potential. As such, it is expected that the adatoms will diffuse independently of each other for as long as they remain separated by more than the cut-off distance.

Dimer searches were carried out on this system using 36 free atoms, namely the 18 atoms within a 5 Å radius of each adatom. As expected, most of the transitions discovered from this configuration involved the atoms around only one of the adatoms, and were the same transitions with the same energy barriers as those that were found with a single adatom.

By moving the adatoms closer together, so that they can interact with each other, some interesting results occur. Four different displacements between the two adatoms have been used as initial states, and the results are described in the following sections. The first two configurations are those in which the adatoms are placed at adjacent hollow sites on the surface, forming an *addimer*. The remaining two setups used initial states of two adatoms separated by a single vacant hollow site. Figure 4.6 illustrates the four states. The [011] addimer (a) is the most energetically favourable of these four states, with an energy level 0.38 eV below the other three states (b, c, d), which are all within 0.01 eV of each other.

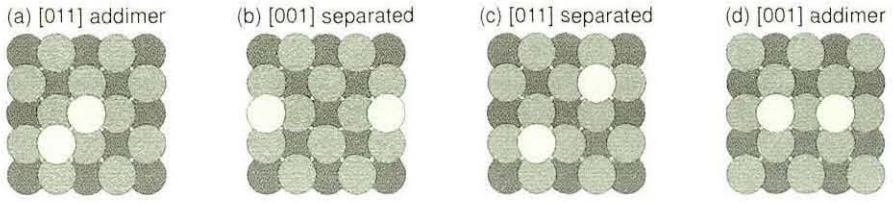


Figure 4.6: The four two-adatom initial states from which dimer searches were conducted, in ascending order of potential energy.

4.2.1 Adatoms at orthogonally adjacent sites

A number of dimer searches were performed from an initial state with two adatoms at orthogonally adjacent hollow sites, forming a [001] addimer. Figure 4.7 shows the transitions found. These are essentially the same as those found with a single adatom, but with different energy barriers. In other words, if a non-moving adatom is removed, then the transition becomes identical to one of those that was found in the single-adatom system.

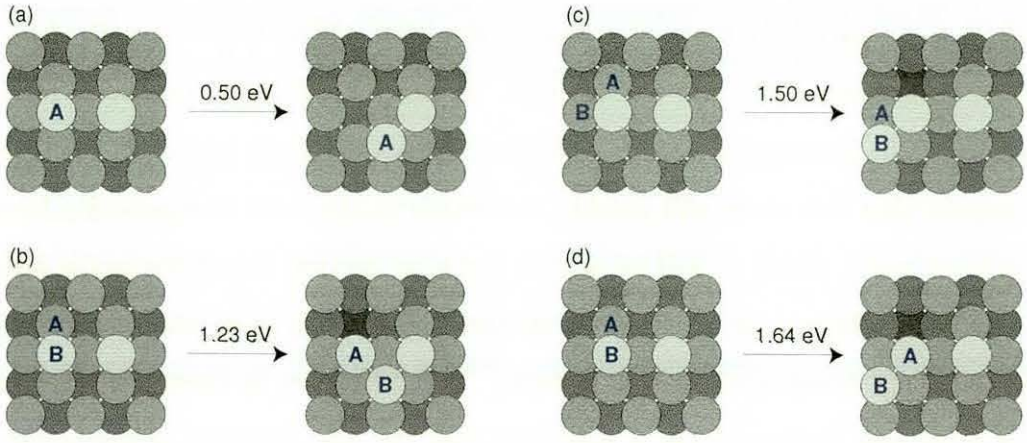


Figure 4.7: Transitions found in the [001] addimer system.

Two-adatom transition		Equivalent one-adatom transition	
In Figure 4.7	Barrier (eV)	In Figure 4.2	Barrier (eV)
(a)	0.50	(a)	0.73
(b)	1.23	(c)	1.35
(c)	1.50	(d)	1.55
(d)	1.64	(c)	1.35

Table 4.4: Comparison of energy barriers between transitions in the [001] addimer system and their equivalent one-adatom transitions given by removing a non-moving adatom.

	Energy barrier	Rate constant (s^{-1})		
In Figure 4.7	(eV)	300 K	500 K	700 K
(a)	0.50	4.1×10^{4}	9.3×10^{7}	2.6×10^{9}
(b)	0.54	9.9×10^{3}	4.0×10^{7}	1.4×10^{9}
(c)	1.23	2.1×10^{-8}	3.9	1.4×10^{4}

Table 4.5: The rate constants for the three lowest-energy transitions in the [001] addimer system at 300 K, 500 K and 700 K.

Table 4.4 compares the energy barriers in the [001] addimer system with those of the equivalent transitions with only one adatom. It can be seen that, except in the case of Figure 4.7 (e), the presence of a second adatom lowers the energy barrier of the transition. This suggests that this state will quickly transition to the more energetically favourable [011] addimer state, which as will be seen in the next section, is a more stable state than either the [001] addimer or an isolated adatom. Table 4.5 gives the rate constants of the lowest-energy transitions in this system. The difference in energy barrier between Figure 4.2 (a) and Figure 4.7 (a) is sufficient that the rate constant of the latter is four orders of magnitude higher at 300 K. The transition rate of Figure 4.7 (c) is still much lower than that of (a) and (b), so that diffusion remains almost entirely driven by adatom hops especially at lower temperatures.

4.2.2 Adatoms at diagonally adjacent sites

Searches have also been carried out starting from the [011] addimer configuration. Figure 4.8 shows the transitions that were found in this system. Two distinct adatom hops were discovered, namely in the $[01\bar{1}]$ (a) and [011] (c) directions; however, both have higher energy barriers than the 0.73 eV barrier of the isolated adatom hop. On the other hand, the two instances of a two-atom concerted exchange, one effecting a rotation of the addimer (b) and the other moving the two adatoms apart (e), both have lower barriers than the equivalent transition in the isolated adatom system (1.58 eV). Unlike with a

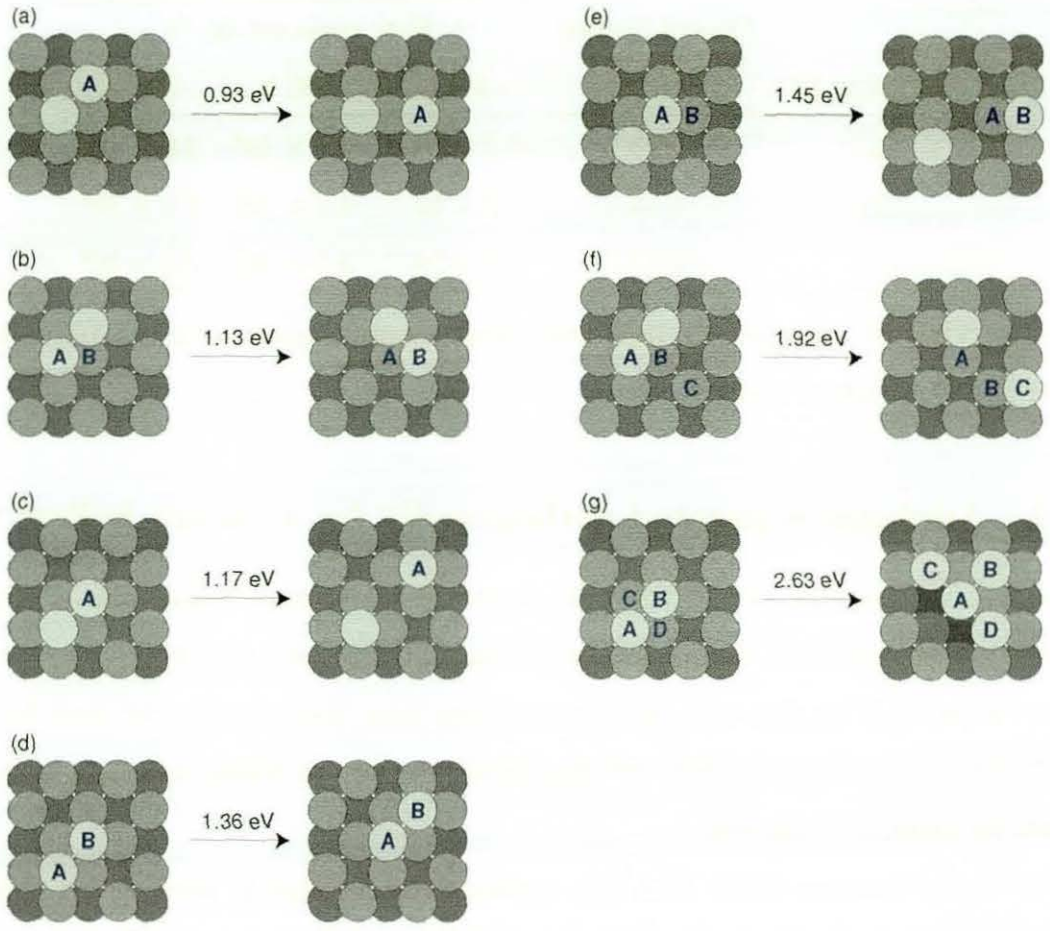


Figure 4.8: Transitions found in the [011] addimer system.

single adatom or [001] addimer, no transition leaving a surface vacancy was discovered except for (g), which is too high-energy to be expected to occur.

The transition barriers from this state and the [001] addimer state together imply that an addimer will generally diffuse by moving between these two states. As Table 4.6 shows, at a temperature of 300 K, even the lowest-energy transition from this state has a rate constant of 2.3×10^{-3}. This suggests that the [011] addimer is a more stable state than the isolated adatom and [001] addimer configurations. It follows that the addimer will diffuse more slowly than an isolated adatom, mainly by moving back and forth between the [011] and [001] orientations.

	Energy barrier	Rate constant (s^{-1})		
In Figure 4.8	(eV)	300 K	500 K	700 K
(a)	0.93	2.3×10^{-3}	4.1×10^{3}	2.0×10^{6}
(b)	1.13	1.2×10^{-6}	4.5×10^{1}	7.8×10^{4}
(c)	1.17	1.9×10^{-7}	1.5×10^{1}	3.5×10^{4}

Table 4.6: The rate constants for the three lowest-energy transitions in the [011] addimer system at 300 K, 500 K and 700 K.

4.2.3 Adatoms separated orthogonally by a vacant hollow site

The initial configuration of two adatoms separated by a single hollow site in the [001] direction was also studied. Figure 4.9 shows some of the transitions found from this state. Unlike the setup of section 4.2.1, most transitions from this arrangement were found to have higher energy barriers than the equivalent transitions where only one adatom is present, as Table 4.7 indicates.

The higher barriers imply that this configuration is slightly more stable than the isolated adatom state. However, the barrier of Figure 4.9 (b) is lower than that of the same transition for an isolated adatom, making it the second lowest barrier for this setup. The transition leads to the [001] addimer configuration, which can then be expected to lead to a sequence of [011] and [001] addimer states as was determined in sections 4.2.1 and 4.2.2.

4.2.4 Adatoms separated diagonally by a vacant hollow site

The fourth initial state from which dimer searches were performed was that of two adatoms separated by a single vacant hollow site in the [011] direction. The transitions shown in Figure 4.10 were found from this configuration.

In the most likely transition, one of the adatoms moves towards the other to form the more stable [011] addimer configuration, as shown in Figure 4.10 (a). The energy barrier for this to happen is similar to the barrier of the isolated adatom hop. Another

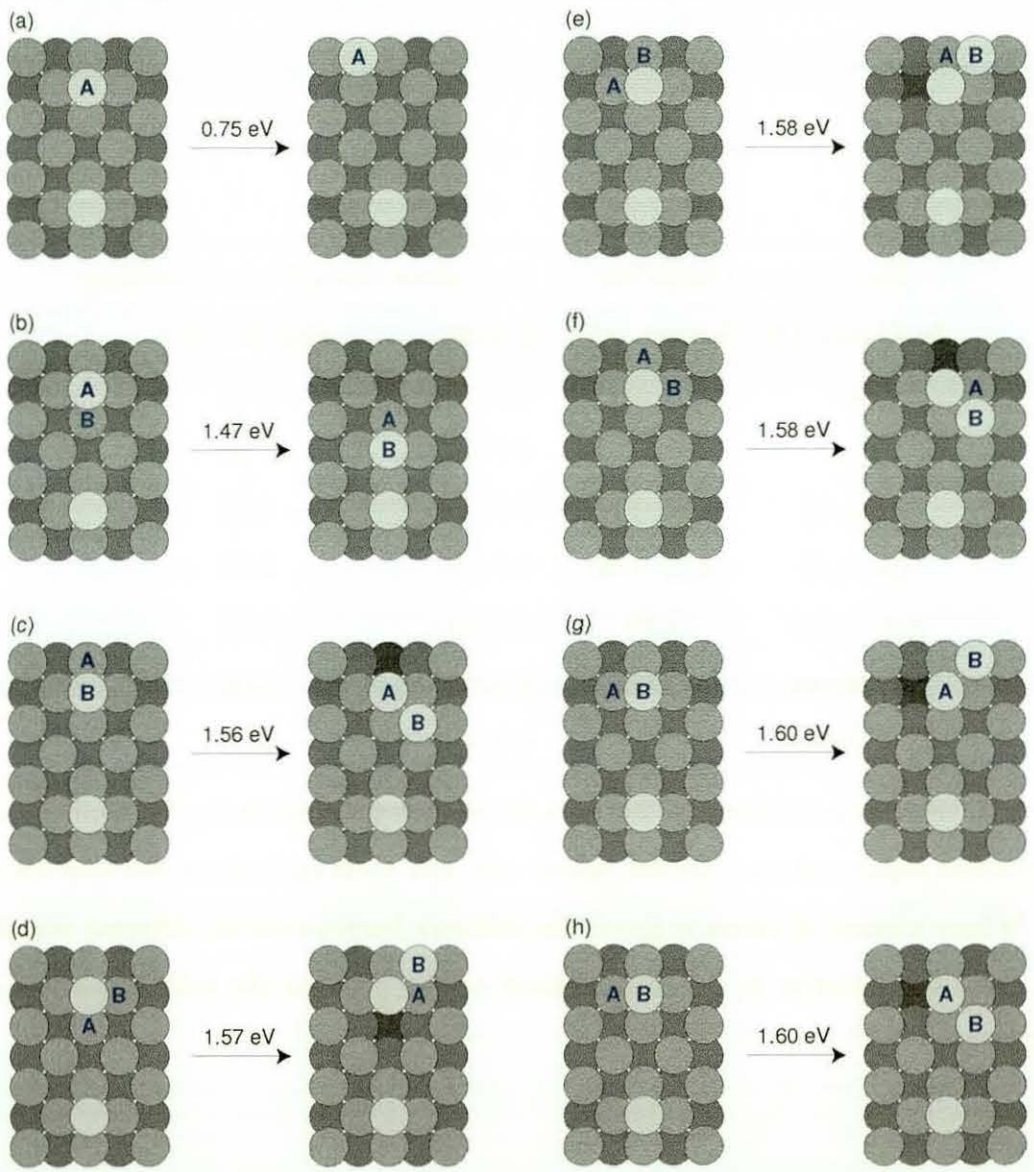


Figure 4.9: Transitions found in the system of two adatoms separated by a single hollow site.

Two-adatom transition		Equivalent one-adatom transition	
In Figure 4.9	Barrier (eV)	Figure	Barrier (eV)
(a)	0.75	4.2 (a)	0.73
(b)	1.47	4.2 (e)	1.58
(c, g, h)	1.56–1.60	4.2 (c)	1.35
(d, e, f)	1.57–1.60	4.2 (d)	1.55
not shown	1.89	4.2 (f)	1.85
not shown	2.05	4.2 (g)	2.03
not shown	2.10	4.3 (d)	1.70

Table 4.7: Comparison of energy barriers between transitions from the configuration of two adatoms separated by a vacant hollow site and their equivalent one-adatom transitions. Where a range of values is given, the different barriers are for different orientations of the transition relative to the displacement vector between the initial locations of the two adatoms.

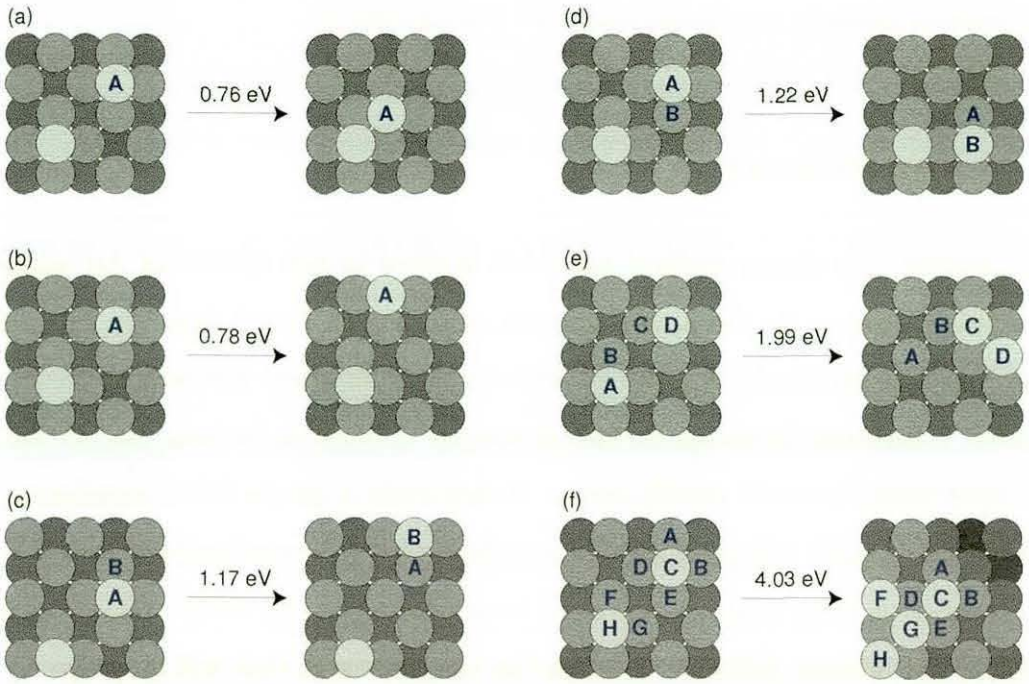


Figure 4.10: Transitions found from the initial state of two adatoms separated diagonally by a single hollow site.

adatom hop from this two-adatom state is to a configuration in which the adatoms are displaced from each other by a [013] vector (b). This also has a similar barrier, and so it can be expected to occur with almost the same probability as (a). Two instances of a two-atom concerted exchange were also observed (c, d). These are both lower-energy than the same mechanism is for an isolated adatom. Also discovered in this setup is a four-atom concerted exchange mechanism involving both adatoms (e), and an interesting seven-atom transition (f). The latter has a barrier close to the cohesive energy of Au, and as such, is unlikely ever to be observed even at 700 K.

4.3 Conclusions

In this chapter, the dimer method has been applied to the system of Au adatoms on the Au(100) surface. In the single-adatom case, it was found that the motion of an adatom will consist almost entirely of adatom hops at lower temperatures. At 400 K and below, transitions involving surface atoms are unlikely to be seen, but at 500 K and above these transitions will readily occur. Hence while a simple KMC simulation with a predefined event table will produce a reasonable result at lower temperatures, for higher temperatures it is likely that one will need to employ on-the-fly KMC techniques in order to simulate the adatom diffusion through the various states that will crop up.

The results on the one-adatom system have been compared with the previous studies of an Al adatom on Al(100). It has been seen that, although both materials have the same crystalline structure, they have very different diffusion dynamics. In Al, it was seen that concerted exchanges involving up to four atoms are among the lowest-energy transitions. In Au, on the other hand, adatom hops and other transitions involving few atoms are the most energetically favourable transitions.

Some interesting results have also been discovered for an Au(100) surface on which two adatoms are present and close enough to interact with each other. It has been seen that two-adatom configurations can be either more stable or less stable than one-adatom configurations. In particular, of the initial states studied in this chapter, the [011] addimer

is the most stable configuration to have been found. The diffusion of the addimer consists mainly of hops between the [011] and [001] addimer states. Because the hop from the [011] to the [001] addimer has a somewhat lower rate constant than the isolated adatom hop, this diffusion process occurs more slowly than the diffusion of an isolated adatom.

The low energy level of the [011] addimer, relative to the other two-adatom configurations, suggests that a property of Al, namely that compact adatom clusters are more energetically favourable than scattered adatoms [9], may also hold true for Au.

The results presented in this chapter have identified the transitions and energy barriers from initial configurations of one or two adatoms with no defects in the surface or lower layers. The methodology can be easily adapted to study transitions from other configurations, such as those involving surface vacancies or larger clusters of adatoms. This work can also lead towards more complex studies of Au surface effects, such as the simulation of island ripening and crystal growth.

Chapter 5

Transitions in Fe and Fe–P systems

The previous chapter was concerned with transitions on the surface of a material. In this chapter, the dimer method is applied to the study of defect transitions in bulk material. Systems of α-Fe with isolated point defects have been chosen for this work. These include two defect systems involving a P impurity atom.

An understanding of the diffusion of P impurities in Fe is important to the study of the durability of steel nuclear pressure vessels (NPVs). This is because steel is an alloy containing a number of elements in addition to Fe, some of which are impurities. Of these impurities, P atoms in particular are known to segregate towards the grain boundaries of the steel. When such a vessel is subjected to radiation or high temperature, the impurities become more mobile and the segregation of P atoms to grain boundaries is enhanced. This is believed to be a major cause of embrittlement in the NPVs [11, 12, 13, 14]. It is therefore of interest to understand the mechanisms behind this process.

To simulate the damage to a crystal caused by radiation, a collision cascade is typically used. This involves performing an MD simulation from an initial configuration in which one atom, known as the primary knock-on atom (PKA), has been given a significant amount of kinetic energy, typically of the keV order of magnitude [2].

The defect structures and transitions in the system of Fe containing P have also been studied using molecular statics [2]. Results from these studies include that a substitutional P atom is the most stable of the P defect structures; displacement of the P atoms

from substitutional positions has been observed only during a collisional event or by raising the temperature to near the melting point. During a collision cascade, the P atoms often form ⟨110⟩ Fe–P dumbbells (see section 5.1 for more about defect structures). Another important observed effect is that clusters of interstitial Fe atoms can aggregate around a substitutional P atom.

For this work, two different models of the interactions between the atoms involved were used. The first of these is to use the Ackland Fe potential [18] to model the Fe–Fe interactions, and the Morse potential [25] to model the Fe–P and P–P interactions; this combination will subsequently be referred to as Ackland/Morse. The second model tested is the newer Ackland potential designed for the modelling of Fe containing P impurity atoms [19]; the details of this potential have been given in section 2.1.2.

Other studies of P impurities in Fe carried out in recent years include an MD investigation of the interaction of the P atoms with grain boundaries [52], both using Ackland/Morse and using a modified EAM (MEAM) potential [53, 54]. There has also been an *ab initio* study of the transition mechanisms in this system [55].

5.1 Defect structures

The structure of α-Fe is body-centred cubic (bcc). Figure 5.1 (a) shows a perfect bcc unit cell. The atoms at the vertices of the cell adjoin the adjacent unit cells, such that there are are two atoms per unit cell in a bcc lattice. Point defects generally occur in three basic kinds: vacancy, interstitial and substitutional. Figure 5.1 shows typical structures of vacancy and interstitial defects. Other interstitial defect structures sometimes seen in bcc materials, but which were not observed in this work, include ⟨100⟩ and ⟨111⟩ dumbbell configurations.

The two kinds of defect involving a P atom that were studied here are the P–vacancy complex and the P interstitial defect. The P–vacancy complex is a combination of a vacancy and a P substitutional atom, close enough to interact with each other and treated as a single defect. The P interstitial defect is created from a locally perfect Fe lattice by

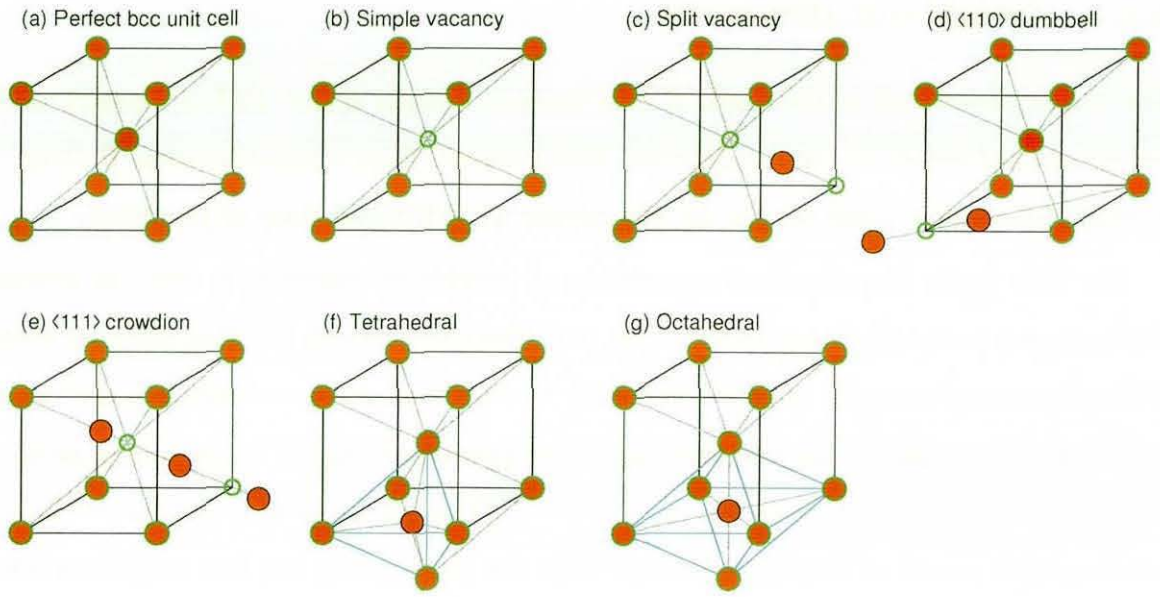


Figure 5.1: Some defect structures commonly found in bcc materials. The green circles denote lattice sites. The structures shown are a perfect lattice cell (a), two vacancy defect structures (b, c) and four interstitial defect structures (d, e, f, g). The tetrahedron (f) and octahedron (g) defining the respective interstitial positions have been outlined in blue.

adding a P atom to it; this defect usually forms a $\langle 110 \rangle$ dumbbell with an Fe atom, but as will be seen, it can also assume a tetrahedral position or become decomposed into a P substitutional atom and Fe self-interstitial defect.

In this work, the dimer searches were performed on an Fe lattice of $14 \times 14 \times 14$ bcc unit cells (giving a total of 5488 atoms) to which a defect has been added. Because we are working with bulk material, periodic boundary conditions have been used in all dimensions. When using the dimer method to find transitions in the bulk material, it turned out that more free atoms were necessary to obtain accurate energy barriers than for the surface studies of the previous chapter. Generally the searches were performed with 200 to 500 free atoms and 40 to 100 atoms in the initial dimer vector. The free atoms and the atoms in the initial dimer vector were selected by defining cubic regions around the defect that has been created in the lattice to form the initial state.

5.1.1 Neighbour distances

The states discovered in the work of this chapter include states with a vacancy or Fe self-interstitial defect at various different displacements from a substitutional P atom. It is necessary to have a convention for identifying an individual state of this kind.

For this work, the standard convention of neighbour numbers is used to identify displacement vectors between lattice sites, and hence distinguish between different states of the P–vacancy complex and of a $\langle 110 \rangle$ Fe–Fe dumbbell near a substitutional P atom. Given any lattice point, the surrounding lattice points are ranked in ascending order of their Euclidean distances from the given point. Thus the immediately adjacent lattice points, which are all at the same distance from the given point, are first neighbour sites. Those lattice points that are at the next smallest distance are the second neighbours, and so on.

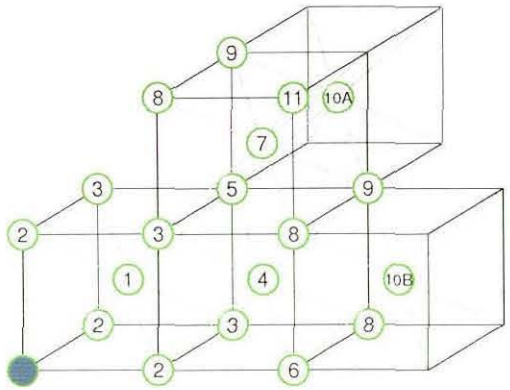


Figure 5.2: Nearest neighbour numbers corresponding to small displacements from a given lattice point, represented by a blue sphere.

Figure 5.2 gives neighbour numbers for various displacements in the bcc structure. In this structure, the first neighbours of a given lattice point are those points that have a $\frac{1}{2}\langle 111 \rangle$ displacement from that point, giving a separation distance of $\frac{\sqrt{3}}{2}a \approx 0.87a$, where a is the lattice parameter. The second and third neighbours are those with $\langle 100 \rangle$ and $\langle 110 \rangle$ displacements respectively, and distances of a and $\sqrt{2}a \approx 1.41a$. The tenth

neighbour distance, equal to $\frac{3\sqrt{3}}{2}a$, occurs for two distinct displacement vectors, $\frac{3}{2}\langle 111\rangle$ and $\frac{1}{2}\langle 511\rangle$, which are identified as 10A and 10B respectively in the figure.

5.2 Results from Ackland Fe and Morse potentials

5.2.1 Fe self-interstitial defect

Using the Ackland pure Fe potential [18], the diffusion of self-interstitial atoms in a system containing only Fe atoms was studied. Only two interstitial defect structures were found to be metastable, namely the $\langle 110\rangle$ dumbbell and the $\langle 111\rangle$ crowdion shown in Figure 5.1 (d, e).

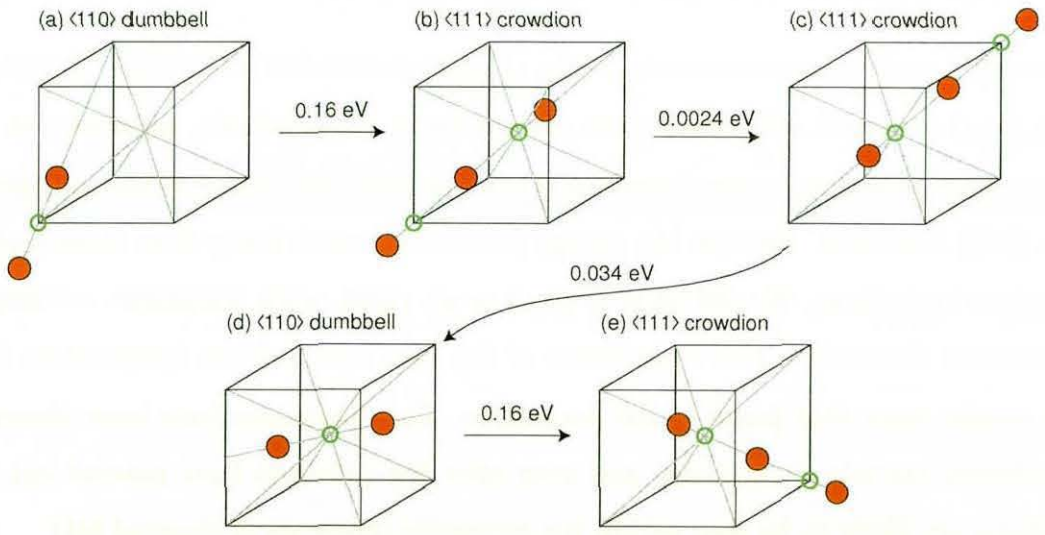


Figure 5.3: A sample sequence of transitions of an Fe self-interstitial defect.

The dimer method was used to find transitions from both of these initial states. The diffusion of the Fe self-interstitial was found to consist of transitions between the $\langle 110\rangle$ dumbbell and the $\langle 111\rangle$ crowdion states and translation of the $\langle 111\rangle$ crowdion along its axis. Figure 5.3 gives a sample sequence of transitions. The $\langle 110\rangle$ dumbbell can become a $\langle 111\rangle$ crowdion in either of two $\langle 111\rangle$ axes passing through its centre, and in two positions along each. The $\langle 111\rangle$ crowdion can become a dumbbell again at either of the lattice sites

it leaves vacant, and in three different ⟨110⟩ axes through each site. By a sequence of these moves, all ⟨110⟩ and ⟨111⟩ directions are accessible, and therefore the self-interstitial defect can move freely in the three-dimensional bulk of the material.

	Energy barrier	Rate constant (s^{-1})		
Transition	(eV)	300 K	500 K	700 K
crowdion translation	0.0024	9.1×10^{12}	9.5×10^{12}	9.6×10^{12}
crowdion to dumbbell	0.034	2.7×10^{12}	4.5×10^{12}	5.7×10^{12}
dumbbell to crowdion	0.16	2.1×10^{10}	2.4×10^{11}	7.0×10^{11}

Table 5.1: The rate constants for the transitions of an isolated Fe self-interstitial defect found using the Ackland pure Fe potential.

Table 5.1 gives the rate constants for the transitions found in this system. Translations of the ⟨111⟩ crowdion will often occur several times consecutively, especially at lower temperatures, producing a one-dimensional random walk until the crowdion changes back into a ⟨110⟩ dumbbell. Because the energy barriers are much lower than those found for Au surface transitions, the diffusion is much more rapid, with transitions occurring on a picosecond timescale. The dependence of the rate constants on temperature is also much smaller than that found in the Au system. The transitions have been observed in MD collision cascade simulations, and even after the collisions have petered out, some transitions are likely to be seen within the accessible timescale of classical MD.

5.2.2 P interstitial defect

This section describes the transitions that were found in the system of Fe with an interstitial P defect, using the Ackland/Morse model. Systems consisting of a P atom added to an Fe lattice have been shown to exhibit a variety of metastable states and transitions between them (Figure 5.4). The states that were found include two unusual dumbbell configurations, denoted by the approximate ratios of the components, along the cubic cell axes, of the displacement between the Fe and P atoms.

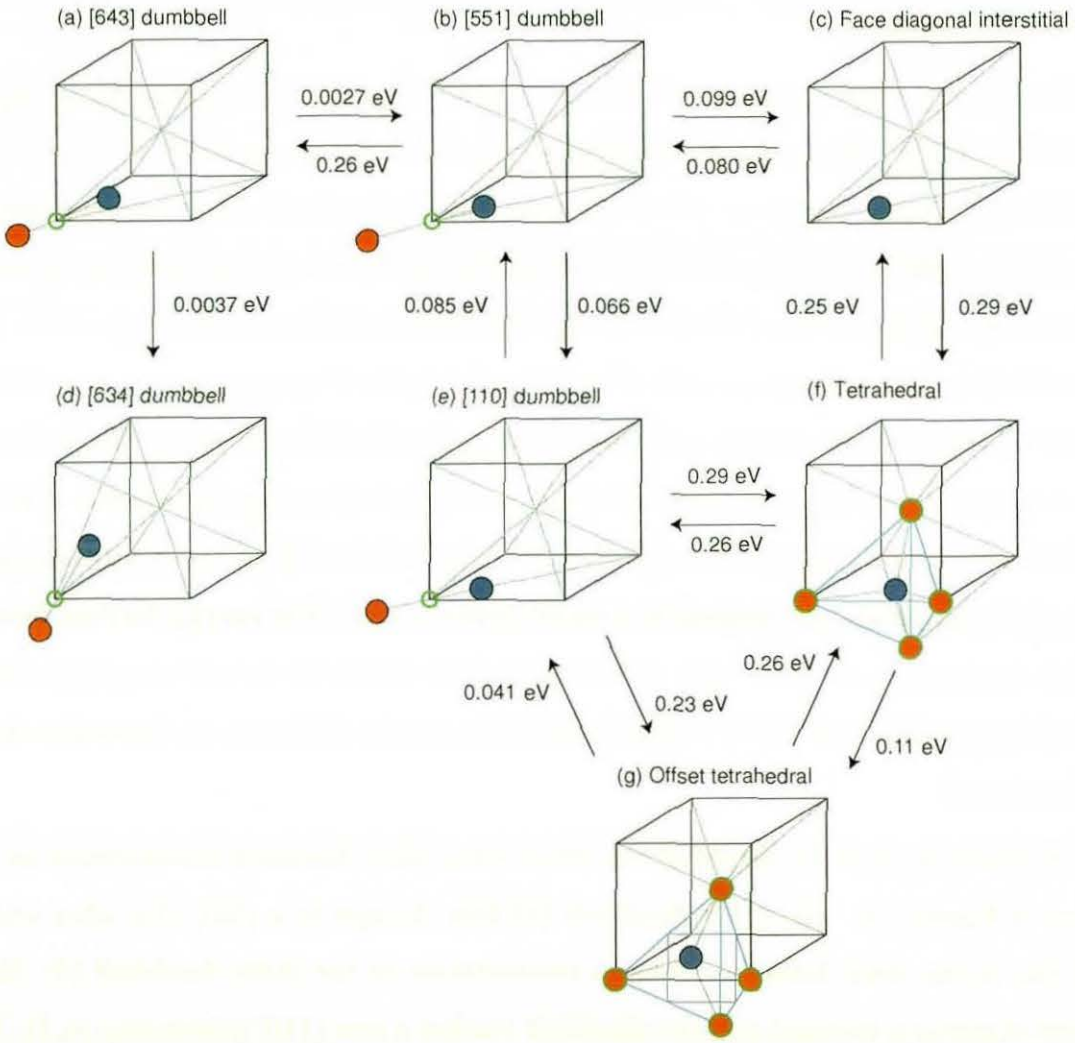


Figure 5.4: The transitions found in the P interstitial system using the Ackland/Morse model, with energy barriers.

Dimer searches were performed on this system with up to 1730 free atoms, which generally gives convergence of the energy barriers to the 10^{-3} eV level. Various sequences of transitions that enable the P atom to diffuse through the lattice have been found. With the ⟨110⟩ dumbbell position considered as a starting point, we have found two distinct diffusion mechanisms.

The simplest mechanism is a P atom in a ⟨110⟩ dumbbell configuration moving into a tetrahedral site. From the dumbbell position the P atom can move to either of two possible tetrahedral positions. From the tetrahedral site the P atom can then move to one of four possible ⟨110⟩ dumbbells, including the one from which it came. By a series of such moves, the P atom can diffuse through the lattice as illustrated in Figure 5.5. From a dumbbell position about a vertex of a unit cell (a), the P atom moves to a tetrahedral site on the face of the cell (b), and then forms a dumbbell in the centre of the face (c). It can then move to another tetrahedral position (d), before forming another dumbbell in the centre of an adjacent face (e). The P atom is thus able to move throughout the bulk of the material, not limited to a single plane or cell. The energy barriers involved in this mechanism are noticeably higher than those found for Fe self-interstitial defects, and at temperatures of 300 K to 500 K, diffusion events will occur on timescales around a nanosecond.

Alternatively, a ⟨110⟩ dumbbell can pivot via a ⟨643⟩ dumbbell configuration as illustrated in Figure 5.6. The ⟨110⟩ dumbbell (a) first changes to a ⟨551⟩ (b), after which it switches to the ⟨643⟩ form (c). After a reorientation of the ⟨643⟩ dumbbell (d), the sequence of states is reversed and the dumbbell reaches a new ⟨110⟩ orientation (e, f). From the initial ⟨110⟩ dumbbell, there are two possible ⟨551⟩ positions, each of which leads to either of two ⟨643⟩ orientations. Consequently, there are four possible orientations to which the ⟨110⟩ dumbbell can rotate by this mechanism. The main energy barrier in this process is the 0.26 eV barrier between states (b) and (c), which has a transition time of 2.1 ns at 300 K. However, from state (c) there is a moderate probability that the defect will slip back to state (b) (0.0027 eV) instead of moving on to (d) (0.0037 eV). It may also flip several times between states (c) and (d) before going either to (b), leading back

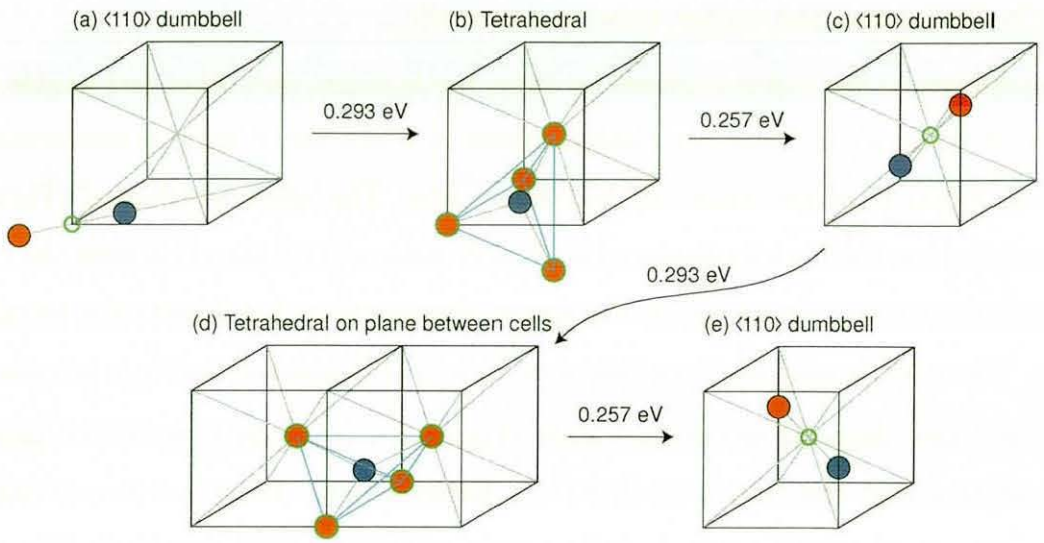


Figure 5.5: Sample sequence of transitions between ⟨110⟩ dumbbell and tetrahedral configurations.

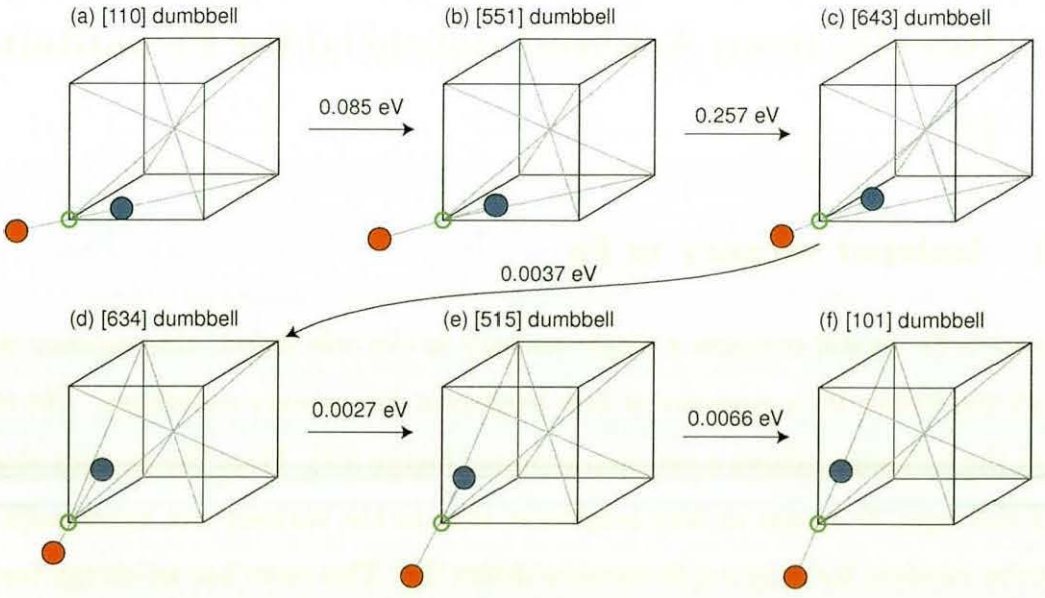


Figure 5.6: The rotation of a ⟨110⟩ dumbbell via a ⟨643⟩ dumbbell configuration.

to (a), or (e), leading to a new state. By this mechanism, the dumbbell can only rotate about its lattice site, and cannot move to a new site.

In addition to the states explored by these mechanisms, two other metastable states have been discovered. These are configurations in which the P atom is interstitial, but the surrounding Fe atoms remain at their lattice sites. The 'offset tetrahedral' (Figure 5.4 (g)) position has the P atom displaced by 0.28 Å along a ⟨110⟩ direction from the regular tetrahedral location, remaining on the square cross-section that bisects the tetrahedral region. There are transitions from this state to a ⟨110⟩ dumbbell and to a tetrahedral.

In the 'face diagonal interstitial' state (Figure 5.4 (c)), the P atom lies along the diagonal of a unit cell face, 1.33 Å from a lattice point. There are transitions from this position to the regular tetrahedral position and to the ⟨551⟩ dumbbell. The energy barriers to and from this state are similar to those involving the ⟨110⟩ dumbbell; hence it is moderately probable that the system will enter this state instead of the ⟨110⟩ dumbbell state at the end of either of the diffusion mechanisms.

5.3 Results from Ackland potential for Fe containing P

5.3.1 Isolated vacancy in Fe

When an α-Fe crystal contains a single vacancy as the sole defect, this vacancy diffuses through the lattice by a sequence of first neighbour Fe–vacancy exchanges. The mechanism of the exchange follows a two-step process (Figure 5.7). From the vacancy state (a), any of the eight Fe atoms at first neighbour sites to the vacancy can move halfway towards the vacancy, forming a split vacancy defect (b). This move has an energy barrier of 0.640 eV. The split vacancy then transitions back into a simple vacancy with an energy barrier of 0.091 eV, by the Fe atom moving either back to its previous position or into the vacancy of the previous state, leaving its own previous lattice site vacant (c).

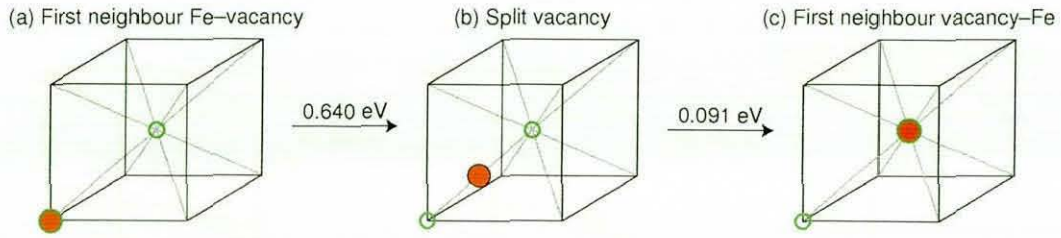


Figure 5.7: The mechanism of an isolated vacancy (green, unfilled circle) exchanging with an adjacent Fe atom (red sphere), via a split vacancy state (interstitial Fe atom shown with a black outline).

Since the vacancy can move in any of eight directions, and these directions span the three-dimensional vector space, the vacancy can freely explore the three-dimensional bulk of the material.

5.3.2 The P–vacancy complex

When a P atom and a vacancy are close to each other, the position of the P atom relative to the vacancy influences the diffusion mechanisms and associated energy barriers.

The most frequent diffusion mechanism is the exchange of the P atom and vacancy when they are at first neighbour positions (Figure 5.8). From the starting position (a), the P atom moves in the $\langle 111 \rangle$ direction towards the vacancy (b) and then to a position slightly off the line connecting the two lattice points (c). Position (b) is 0.9 Å from the lattice site, compared with first neighbour distance of 2.5 Å. In position (c), the P atom has coordinates of $(0.7, 0.7, 0.4)$ Å, giving a distance of 1.1 Å, relative to the lattice site in which it started. To complete the exchange, the system passes through the same three states in reverse order (d, e, f) with the P atom now in the area of the original vacancy. This mechanism does not by itself enable the P atom to diffuse through the lattice.

Other transition mechanisms enable the vacancy to move by exchanging with Fe atoms. Most of these mechanisms involve a split vacancy state similar to that found in the diffusion mechanism of an isolated vacancy. A frequently occurring example of this is

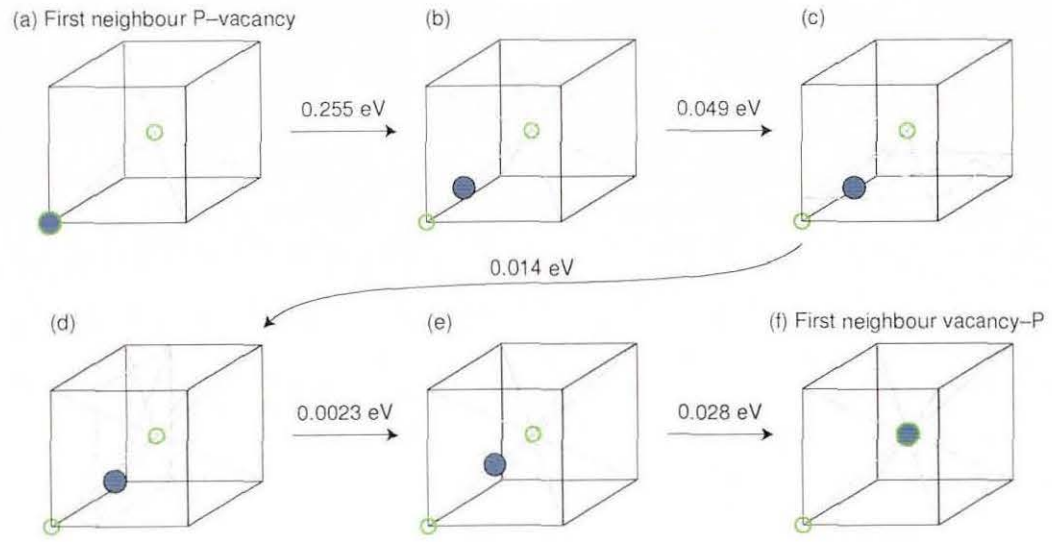


Figure 5.8: The mechanism of a P atom (blue sphere) exchanging with a vacancy (unfilled circle) when they are in first neighbour positions.

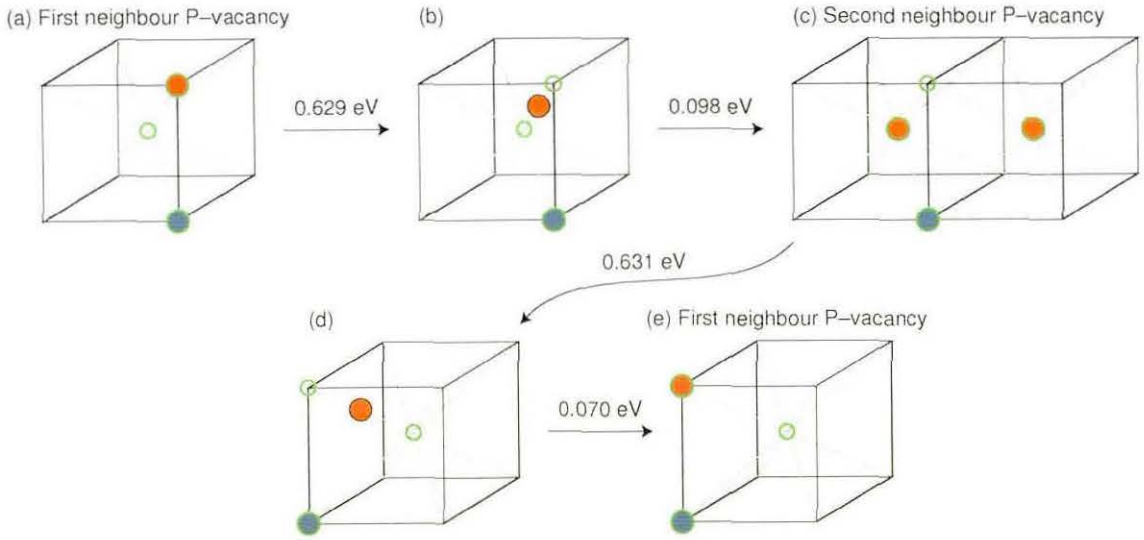


Figure 5.9: The mechanism of the first neighbour P–vacancy reorientation via the second neighbour state, with energy barriers. Fe atoms (red spheres) move to exchange with the vacancy (unfilled circle), while the P atom (blue sphere) does not move.

shown in Figure 5.9. The defect transitions from the first neighbour P–vacancy state (a) through the split vacancy state (b) so that the vacancy becomes a second neighbour of the P atom (c). There are three possible second neighbour sites at which the vacancy can become situated from a given first neighbour state. From the second neighbour P–vacancy state, the vacancy moves by a similar process (d, e) to any of four first neighbour sites, including the one from which it came. This mechanism enables the vacancy to circle around the P atom. By a sequence of these moves combined with the P–vacancy exchange mechanism, the P atom can diffuse through the lattice.

State	Energy level (eV)
1st	0.000
2nd	0.026
3rd	0.433
4th	0.385
5th	0.371
6th	0.360
7th	0.368
8th	0.364
9th	0.366

Table 5.2: Potential energy levels of states with the vacancy separated at different distances from the P atom in the P–vacancy system. The first neighbour state has been taken as a base level.

There are also transitions in which the vacancy moves further from the P atom. The energy levels of simple vacancy states out to the ninth neighbour position are given in Table 5.2, and the energy barriers of the transitions between these states are given in Table 5.3. Since all of the states are within 0.5 eV of each other and the energy barriers do not exceed 0.65 eV, it should be expected that the system explores a wide variety of states. Figure 5.10 shows the potential energy profiles of a selection of transition paths in

		Energy barriers (eV)	
Initial state	Final state	Initial → intermediate	Intermediate → final
1st	2nd	0.629	0.098
1st	3rd	1.065	
1st	5th	0.582	0.388
2nd	1st	0.631	0.070
2nd	4th	0.592	0.165
3rd	1st	0.632	
3rd	4th	0.622	0.076
3rd	7th	0.634	0.068
4th	2nd	0.307	0.090
4th	3rd	0.646	0.099
4th	5th	0.615	0.093
4th	6th	0.639	0.077
4th	8th	0.599	0.102
4th	9th	0.639	0.055
5th	1st	0.528	0.071
5th	4th	0.637	0.086
5th	7th	0.632	0.094
6th	4th	0.647	0.094
7th	3rd	0.677	0.090
7th	5th	0.640	0.088
7th	8th	0.635	0.089
7th	9th	0.640	0.088
8th	4th	0.654	0.070
8th	7th	0.636	0.091
9th	4th	0.624	0.090
9th	7th	0.640	0.091

Table 5.3: Transition barriers found in the P–vacancy system. The transition between the first and third neighbour states was not found to have an intermediate state. Hence for this case only one energy barrier is given in each direction.

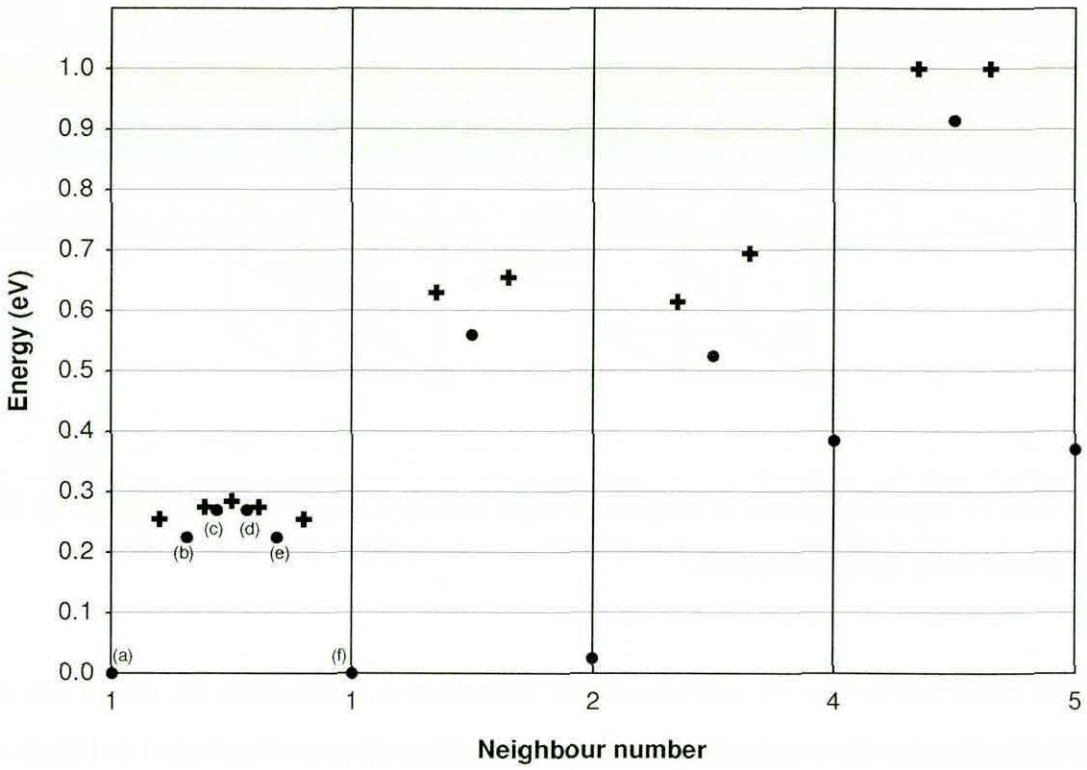


Figure 5.10: The energy profiles of the P–vacancy exchange mechanism and three Fe–vacancy exchange mechanisms between different neighbour sites of the P atom. Only the metastable states (●) and the saddle points (✚) are plotted. The letters on the profile of the P–vacancy exchange refer to the stages of the exchange illustrated in Figure 5.8.

this system. Beyond the eighth neighbour position, the energy barriers show very little deviation from those calculated for an isolated vacancy. This shows that the diffusion of the vacancy at these greater distances from the P atom can be approximated by that of an isolated vacancy in Fe.

5.3.3 Isolated Fe self-interstitial

Self-interstitial defects in α-Fe usually assume a $\langle 110 \rangle$ dumbbell configuration. The preferred mechanism for diffusion of this defect is a simultaneous rotation and translation to a first neighbour site (Figure 5.11). From the [110] orientation, the dumbbell can move

to any of four first neighbour sites, those with a displacement of $[111]$, $[11\bar{1}]$, $[\bar{1}\bar{1}1]$ and $[\bar{1}\bar{1}\bar{1}]$. At each site the dumbbell can rotate to either of two orientations: if it moves by $[111]$ then the dumbbell will have orientation $[101]$ or $[011]$.

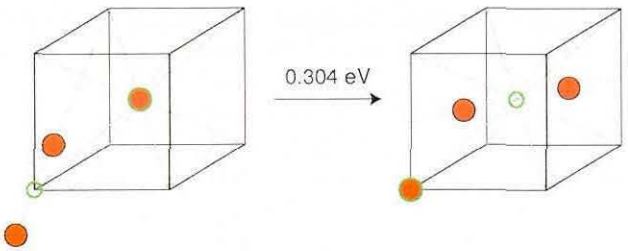


Figure 5.11: The mechanism of migration of an isolated Fe–Fe $\langle 110 \rangle$ dumbbell by simultaneous rotation and translation.

Also observed in the Fe self-interstitial system is a mechanism by which the $\langle 110 \rangle$ dumbbell migrates via a metastable octahedral configuration as illustrated in Figure 5.12. The octahedral can form at either of two sites, namely those that are adjacent along the dumbbell axis at either end. The octahedral atom can then move to form a $\langle 110 \rangle$ dumbbell at any of the four lattice sites that are nearest in a $\langle 110 \rangle$ direction.

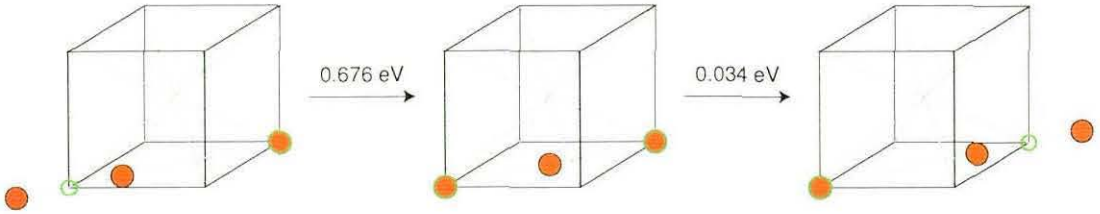


Figure 5.12: The mechanism of migration of an isolated Fe–Fe $\langle 110 \rangle$ dumbbell via an octahedral configuration.

5.3.4 P interstitial defect

When a P atom is added to an otherwise locally perfect Fe lattice, a variety of states exist. Most of the time, the defect is found in a $\langle 110 \rangle$ Fe–P dumbbell configuration.

In the most common diffusion mechanism, the Fe atom in the dumbbell returns to its lattice site and the P atom moves to either of two tetrahedral interstitial sites. From this position, the P atom can form a dumbbell with any of four Fe atoms, including the one from which it came. By a sequence of these moves as shown in Figure 5.13, P atom can explore the bulk of the Fe lattice in all three dimensions.

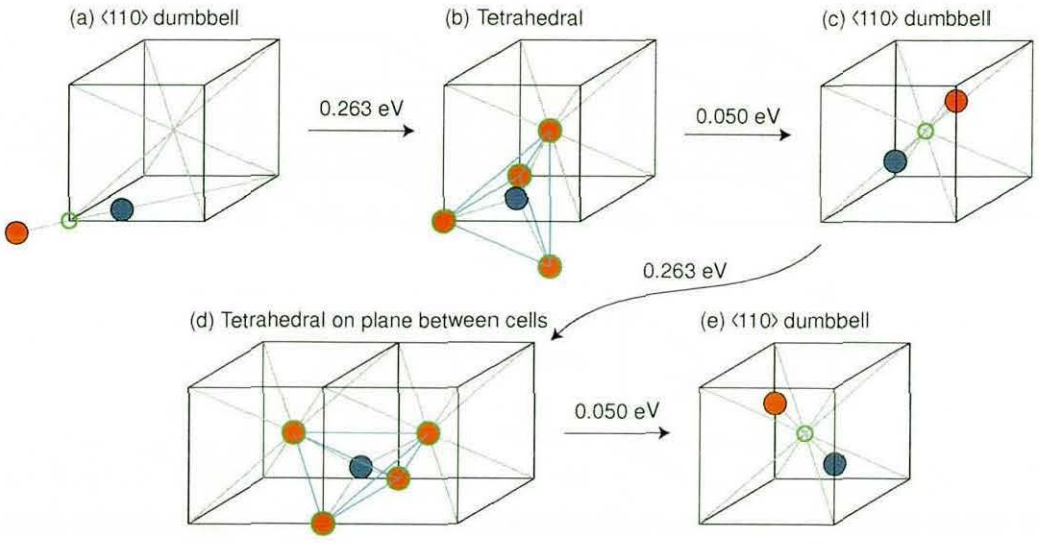


Figure 5.13: The mechanism of a P atom moving by changing between the ⟨110⟩ dumbbell and tetrahedral states.

From the ⟨110⟩ Fe–P dumbbell state, the P atom can also become substitutional, with the Fe atom from the dumbbell moving to form a ⟨110⟩ dumbbell with a first neighbour Fe atom. The resulting Fe–Fe dumbbell can then move out to lattice sites further from the P atom. The table of transitions is more complicated than that of the P–vacancy complex, since different orientations of the dumbbell relative to its displacement from the P atom are possible. Figure 5.14 shows some of the transitions that were found between dumbbell states.

Unlike in the P–vacancy system, there are relatively few states with low energy levels relative to the ⟨110⟩ Fe–P dumbbell state (Table 5.4). Most Fe–Fe dumbbell states have energy levels at least 1 eV above that of the Fe–P dumbbell. This implies that an

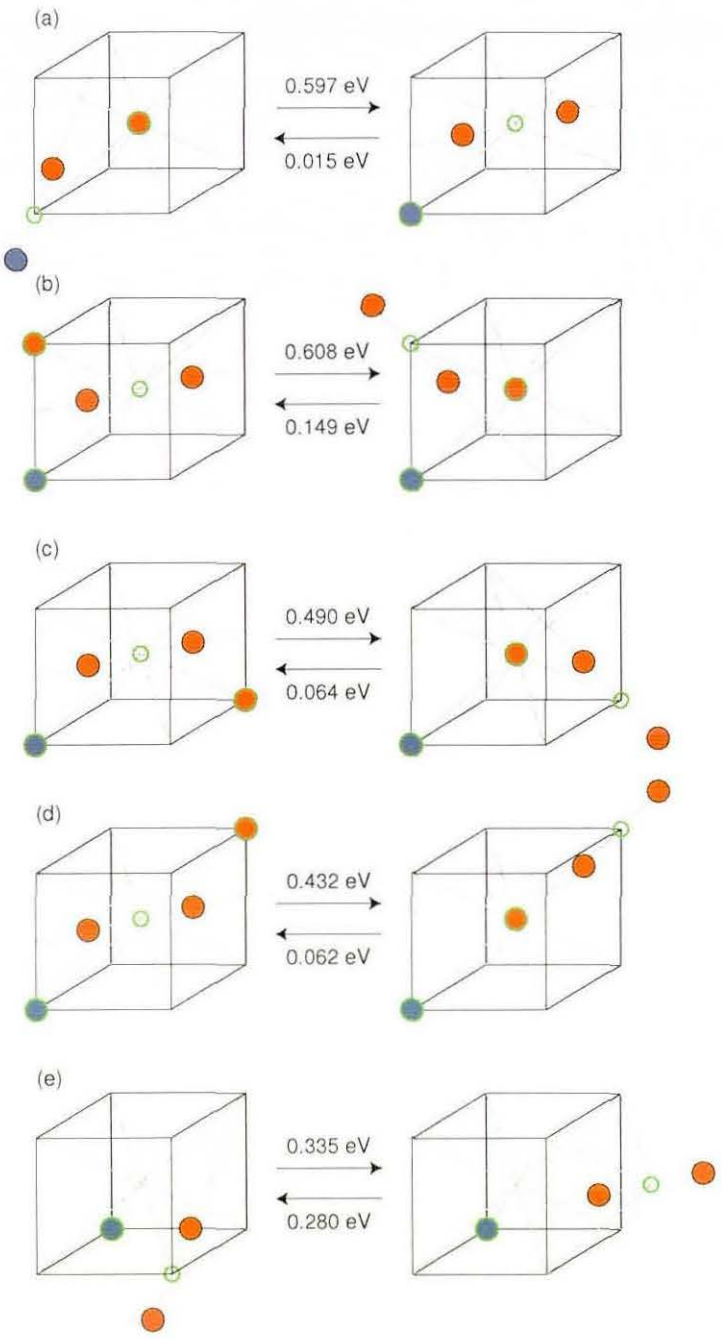


Figure 5.14: A selection of transitions between ⟨110⟩ dumbbells in the P interstitial defect system, with energy barriers.

State	Energy level (eV)
⟨110⟩ Fe–P	0.000
Tetrahedral	0.212
[110] at 1st [111]	0.581
$[1\bar{1}0]$ at 1st [111]	1.330
[110] at 2nd [100]	1.040
[011] at 2nd [100]	1.248
[110] at 3rd [110]	1.441
$[1\bar{1}0]$ at 3rd [110]	1.058
[101] at 3rd [110]	1.006

Table 5.4: *Potential energy levels of different states in the P interstitial defect system, relative to the ⟨110⟩ Fe–P dumbbell state. All states except for the first two are Fe–Fe dumbbell positions near a substitutional P atom. For each of these, two direction vectors are given: the orientation of the dumbbell and the direction of the displacement of the dumbbell centre from the P atom.*

interstitial defect is more strongly attracted to a P atom than a vacancy is. It should therefore be expected that the system spends most of its time moving between the Fe–P dumbbell and tetrahedral states. Also, since the main energy barriers for P migration are lower than those involved in the P–vacancy system, and it also takes fewer events for the P atom to move the same distance, it is to be expected that the P atom will diffuse more rapidly by moving as an interstitial defect than as a P–vacancy complex.

5.4 Conclusions

In this chapter, two different potentials have been used to study defect transitions in bulk Fe with and without P impurities. In carrying out this work, it was discovered that the choice of potential can have a significant effect on what transitions are found and on their energy barriers. This demonstrates that in order to obtain realistic results, it is

necessary to choose an accurate interatomic potential. The Ackland Fe–P potential used in section 5.3 is fitted to *ab initio* data, unlike the older potentials used for section 5.2, and developed with the aim of accurately modelling common defect structures with and without the P impurity atoms. In this work, the energy barriers found were close to those determined by *ab initio* calculations [55, 56]. As such, it appears that the newer Ackland Fe–P potential is more sophisticated than the older set of potentials used to model this system.

A difficult aspect of the study of transitions between states in atomistic systems is knowing which metastable states really exist and which are artefacts of the interatomic potential being used. In the older potential, several states were found to exist that did not occur when using the newer potential or in the *ab initio* studies, including the unusual dumbbell orientations. As such, it is very likely that these states are artefacts of the Ackland/Morse model used in section 5.2. This is further evidence that the newer Ackland Fe–P potential is more sophisticated for the modelling of Fe in which P impurity atoms may be present. Therefore, the results from the Ackland Fe–P potential will be used to draw the remaining conclusions and for the kinetic Monte Carlo work of the next chapter.

In the studies of this chapter, the lowest energy barriers were generally those found in transition mechanisms that keep the vacancy close to the P atom in the P–vacancy complex, and that keep the P atom moving as part of the interstitial defect. It follows that, in agreement with previous studies, P atoms act as centres of attraction for both vacancy and interstitial defects in Fe.

The results presented in this chapter provide a catalogue of transitions between various configurations of point defects in α-Fe with and without a P impurity. In the next chapter, these results are used to perform kinetic Monte Carlo on three different defect systems and to calculate diffusion constants for each.

Chapter 6

Kinetic Monte Carlo

The kinetic Monte Carlo (KMC) technique was briefly described in section 2.3.4. This chapter explains in more detail how KMC works and how it has been implemented for this project. The technique is then applied to three of the Fe and Fe–P systems studied in the previous chapter, and used to calculate the diffusion constants of the defect systems.

6.1 Methodology

KMC is an event-driven simulation technique, *i.e.* it advances the system by one transition at a time. This is in contrast to the other accelerated dynamics methods described in section 2.3 which, like classical MD, still involve simulating an MD trajectory over time.

From a given initial state, a transition is chosen randomly from those that are possible. The system is advanced to the new state, and the simulation time clock is incremented discretely. The process is then repeated from the new state, and continued to produce a sequence of states through which the system passes.

The probability distributions used here to select a random transition and to calculate the escape time are based on harmonic transition state theory (hTST). The relative probabilities of the transitions and the mean escape time are determined by the rate constants of the transitions, which are in turn calculated from the energy barriers and the temperature of the system. The rate constant of a transition is the mean frequency

of the transition over the time that the system spends in the transition's initial state. For a transition from state i to state j, the rate constant is given by the Arrhenius equation:

$$r_{ij} = \nu \, \mathrm{e}^{-E_{ij}/k_B T} \tag{6.1}$$

where E_{ij} is the transition's energy barrier, k_B is Boltzmann's constant, T is the temperature and ν is the attempt frequency. As with the rate constants that have been given in previous chapters, the typically assumed attempt frequency $\nu = 10^{13}$ s^{-1} has been used for the rate constant calculations in the work of this chapter.

In the hTST model, the time for which the system remains in a given state is exponentially distributed and independent of the state to which the system is to exit. The mean of this distribution, for a state i, is

$$\tau_i = \frac{1}{\sum_j r_{ij}} \tag{6.2}$$

where the sum is taken over all possible transitions from state i. Consequently, given a second state j, the probability that a given transition is next to occur is proportional to its rate constant and is given by

$$p_{ij} = r_{ij}\tau_i. \tag{6.3}$$

These equations give rise to the formula used to increment the simulation time clock. Given a uniform random variable $R \in (0, 1)$, the increment is

$$\Delta t = -\tau_i \ln R. \tag{6.4}$$

6.1.1 Dealing with symmetry

Each distinct state exists in a number of different orientations. Different states also have different symmetries. This influences the number of possible orientations of the transition from one state to another, and needs to be taken into account when weighting the transition probabilities, choosing a transition and updating the coordinates.

For example, in considering an impurity, located at the blue site in Figure 6.1, the first neighbour site of an impurity–vacancy complex has eight orientations, each of which has

triangular symmetry about a ⟨111⟩ axis. From a given first neighbour orientation, there are three possible orientations of the second neighbour impurity–vacancy state towards which a transition can be made, three orientations of the third neighbour state and one of the fifth neighbour state. This is due to the numbers of lattice sites of each kind adjacent to the first neighbour site (Figure 6.1 (a)). The second neighbour state, on the other hand, has six orientations and is symmetrical about a ⟨100⟩ axis. Of its eight first neighbours, four are first neighbours of the impurity, and four are fourth neighbours of it (Figure 6.1 (b)). The defect can transition from the second neighbour state towards any of these eight adjacent sites.

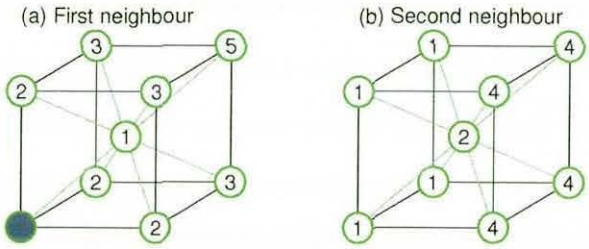


Figure 6.1: Neighbour numbers of bcc lattice sites adjacent to the first (a) and second (b) neighbour positions.

Essentially, each orientation of a state should be considered as a state in its own right. However, to simplify the process of configuring the simulation, it is useful if the user does not have to duplicate information for each orientation of every state and transition. This is achieved by implementing symmetry checking at run time. A variation of equation (6.1) is therefore used:

$$r_{ij} = \nu \, S_{ij} \, \mathrm{e}^{-E_{ij}/k_B T} \tag{6.5}$$

where S_{ij} is the number of orientations of the final state that can be reached from a given orientation of the initial state. For example, Figure 6.1 shows that $S_{12} = S_{13} = 3$, $S_{15} = 1$ and $S_{21} = S_{24} = 4$, where the subscripts denote the neighbour numbers of impurity–vacancy states. In practice, these transitions often involve intermediate split vacancy states, but this does not affect the number of orientations, since the symmetry

of a transition does not depend on how far the defect moves in a given direction. The following section describes how transition symmetries, and hence the S_{ij} values, can be calculated.

Enumerating the orientations of a transition

Impurity–vacancy complex

Algorithmically, determining the symmetry of states and transitions is a matter of enumerating the transformations of the displacement vector dimensions that yield compatible but not identical results. In the impurity–vacancy case, three vectors are involved in any transition: the displacement of the vacancy from the impurity atom in the initial state, the vector by which the impurity atom moves and the vector by which the vacancy moves. For split vacancy states, the location of the vacancy is taken to be that of the interstitial atom between the two vacancies. Compatible orientations are those transformations of the three vectors simultaneously that leave the impurity–vacancy displacement unchanged. Because of the cubic lattice structure, only permutations of the three dimensions and reflections in coordinate planes are to be considered.

For example, in the first neighbour impurity–vacancy state, the impurity–vacancy displacement is a [111] vector. In the transition to second neighbour, the impurity atom remains fixed and the vacancy moves in a $[1\bar{1}\bar{1}]$ direction. In this case, permutations of the dimensions preserve the impurity–vacancy displacement but reflections in coordinate planes do not, so the possible vacancy moves to a second neighbour state are $[1\bar{1}\bar{1}]$, $[\bar{1}1\bar{1}]$ and $[\bar{1}\bar{1}1]$.

The second neighbour state has a [100] impurity–vacancy displacement, and the vacancy moves to first neighbour through a $[\bar{1}11]$ vector. Since the first component of the impurity–vacancy displacement differs from the others, it cannot interchange with the others in the vacancy move vector. However, since the other components of the impurity–vacancy displacement are zero, the transition can be reflected in the coordinate planes corresponding to these, so that there are four orientations of the transition: $[\bar{1}11]$, $[\bar{1}1\bar{1}]$, $[\bar{1}\bar{1}1]$ and $[\bar{1}\bar{1}\bar{1}]$.

In configuring the KMC simulation, it was useful to devise a canonical form for each state, and to classify states and transitions by symmetry. For the impurity–vacancy complex, the displacement vector between the impurity atom and the vacancy is the only item of information required to identify a state. In the canonical form, this vector is given with all components non-negative, and in descending order of value. As an exception, since for an $\langle abb \rangle$ vector with $a, b \neq 0$ there is no need to distinguish the cases $a < b$ and $a > b$, both are canonicalised as $[abb]$.

Interstitial defect

The system of an interstitial defect involving an impurity atom is more complicated, because in the defect structures consisting of a substitutional atom close to a self-interstitial dumbbell, the orientation of a dumbbell relative to the impurity atom also needs to be considered.

Each state is determined by two vectors as illustrated in Figure 6.2. For dumbbell defects, the first vector is the displacement of the vacancy at the centre of the dumbbell from the impurity atom, and the second is the displacement of one of the atoms forming the dumbbell from the impurity atom. For an impurity atom in a tetrahedral position, the zero vector is used for the first vector and the displacement to the nearest lattice point for the second. For a self-interstitial tetrahedral or octahedral defect, the two vectors are both equal to the displacement of the defect from the impurity atom.

There are then three vectors used to identify each transition. These correspond to the movement of the impurity atom and the movement of the coordinates denoted by the first and second state vectors respectively.

Because there is an extra vector to consider, canonicalisation of the interstitial defect states is a little more complicated than it is for the impurity–vacancy complex. First, all negative components of the first vector are made positive, and the corresponding components of the second vector are inverted in sign. Second, any components of the second vector corresponding to zero components of the first vector are made positive. Then the dimensions of the resulting vectors are ranked in descending order of the component of

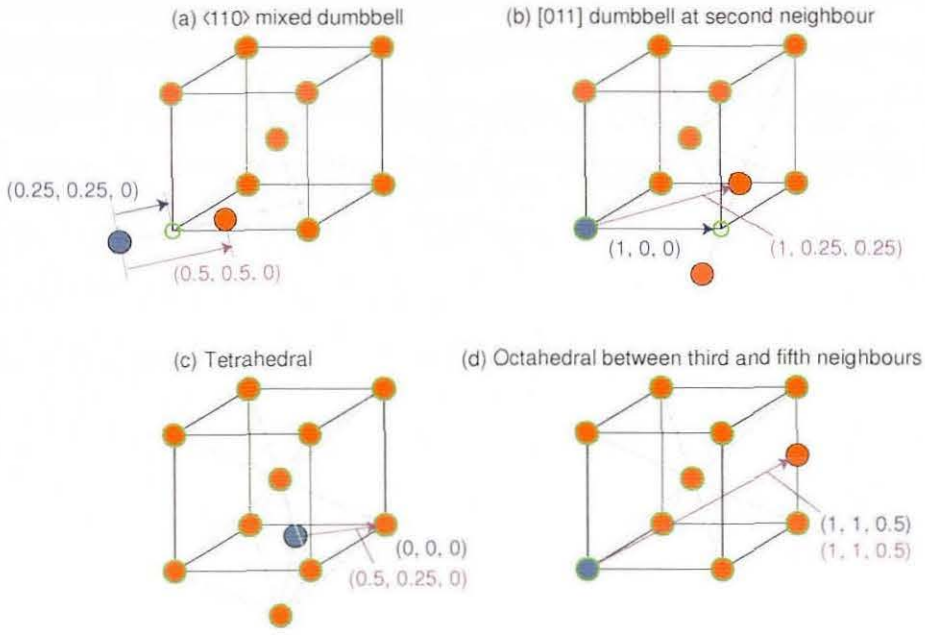


Figure 6.2: Some examples of vectors used to denote states in the interstitial defect system. Vectors have been given in units of the lattice parameter. In each case, the first vector is given in blue, and the second in purple.

the first vector, using the corresponding component of the second vector as a tie-breaker. As a further canonicalisation, of the two atoms forming a self-interstitial dumbbell, the one whose displacement is given by the second vector is that which has the greater first coordinate.

When the vectors defining the state are in canonical form, the orientations of the transitions from the state can be enumerated in the same basic manner as for the impurity–vacancy complex. However, a few special cases need to be considered. In some states, such as the $[1\bar{1}0]$ dumbbell at a first neighbour bcc site, the positions of the two dumbbell atoms are equivalent, and so it was necessary to consider that transitions can occur from this state in either direction. Also, because of the way in which a tetrahedral impurity atom is specified, it is necessary to treat this as a special case when determining the compatible orientations.

Updating the coordinates

Having chosen a transition to apply, an orientation of the transition is then randomly chosen. To update the coordinates, it is necessary to calculate the correct transformation of the transition vectors for both the state orientation and the orientation chosen for the transition.

The orientation process is done in two stages. The first is to match the coordinates of the transition vectors to the coordinates of the actual state. This is done by forming a one-to-one mapping of the dimensions of the original vectors to those of the current state, and noting where a negative needs to be taken. This provides a view of the current coordinates in a vector basis that makes the coordinates match those of the canonical form. Table 6.1 gives a sample sequence of vector basis transformations to canonicalise a set of coordinates.

Vector basis	Impurity coordinates	Vacancy coordinates	Displacement
$[xyz]$	$(1, 2, 3)$	$(1, 1, 4)$	$(0, -1, 1)$
$[zyx]$	$(3, 2, 1)$	$(4, 1, 1)$	$(1, -1, 0)$
$[z\bar{y}x]$	$(3, -2, 1)$	$(4, -1, 1)$	$(1, 1, 0)$

Table 6.1: Vector basis transformation of the impurity and vacancy coordinates of a sample impurity–vacancy complex in a bcc third neighbour position. The vectors are given in units of the lattice parameter. The first line gives the raw coordinates of the impurity atom and vacancy. These coordinates are transformed first by swapping the x and z coordinates (second line) and then by inverting the sign of y (third line) so that the impurity–vacancy displacement vector is in canonical form.

The second stage is to orient the coordinates in order to effect the transition in its randomly chosen orientation. The orientations of a transition correspond to the transformations of the state coordinates and move vectors together that retain the displacement vector(s) defining the state impurity–vacancy displacement (or the two vectors defining an impurity–interstitial state) in canonical form. For example, for the impurity–vacancy

complex in a bcc lattice, the transition from the third to the fourth neighbour state has an impurity move vector of $(0, 0, 0)$ and a vacancy move vector of $(0.5, -0.5, 0.5)$. This has four such transformations, given in Table 6.2.

Vector basis	Impurity coordinates	Vacancy coordinates	Vacancy move vector
$[z\bar{y}x]$	$(3, -2, 1)$	$(4, -1, 1)$	$(0.5, -0.5, 0.5)$
$[\bar{y}zx]$	$(-2, 3, 1)$	$(-1, 4, 1)$	$(-0.5, 0.5, 0.5)$
$[z\bar{y}\bar{x}]$	$(3, -2, -1)$	$(4, -1, -1)$	$(0.5, -0.5, -0.5)$
$[\bar{y}z\bar{x}]$	$(-2, 3, -1)$	$(-1, 4, -1)$	$(-0.5, 0.5, -0.5)$

Table 6.2: Vector basis transformations of the transition from third to fourth neighbour P–vacancy states that preserve the canonical form of the P–vacancy displacement and hence are possible transitions from the state orientation being considered.

Internally, three pointers are mapped to the data relating to the x, y and z dimensions respectively, and each has an associated sign flag. To prepare for each transition event, the pointers are arranged to view the current state in canonical form. An orientation of the chosen transition is then chosen randomly, and the pointers and sign flags are updated by applying the required basis transformation. To update the coordinates, the simulation then applies the transition's move vectors through these pointers.

6.1.2 Dissociation

It is possible for the vacancy or interstitial defect to become dissociated from the impurity atom. To allow for this, a dissociation distance is defined. When the vacancy or interstitial is further from the impurity atom than this distance, it is considered to have dissociated, and then it behaves as an isolated vacancy or an isolated self-interstitial, not influenced by the impurity.

For a dissociated defect, the canonical form of a state is based on a displacement vector connecting the defect to a lattice point, ignoring the location of the impurity atom.

When the defect is dissociated, it may later reassociate with the same impurity atom or associate with another one. To model the finite concentration of impurity atoms in the material, periodic boundary conditions were used. Each periodic cell contains a single impurity atom, and the size of the cell is chosen to represent the concentration of impurity atoms being modelled. For example, a periodic cell of $11 \times 11 \times 11$ bcc unit cells contains 2662 atoms, giving a 0.038% concentration of impurity atoms.

6.2 Calculating diffusion constants

The diffusion constant is a commonly used measure of the rate at which a defect diffuses. It is calculated by the formula

$$D = \frac{\langle |\mathbf{r}|^2 \rangle}{6t_{sample}} \tag{6.6}$$

where $\mathbf{r}$ is the displacement of the defect from its starting position after time t_{sample}. In cases such as those studied here, where the defect explores the bulk of the material in three dimenisons, the diffusion constant is independent of the interval t_{sample} that has been chosen. To obtain an accurate diffusion constant, many uncorrelated samples must be taken. In this work, the KMC simulation is run for a long time (see section 6.3), with samples taken at fixed simulation time intervals of t_{sample}, with each $\mathbf{r}$ taken to be the displacement of the defect from where it was at the previous multiple of t_{sample}. For computational convenience, the equivalent form

$$D = \frac{\sum |\mathbf{r}|^2}{6t_{run}} \tag{6.7}$$

is used, where t_{run} is the total amount of time for which the simulation has run.

Although the term *diffusion constant* is used, the measure varies with temperature. In a typical system, this dependence is governed by the Arrhenius relation

$$D = D_0 \mathrm{e}^{-E/k_B T} \tag{6.8}$$

where D_0 and E are system-dependent parameters. The Arrhenius relation can be visualised by plotting the diffusion constant on a logarithmic scale against the temperature

on a reciprocal scale, thereby producing an *Arrhenius plot*. For a system that obeys the Arrhenius relation, the points plotted will lie on a straight line.

One of the aims of this work is to determine whether this relation holds for the systems being studied and to calculate the values of D_0 and E.

6.3 Results for Fe and Fe–P systems

The KMC method has been used to perform simulations on three different defect systems in Fe. The first of these systems is an isolated vacancy. The other two systems involve a P impurity atom. The first of these systems is the P–vacancy complex, and the second is the P interstitial defect system.

For these simulations, the transition table was supplied in advance, and composed of the states and transitions discovered in section 5.3 using the Ackland Fe–P interatomic potential.

In order to calculate accurate diffusion constants, it is important both to use a sufficiently large t_{sample} interval for the $\mathbf{r}$ samples to be meaningful, and sufficiently many samples to achieve convergence of the calculated diffusion constant. Each simulation was run for at least 10000 samples, which was found to give diffusion constants converged to two significant figures. At 350 K, where events are infrequent, we found that a sampling interval of 1 ms gave the best results. At higher temperatures, events happen more frequently, so shorter sampling intervals (down to 10 μs) and total simulation times of about 0.1 s were found to be adequate. Altogether, the total length of time covered in each simulation ranged from 0.1 s for the simulations run at 700 K, to 10 s as was used at 350 K. In this time, a total number of events in the order of 10^{10} was generally observed.

6.3.1 Isolated vacancy

As was discovered in section 5.3.1, an isolated vacancy diffuses by alternating between two states, namely the simple vacancy state and a metastable split vacancy configu-

Temperature (K)	Diffusion constant (m^2 s^{-1})
350	2.5×10^{-16}
400	3.6×10^{-15}
500	1.5×10^{-13}
600	1.8×10^{-12}
700	1.0×10^{-11}

Table 6.3: Diffusion constants for an isolated vacancy in Fe.

ration. KMC simulations have been performed on this defect system at five different temperatures, and the diffusion constants have been calculated.

Table 6.3 gives the diffusion constants at these five temperatures. It can be seen that as the temperature is increased, the diffusion constant increases in order of magnitude. It can be seen from Figure 6.3 that this system obeys the Arrhenius relation. The constants involved in equation (6.8) for the isolated vacancy system are $D_0 = 4.2 \times 10^{-7}$ m^2 s^{-1} and $E = 0.64$ eV.

6.3.2 P–vacancy complex

Simulations were performed on a P–vacancy complex, with the P atom and vacancy initially at first neighbour positions. As was seen in section 5.3.2, convergence of the energy barriers towards those of an isolated vacancy was seen beyond the eighth neighbour distance.

The exchange of the P atom and vacancy when they are at first neighbour positions, as was given in Figure 5.8, turned out as expected to be the most frequently observed mechanism. Overall, the system overcomes an energy barrier of 0.287 eV during this exchange. At a temperature of 300 K, only the transitions corresponding to this mechanism were observed, hence the defect did not diffuse through the lattice.

KMC simulations in this system have been performed using two different periodic cell sizes at five different temperatures. The first cell size is $11 \times 11 \times 11$ bcc unit

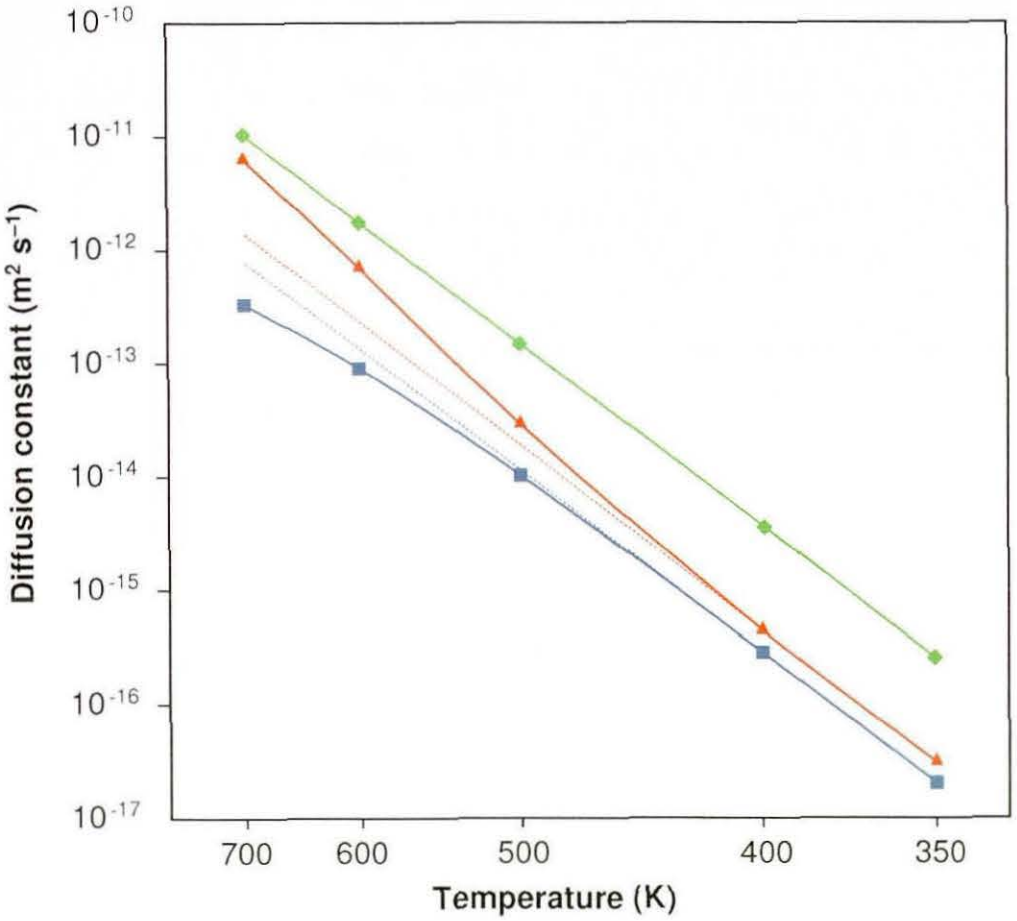


Figure 6.3: Arrhenius plot of the diffusion constants of an isolated vacancy in Fe (◆), and of a P atom (■) and a vacancy (▲) in a P–vacancy complex with a periodic cell size of $15 \times 15 \times 15$ bcc unit cells. The fine dotted lines show the extrapolation from the diffusion constants in the P–vacancy system calculated at 350 K and 400 K if the Arrhenius relation is assumed.

Temperature (K)	Diffusion constants (m^2 s^{-1}) P atom	Vacancy	Time associated
350	2.3×10^{-17}	1.2×10^{-16}	99.8%
400	3.2×10^{-16}	1.3×10^{-15}	99.4%
500	1.1×10^{-14}	1.9×10^{-14}	95.2%
600	1.1×10^{-13}	4.1×10^{-13}	83.9%
700	5.1×10^{-12}	4.1×10^{-12}	66.9%

Table 6.4: Diffusion constants and association times for the P–vacancy complex with a PBC cell of $11 \times 11 \times 11$ bcc unit cells, representing a 0.038 at.% P concentration. The time associated is defined here as the percentage of the time for which the vacancy was within 6.8 Å of the P atom, *i.e.* at up to the eighth neighbour position.

cells (31.5 Å), representing a P concentration of 0.038 at.%. The results are shown in Table 6.4. The P atom diffusion constants calculated at this cell size for 500 K and 600 K agree with those calculated by Barashev [57]. The second cell size that was studied is $15 \times 15 \times 15$ bcc unit cells (43.0 Å), representing a P concentration of 0.015 at.%; the results at this concentration are given in Table 6.5. For both concentrations and for all five temperatures, it can be seen that the P impurity makes the vacancy less mobile than it is in pure Fe. Also, on the Arrhenius plot (Figure 6.3) for this system, neither the diffusion constants of the P nor those of the vacancy lie on a straight line. The P–vacancy system therefore does not satisfy the Arrhenius relation.

Tables 6.4 and 6.5 also compare the proportions of the time for which the vacancy is associated with the P atom. As the temperature is increased, the amount of time over a given interval for which the system is dissociated also increases. This shows that, while the vacancy is strongly attracted to a substitutional P atom, the effect of this attraction weakens as the temperature increases. This is as expected, since the higher-energy transitions are more probable at higher temperatures.

This increase in the time that the P atom and vacancy spend dissociated also means

Temperature (K)	Diffusion constants (m^2 s^{-1}) P atom	Vacancy	Time associated
350	2.0×10^{-17}	3.0×10^{-17}	99.4%
400	2.8×10^{-16}	4.3×10^{-16}	98.7%
500	1.0×10^{-14}	2.8×10^{-14}	88.5%
600	8.9×10^{-14}	6.8×10^{-13}	66.9%
700	3.3×10^{-13}	6.2×10^{-12}	43.6%

Table 6.5: Diffusion constants and association times for the P–vacancy complex with a PBC cell of $15 \times 15 \times 15$ bcc unit cells, representing a 0.015 at.% P concentration. The time associated is defined here as the percentage of the time for which the vacancy was within 6.8 Å of the P atom, *i.e.* at up to the eighth neighbour position.

that, as the temperature increases, there is less time in which the P can diffuse. On the other hand, the vacancy is more mobile when dissociated than when moving in the P–vacancy complex. This accounts for the non-Arrhenius behaviour, since as dissociation time increases, the overall P mobility is decreased, and the overall vacancy mobility is increased, compared with what would otherwise be expected with the variation in temperature.

The diffusion constants were also compared at 500 K for five different periodic cell sizes, representing P concentrations from 0.038 at.% down to 0.006 at.%. The results are in Table 6.6. As the P concentration decreases, the diffusion constant of the P atom decreases, and the diffusion constant of the vacancy increases. The time for which the vacancy and P atom are associated also decreases at lower P concentrations. This illustrates that when the P concentration is lower, the vacancy diffuses further before associating with another P atom. It follows that the vacancy is more mobile, hence the increased diffusion constant, but the P atom has less opportunity to diffuse, leading to a decreased diffusion constant.

Periodic cell size (bcc unit cells)	P atom concentration	Diffusion constants (m^2 s^{-1}) P atom	Vacancy	Time associated
$11 \times 11 \times 11$	0.038 at.%	1.1×10^{-14}	1.9×10^{-14}	95.3%
$13 \times 13 \times 13$	0.023 at.%	1.1×10^{-14}	2.2×10^{-14}	92.3%
$15 \times 15 \times 15$	0.015 at.%	1.0×10^{-14}	2.8×10^{-14}	88.5%
$17 \times 17 \times 17$	0.010 at.%	9.6×10^{-15}	3.4×10^{-14}	84.1%
$20 \times 20 \times 20$	0.006 at.%	8.7×10^{-15}	4.5×10^{-14}	76.1%

Table 6.6: Diffusion constants for the P–vacancy complex at 500 K with different periodic cell sizes, representing different concentrations of P atoms in Fe.

6.3.3 P interstitial defect

The system of an interstitial defect involving a P atom was studied using KMC simulations at the same five temperatures as those used for the studies of isolated vacancy and P–vacancy diffusion, using a periodic cell size of $11 \times 11 \times 11$ bcc unit cells.

In section 5.3.4, it was discovered that the attraction of an Fe interstitial defect to a P impurity atom is stronger than the attraction that holds the P–vacancy complex together. Indeed, at all temperatures the KMC simulations showed that the interstitial defect spends over 98% of its time in the ⟨110⟩ Fe–P dumbbell state, with the P atom migrating via the P tetrahedral state. At 350 K and 400 K, the system was moving between these two states for over 99.99999% of the time, becoming an Fe–Fe dumbbell with the P atom in a substitutional position for only a few transition events at a time. Consequently, the interstitial defect did not become dissociated from the P atom.

This system has previously been studied using preliminary MD simulations [19]. At 1200 K, the interstitial defect dissociated from the P atom after 80 ps; however, at 600 K no dissociation was seen. However, by extending the simulation beyond the accessible MD timescale, we have seen that dissociation can occur at temperatures as low as 500 K, at which the interstitial went beyond the sixth neighbour position approximately 60 times per second, but only for short periods of the order of nanoseconds. At higher temperatures

Temperature (K)	Diffusion constant (m^2 s^{-1})	Time associated
350	2.9×10^{-12}	100%
400	8.3×10^{-12}	100%
500	3.8×10^{-11}	99.999987%
600	1.1×10^{-10}	99.999627%
700	2.2×10^{-10}	99.993761%

Table 6.7: Diffusion constants and association times for the P interstitial defect system with a PBC cell of $11 \times 11 \times 11$ bcc unit cells. The time associated is defined here as the percentage of the time for which the interstitial defect was within 4 Å of the P atom, *i.e.* at up to the sixth neighbour position.

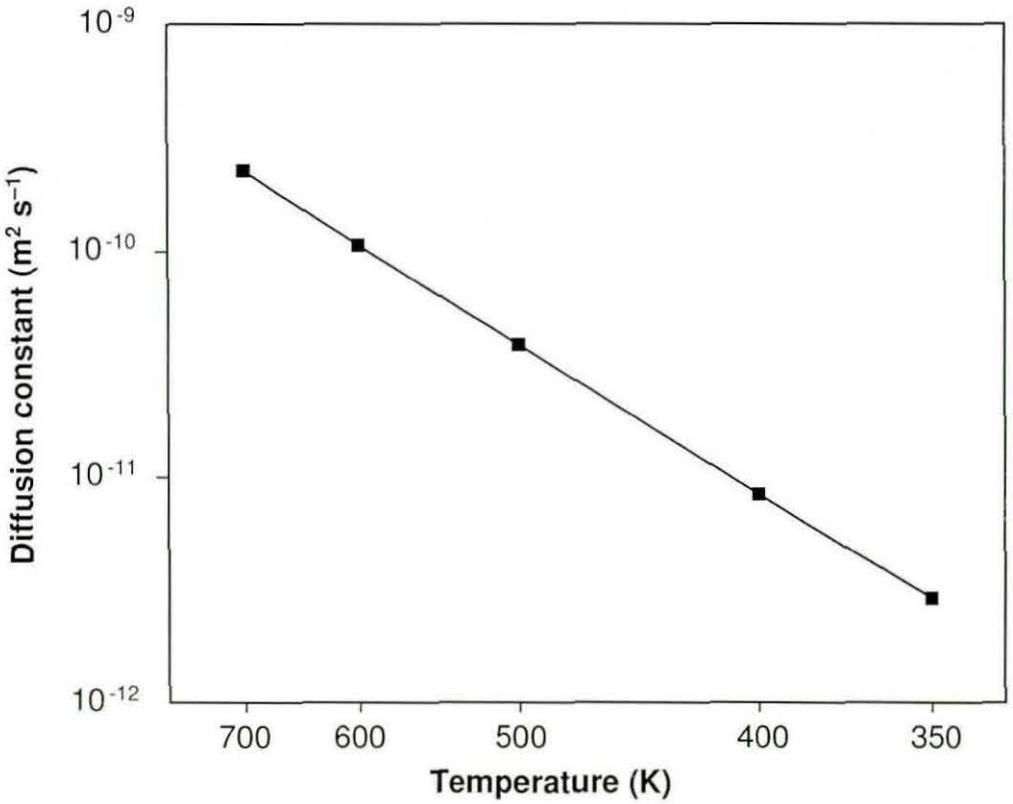


Figure 6.4: Arrhenius plot of the diffusion constants of a P interstitial defect in Fe.

dissociations are more frequent, but even at 700 K, the dissociation was for only 0.006% of the time.

The diffusion constants for this system at a periodic cell size of $11 \times 11 \times 11$ are given in Table 6.7. In this system, there was no noticeable difference between the diffusion constant of the P atom and that of the interstitial. As Figure 6.4 shows, there was also no noticeable deviation from the Arrhenius relation, with $D_0 = 1.7 \times 10^{-8}$ m^2 s^{-1} and $E = 0.26$ eV.

6.4 Conclusions

In the work of this chapter, the kinetic Monte Carlo method has been adapted to study the diffusion of isolated defects in bulk material. The technique has been applied to the diffusion of three different defect systems in α-Fe, and used to calculate the diffusion constants for these systems at various temperatures and concentrations of P impurity atoms.

It has been noted that in many systems, the Arrhenius relation governs the dependence of the diffusion constant on temperature. This relation was confirmed in the diffusion of an isolated vacancy and of a P interstitial defect in α-Fe over the range of temperatures that was studied. However, a deviation from this relation was seen in the diffusion of a P–vacancy complex.

In the P–vacancy system, the vacancy is attracted to the P atom. As the temperature increases, the vacancy becomes more easily able to overcome this attraction and become dissociated from the P atom. This attraction is even stronger in the P interstitial defect system, so that at this temperature, even at 700 K dissociation was seen for less than 0.01% of the simulation time.

The diffusion constants that were calculated are very different between the three defect systems. It has been found that, of the three defect systems studied in this chapter, the P–vacancy complex is the least mobile and the P interstitial defect is the most mobile.

This implies that the P interstitial defects in Fe are an especially important process by which P atoms can migrate to grain boundaries in NPVs subjected to radiation damage.

It has also been seen in this work that diffusion constants calculated using simulation methods can differ by orders of magnitude from those often used by nuclear engineers in rate theory models [11] to assess NPV structural integrity. This suggests that the results of such calculations need to be carefully interpreted. Although the processes described here are only some of those important in determining the build-up of impurities at grain boundaries, the methodology has great promise as a means to provide a consistent set of diffusion data for such rate theory models and hence a more accurate quantification of NPV structural integrity and reactor lifetimes.

Chapter 7

Conclusions and future work

The dimer method, described in Chapter 3, was the main driving force behind the work of this thesis. As such, the first part of this work focussed on implementing the dimer method and adapting it to the systems that were to be studied.

While implementing the dimer method, a number of modifications to the algorithm were considered, including:

- different formulae for rotating the dimer
- periodically relaxing the fixed atoms
- a modified sequence of rotation and translation steps, with rotation separated from choosing a plane in which to rotate
- parallelisation of the algorithm.

These modifications were found to improve the efficiency of the dimer method. However, a continuing difficulty is in determining an optimum set of dimer method parameters to find transitions in a given system. A good set of parameters is one that enables the dimer method to:

- successfully find a transition in a high proportion of searches
- perform a relatively low average number of force evaluations per search

- yield accurate energy barriers
- find as many as possible of the transitions that are possible from the given initial state.

Good parameter sets can be found by trial and error, but the computational cost of trying many sets of parameters is not always outweighed by the increased efficiency of whichever parameters turn out to be optimal. It would therefore be of interest to users of the dimer method if a systematic approach to choosing a good set of parameters in the first place can be developed.

The dimer method has been applied to two very different atomistic diffusion problems. The first of these is the diffusion of an Au adatom on the Au(100) surface. This is similar to a previous study of Al adatom diffusion on Al(100) [10] in both the way in which it was carried out and that both materials have the face-centred cubic structure. The results of the Au study showed that the energy barriers of the transitions involved are generally higher than those for the corresponding transitions in Al, and that the lowest-energy transitions are very different between the two systems. For Au, these lowest-energy transitions are mainly those in which small numbers of atoms move between sites. For Al, concerted displacements are the preferred mechanism for an adatom to diffuse. This is therefore an example of how two materials of the same structure can behave very differently on the atomic scale. However, the studies of two adatoms on Au suggest that there may be a similarity between Au and Al, namely that adatoms will tend to form compact clusters rather than scatter over a (100) surface.

The second diffusion problem that has been studied is the diffusion of point defects in bulk α-Fe. Both defects in pure Fe and defects involving a P impurity atom were studied. For this study, two different interaction models were used, and the results were compared. The first model is the Ackland potential for pure Fe [18] together with a Morse potential [25] to model the Fe–P and P–P interactions. The second model is the newer Ackland potential designed for the modelling of P impurities in Fe [19]. With each potential, different transition mechanisms were discovered, and different energy barriers

were calculated. Thus it has been demonstrated that in order to model materials on the atomic scale, a good choice of interatomic potential is essential. In this case, the newer Ackland Fe–P potential has produced the more accurate results, as compared with experimental and *ab initio* studies of the behaviour of Fe with and without P impurities. As such, it is believed that there is plenty of opportunity for this more sophisticated potential to be used in more studies of Fe containing P, and to produce more accurate results than have been derived from studies involving older potentials.

In studying surface diffusion and bulk diffusion, a number of differences between the two problems were seen. The number of free atoms required to obtain meaningful results using the dimer method is much higher in bulk material than on surfaces. This is because bulk defects often create strain fields of significant size around them, and calculating accurate energy barriers depends on the algorithm being able to determine how the atoms in the strain field move in the lowest-energy path. On the other hand, an adatom on a surface has a much smaller strain field, and can therefore be modelled using smaller numbers of free atoms. A second difference is that on surfaces there are more transitions possible from a given state than in the bulk. This is due to the possibility of atoms entering and leaving the substrate in surface configurations, leading to a rich variety of co-operative transitions such as concerted displacements. On the other hand, diffusion in bulk material is simpler, with only a small number of atoms moving between lattice sites in any transition.

Kinetic Monte Carlo simulations have been carried out on three different defect systems in bulk α-Fe. These are an isolated vacancy, a P–vacancy complex and a P interstitial defect. Diffusion constants were calculated for five different temperatures in each system, and the results have been compared with those that would be predicted by the Arrhenius relation

$$D = D_0 e^{-E/k_B T} \tag{7.1}$$

which is typical of many systems. Two of the systems studied were found to obey the Arrhenius relation, and the parameters calculated are $D_0 = 4.2 \times 10^{-7}$ m^2 s^{-1}, $E = 0.64$ eV for the isolated vacancy system and $D_0 = 1.7 \times 10^{-8}$ m^2 s^{-1} and $E = 0.26$ eV for

the P interstitial defect system. The P–vacancy complex, on the other hand, does not obey the Arrhenius relation. This was found to be because the time that the vacancy spends dissociated from the P atom increases with temperature. This is also true of the P interstitial system; however, there was not sufficient dissociation for any deviation to be seen. It is therefore possible that, if P interstitial diffusion is studied at sufficiently high temperatures, non-Arrhenius behaviour will begin to emerge in this system.

7.1 Future work

A limitation of the KMC simulations performed in this work is that a constant attempt frequency of 10^{13} s^{-1} has been used to calculate the rate constant of every transition. This is a typical value assumed in rate theory studies; however, in practice, the attempt frequency is not constant. An improvement would be to implement calculation of the actual attempt frequency of each transition using the method of Vineyard [38]. This involves calculating the eigenvalues of the Hessian matrix at the saddle point and the initial minimum of the potential surface. As such, it would require the calculation of second derivatives of the potential, which is a computationally expensive operation, but this calculation needs to be performed only for the initial state and each saddle point found.

To model Fe containing a finite concentration of P impurity atoms, a periodic cell with a single impurity atom was used for the KMC simulations. By varying the size of the cell, different impurity concentrations were considered. An alternative would be to use a much larger periodic cell with many substitutional impurity atoms distributed within it. This would be more realistic, since it enables more variation to be seen in how far a vacancy or interstitial defect moves between dissociating and reassociating with an impurity atom. A difficulty with this approach is that two impurity atoms may become sufficiently close together that they interact with each other, or both interact with the vacancy or interstitial at the same time. This could be addressed by using the predefined event table when the vacancy or interstitial is associated with only one impurity atom,

and employing an on-the-fly KMC (OFKMC) approach for the more complicated states involving multiple impurity atoms.

The methodology could also be adapted to investigate the evolution of defect clusters created by collision cascades. This would necessarily involve OFKMC simulation, but when there are isolated defects as well as clusters, a possibility would be to combine the event table with one generated on the fly. This would work by searching for transitions of each cluster separately, and combining the sets of transitions that can occur on each cluster, along with those in the predefined event table for the isolated point defects, into a single list of transitions from which one is randomly chosen. The separation of the search effort into clusters produces a divide-and-conquer approach to building the event table, which may improve the efficiency of OFKMC simulations and also has potential for parallelisation.

When the event table is predefined, KMC's event-driven nature means that it is faster than other accelerated dynamics methods, such as hyperdynamics or TAD, which still rely on trajectories evolved in small timesteps as does classical MD. However, OFKMC is a much more computationally intensive method than KMC with a predefined event table. At the moment, it is unknown how OFKMC will compare in its performance to the other techniques in studying the problems that have been covered in this work, but it is likely that the aforementioned divide-and-conquer approach will bring about an improvement in efficiency that may lead to further advances in long-timescale simulation applied to the problems studied here and to other atomistic modelling problems.

At the moment it is unclear how the various accelerated dynamics methods compare to each other in their performance. Although the methods have been studied and reviewed extensively, comparisons of the different methods applied to the same problems have not been found in the literature. Thus it would be useful to carry out a systematic comparison of the accelerated dynamics methods, by applying each method to the same set of problems. This would produce an overview of how efficient and versatile the different techniques are in relation to each other, including typical boost factors, how the boost factor scales with temperature, and which problems each technique works well

on. This would, in turn, help scientists to choose the best accelerated dynamics method for the particular problem being studied, and to gauge the relative performance of new accelerated dynamics methods that may be developed.

It would also be of interest to attempt studies of the problems that have been presented here using *ab initio* models of the interatomic interactions. *Ab initio* calculations have already been used [55] to determine the energy barriers of some transitions in the Fe and Fe–P systems. However, only P–vacancy and P–interstitial separations up to the fifth neighbour position have been covered in the results published to date. These results could be extended into a more comprehensive catalogue of *ab initio*-calculated transition barriers, which can then be compared with the barriers calculated in this work using the Ackland Fe–P potential. Transition barriers calculated using *ab initio* methods can then be used for an event table to perform KMC on this system or other similar systems.

Most *ab initio* work has used molecular statics to calculate the energy barriers between known initial and final states. It would be useful to investigate other applications of *ab initio* molecular modelling, including the use of other barrier calculation methods such as the dimer method. This would mean that *ab initio* models can be applied not only to calculate the energy barriers, but also to discover any transitions that are missed when working with empirical potentials.

Bibliography

[1] S. D. Kenny, D. Mulliah, C. F. Sanz-Navarro, R. Smith, *Philosophical Transactions of the Royal Society A*, **363**, 1949 (2005).

[2] H. Hurchand, S. D. Kenny, C. F. Sanz-Navarro, R. Smith and P. E. J. Flewitt, *Nuclear Instruments and Methods in Physics Research B*, **229**, 92 (2005).

[3] A. F. Voter, F. Montalenti and T. C. Germann, *Annual Review of Materials Research*, **32**, 321 (2002).

[4] A. F. Voter, *Physical Review B*, **57**, 13985 (1998).

[5] M. R. Sørensen and A. F. Voter, *Journal of Chemical Physics*, **112**, 9599 (2000).

[6] F. Montalenti, *Nanotech 2002*, **2**, 132 (2002).

[7] A. F. Voter, *Journal of Chemical Physics*, **106**, 4665 (1997).

[8] A. F. Voter, *Physical Review Letters*, **78**, 3908 (1997).

[9] G. Henkelman and H. Jónsson, *Journal of Chemical Physics*, **115**, 9657 (2001).

[10] G. Henkelman and H. Jónsson, *Journal of Chemical Physics*, **111** (15), 7010 (1999).

[11] R. G. Faulkner, S. H. Song, P. E. J. Flewitt, M. Victoria and P. Marmy, *Journal of Nuclear Materials*, **255**, 189 (1998).

[12] G. R. Odette and G. E. Lucas, *Journal of the Minerals, Metals & Materials Society*, **53** (7), 18 (2001).

[13] Z. Lu, R. G. Faulkner, N. Sakaguchi, H. Kinoshita, H. Takahashi and P. E. J. Flewitt, *Journal of Nuclear Materials*, **329**, 1017 (2004).

[14] Z. Lu, R. G. Faulkner and P. E. J. Flewitt, *Acta Metallurgica Sinica*, **41**, 79 (2005).

[15] B. P. Uberuaga, R. Smith, A. R. Cleave, F. Montalenti, G. Henkelman, R. W. Grimes, A. F. Voter and K. E. Sickafus, *Physical Review Letters*, **92** (11), 115505 (2004).

[16] S. H. Song, R. G. Faulkner, P. E. J. Flewitt, L. Q. Weng, *Journal of Materials Science & Technology*, **21**, 196 (2005).

[17] A. V. Barashev, *Philosophical Magazine Letters*, **82**, 323 (2002).

[18] G. J. Ackland, D. J. Bacon and A. F. Calder, *Philosophical Magazine A*, **75**, 713 (1997).

[19] G. J. Ackland, M. I. Mendelev, D. J. Srolovitz, S. Han and A. V. Barashev, *Journal of Physics: Condensed Matter*, **16**, S2629 (2004).

[20] D. W. Brenner, *Physical Review B*, **42** (15), 9458 (1990).

[21] D. W. Brenner, *Physical Review B*, **46** (3), 1948 (1992).

[22] M. S. Daw and M. I. Baskes, *Physical Review Letters*, **50**, 1285 (1983).

[23] M. S. Daw and M. I. Baskes, *Physical Review B*, **29**, 6443 (1984).

[24] G. J. Ackland et al, *Philosophical Magazine A*, **56** (6), 735 (1987).

[25] P. M. Morse, *Physical Review*, **34**, 57 (1929).

[26] M. Hashimoto, Y. Ishida, R. Yamamoto, M. Doyama and T. Fujiwara, *Scripta Metallurgica*, **16**, 267 (1982).

[27] F. H. Stillinger and T. A. Weber, *Physical Review B*, **31**, 5262 (1985).

[28] J. Tersoff, *Physical Review B*, **37**, 6991 (1988).

[29] J. Tersoff, *Physical Review B*, **39**, 5566 (1989).

[30] J. F. Ziegler, J. P. Biersack and U. Littmark, *The Stopping and Range of Ions in Solids*, Volume 1, Pergamon, New York (1985).

[31] M. W. Finnis and J. E. Sinclair, *Philosophical Magazine A*, **50**, 45 (1984).

[32] S. M. J. Gordon, H. Hurchand, S. D. Kenny and R. Smith, *Nuclear Instruments and Methods in Physics Research B*, **228**, 131 (2005).

[33] C. R. A. Catlow, K. M. Diller and M. J. Norgett, *Journal of Physics C: Solid State Physics*, **10**, 1395 (1977).

[34] W. T. Rankin, *Efficient Parallel Implementations of Multipole Based N-Body Algorithms* http://www.ee.duke.edu/~wrankin/SciComp/Papers/wtr-thesis.html (2005).

[35] R. Smith, *Atomic & Ion Collisions in Solids and at Surfaces*, Cambridge University Press, Cambridge (2005).

[36] D. Mulliah, *Molecular Dynamics Simulations of Nanofriction and Wear*, PhD thesis (2004).

[37] R. Marcelin, *Annals of Physics*, **3**, 120 (1915).

[38] G. H. Vineyard, *Journal of Physics and Chemistry of Solids*, **3**, 121 (1957).

[39] C. F. Sanz-Navarro and R. Smith, *Nuclear Instruments and Methods in Physics Research B*, (180), 130 (2001).

[40] C. F. Sanz-Navarro, *Atomistic Interactions of Clusters on Surfaces using Molecular Dynamics and Hyper-Molecular Dynamics*, PhD thesis (2002).

[41] J. A. Sprague, F. Montalenti, B. P. Uberuaga, J. D. Kress and A. F. Voter, *Physical Review B*, **66**, 205415 (2002).

[42] R. C. Pasianot, A. M. Monti, G. Simonelli, E. J. Savino, *Journal of Nuclear Materials*, **276**, 230 (2000).

[43] G. Simonelli, R. Pasianot, E. J. Savino, *Physica Status Solidi B*, **191**, 249 (1995).

[44] H. Jónsson, G. Mills and K. W. Jacobsen, *Classical and Quantum Dynamics in Condensed Phase Simulations*, 385 (1998).

[45] G. Henkelman, B.P. Uberuaga and H. Jónsson, *Journal of Chemical Physics*, **113**, 9901 (2000).

[46] G. M. Crippen and H. A. Scheraga, *Archives of Biochemistry and Biophysics*, **144**, 462 (1971).

[47] R. L. Hilderbrandt, *Computers & Chemistry*, **1**, 179 (1977).

[48] E. K. P. Chong and S. H. Żak, *An Introduction to Optimization*, Wiley, Chichester (2001).

[49] G. Henkelman, B. P. Uberuaga, D. J. Harris, J. H. Harding and N. L. Allan, *Physical Review B*, **72**, 115437 (2005).

[50] L. Xu, G. Henkelman, C. T. Campbell and H. Jónsson, *Physical Review Letters*, **95**, 146103 (2005).

[51] M. Basham, P. A. Mulheran and F. Montalenti, *Surface Science*, **565**, 289-299 (2004).

[52] H. Hurchand, S. D. Kenny and R. Smith, *Nuclear Instruments and Methods in Physics Research B*, **228** 146 (2005).

[53] B.-J. Lee and M. I. Baskes, *Physical Review B*, **62**, 8564 (2000).

[54] B.-J. Lee, M. I. Baskes, H. Kim and Y. K. Cho, *Physical Review B*, **64**, 184102 (2001).

[55] C. Domain and C. S. Becquart, *Physical Review B*, **71**, 214109 (2005).

[56] C. Domain and C. S. Becquart, *Physical Review B*, **65**, 024103 (2001).

[57] A. V. Barashev, *Philosophical Magazine*, **85**, 1539-1555 (2005).