Nuclear Energy Agency

NEA/NSC/R(2019)2

Unclassified

English text only

11 June 2020

NUCLEAR ENERGY AGENCY
NUCLEAR SCIENCE COMMITTEE

State-of-the-Art Report on Multi-Scale Modelling Methods

Please note that this document is available in PDF format only.

JT03462919

Foreword

The Working Party on Multi-scale Modelling of Fuels and Structural Materials for Nuclear Systems (WPMM) was established under the auspices of the Nuclear Energy Agency (NEA) Nuclear Science Committee (NSC) to review multi-scale models and simulations as validated tools to predict the behaviour and performances of fuels and structural materials, in support of nuclear systems design and fuel fabrication. The WPMM's objective is to promote the exchange of information on theoretical and computational methods, experimental validation, and other topics related to modelling and simulation of nuclear materials. The WPMM established the Expert Group on Structural Materials Modelling in 2009 to provide targeted critical reference reviews of the state of the art in relation to the use of multi-scale modelling so as to describe the changes induced by irradiation in structural nuclear materials. The aim of this expert group is to reliably reproduce experimental data, while providing the keys to understand and interpret existing experimental results, with a view to predicting the behaviour of structural nuclear materials under unexplored conditions and supporting the choice and the development of new materials. The present volume presents a state-of-the-art review of physical multi-scale models rooted in computational physics, to describe the properties and behaviour of materials of interest for the nuclear community.

Acknowledgements

The Nuclear Energy Agency (NEA) wishes to express its sincere gratitude to Dr Veena Tikare (Sandia National Laboratory, United States), Chair of the Expert Group on Multi-scale Modelling Methods, and to Dr Ram Devanathan (Pacific Northwest National Laboratory, United States), both of whom co-ordinated this report. The NEA would also like to thank the members of the Expert Group on Structural Materials Modelling and the authors who contributed to this report. Special thanks are also due to Professor Theodore M. Besmann (United States), Chair of the Working Party on Multi-scale Modelling of Fuels and Structural Materials for Nuclear Systems.

List of authors

M. J. Caturla	Universidad de Alicante, Spain
L. J. Criscenti	Sandia National Laboratories, United States
R. Devanthan	Pacific Northwest National Laboratory, United States
D. M. Duffy	University College London, United Kingdom
N. M. Ghoniem	University of California, United States
M. Krack	Paul Scherrer Institute, Switzerland
T. D. Kühne	University of Paderborn, Germany
N. V. Luzginova	Nuclear Research and Consultancy Group, Netherlands
D. Perez	Los Alamos National Laboratory, United States
J. Y. R. Rashid	ANATECH Corp., United States
D. Seif	University of California, United States
R. E. Stoller	Oak Ridge National Laboratory, United States
V. Tikare	Sandia National Laboratories, United States
A. I. Topuz	Materials Innovation Institute (M2i) and University of Groningen, Netherlands
B. P. Uberuaga	Los Alamos National Laboratory, United States
A. Uldry	Paul Scherrer Institute, Switzerland
E. van der Giessen	University of Groningen, Netherlands
A.F. Voter	Los Alamos National Laboratory, United States
M. J. Welland	Canadian Nuclear Laboratory, Canada
T. R. Zeitler	Sandia National Laboratories, United States

Table of contents

Tables

Figures

List of abbreviations and acronyms

1D	One-dimensional
2D	Two-dimensional
3D	Three-dimensional
AIMD	*Ab-initio* molecular dynamics
aKMC	Adaptive kinetic Monte Carlo
AKMC	Atomistic kinetic Monte Carlo
AMD	Accelerated molecular dynamics
ASPC	Always stable predictor-corrector
BCA	Binary collision approximation
bcc	Body centred cubic
BCMD	Basin constrained molecular dynamics
BKL	Borts-Kalos-Liebowitz
BMH	Born-Mayer-Huggins
BO	Born-Oppenheimer
BOMD	Born-Oppenheimer molecular dynamics
CALPHAD	Computer Coupling of Phase Diagraphs and Thermochemistry
CASL	Consortium for Advanced Simulation of Light Water Reactors
CEA	Commissariat à l'énergie atomique et aux énergies alternatives (Alternative Energies and Atomic Energy Commission, France)
CP	Car-Parrrinello
CPMD	Car-Parrinello molecular dynamics
CRSS	Critical resolved shear stress
DCs	Displacement cascades
DD	Dislocation dynamics
DFT	Density functional theory
DMFT	Dynamical mean field theory
DZVP	Double-zeta valence polarisation
EAM	Embedded-atom method
EFDA	European Fusion Development Agreement
EGM3	Expert Group on Multi-scale Modelling Methods (NEA)
EKMC	Event kinetic Monte Carlo
EOM	Equations of motion

EPRI	Electric Power Research Institute (United States)
fcc	Face centred cubic
FEM	Finite element method
F-P	Fokker-Planck
FPKMC	First passage kinetic Monte Carlo
FS	Finnis-Sinclair
GAWP	Gaussian Augmented Plane Waves
GFE	Gibb's free energy
GGA	Generalised gradient approximation
GPW	Gaussian Plane Waves
HE	Hartree equations
HK	Hohenberg-Kohn
HLW	High-level waste
HV	Helium-vacancy
ICT	Information and communications technology
INL	Idaho National Laboratory
KIM	Knowledgebase of Interatomic Models[1]
KMC	Kinetic Monte Carlo
KP	Kinchin Pease
KS	Kohn-Sham
LANL	Los Alamos National Laboratory
LDA	Local density approximation
LKMC	Lattice kinetic Monte Carlo
LOCA	Loss-of-coolant accident
LWR	Light water reactor
MC	Monte Carlo
MD	Molecular dynamics
MEP	Minimum energy pathway
MPS	Missing-pellet-surface
MOX	Mixed oxides
MQ-NMR	Multiple-quantum magic angle spinning nuclear magnetic resonance
M2i	Materials Innovation Institute (Netherlands)

[1] https://openkim.org

NEA	Nuclear Energy Agency
NEB	Nudged elastic band
ODS	Oxide dispersion strengethened
OKMC	Object kinetic Monte Carlo
OMC	Occupation matrix control
ORNL	Oak Ridge National Laboratory (United States)
OT	Orbital transformation
ParRep	Parallel replica dynamics
ParSplice	Parallel trajectory splicing
PAW	Projector augmented wave
PCI	Pellet cladding interaction
PCMI	Pellet-clad mechanical interaction
PDE	Partial differential equations
PIHD	Path integral hyperdynamics
PKA	Primary knock-on atoms
PNNL	Pacific Northwest National Laboratory (United States)
RCS	Replacement collision sequences
RHS	Right hand side
RIA	Reactivity insertion accident
RT	Rate theory
SEAKMC	Self-evolving atomistic kinetic Monte Carlo
SVV	Stress corrosion cracking
SCF	Self-consistent field
SCK	Smoluchowski-Chapman-Kolmogorov
SFE	Stacking fault energy
SFTs	Stacking fault tetrahedra
SIAs	Self-interstitial atoms
SIC	Self-interaction correction
SpecTAD	Speculative parallelisation temperature accelerated dynamics
SNL	Sandia National Laboratory (United States)
TAD	Temperature accelerated dynamics
TAMMBER	Temperature-accelerated Markov models with Bayesian estimation of rates
TEM	Transmission electron microscopy
TST	Transition state theory

TF	Thomas-Fermi
UFF	Universal force field
VR	Vashista-Rahman
WPMM	Working Party on Multi-scale Modelling of Fuels and Structural Materials for Nuclear Systems (NEA)
XC	Exchange and correlation
ZBL	Ziegler-Biersack-Littmark

Executive summary

Ram Devanathan[1] and Veena Tikare[2]
[1]Pacific Northwest National Laboratory, United States
[2]Sandia National Laboratories, United States

This report was commissioned by the Nuclear Energy Agency Expert Group on Multi-scale Modelling Methods (EGM3) under the Working Party on Multi-scale Modelling of Fuels and Structural Materials for Nuclear Systems (WPMM) of the Nuclear Science Committee (NSC). The goal is to provide to the nuclear energy community an overview of models and computer simulation methods of interest at different length and time scales for the materials process. Given the large scope of such an endeavour, it has taken the dedicated efforts of an international team of experts to do justice to this subject. The integration of these models provides a logical path to develop predictive understandings of the performance of materials used in the current fleet of nuclear reactors and to design new materials for advanced fission and fusion reactors.

The report starts at the electronic structure level with a chapter on density functional theory (DFT) written by Uldry and Krack. The authors introduce the framework of DFT, the independent electrons approximation, Hohenberg and Kohn theorems and the Kohn-Sham equations. The discussion also delves into the choice of exchange correlation functionals and the differences between various implementations of DFT. In the area of nuclear materials, DFT can be used to interpret experimental observations and to provide parameters for the empirical potentials used in larger scale molecular dynamics (MD) simulations. DFT calculations can be used to fill knowledge gaps, such as thermodynamic data for multicomponent systems where reliable experimental data is missing for materials systems of interest.

In the second chapter, Kühne introduces *ab-initio* molecular dynamics (AIMD), evaluating the interaction potential between atoms using parameter-free electronic structure calculations. By treating the ions classically, this method makes it possible to accurately determine both the static and dynamic properties of the system. The chapter presents Born-Oppenheimer MD and Car-Parrinello MD. By way of illustration, the author discusses results from liquid silicon and silica. Given that one can simulate systems with hundreds of atoms over a time scale of the order of nanoseconds, it is possible now to tackle nuclear materials problems using *ab-initio* methods that were previously considered impossible due to computational resource limitations.

It is possible to increase the length and time scales of the simulation, although with some compromise in terms of accuracy by performing classical MD simulations. Interatomic potentials, presented in Chapter 3, written by Zeitler and Criscenti, are at the heart of MD simulations. These potentials are analytical expressions of potential energy as functions of the interatomic distance, bond angles and related measures. These expressions use

parameters that are optimised by fitting them to known experimental properties, such as lattice constants and elastic constants, as well as data from DFT, such as defect formation energies. Such potentials can also be used in kinetic Monte Carlo (KMC) simulations. The authors begin with a discussion of the different functional forms of potentials in use and then focus on the state of the art in potentials for nuclear materials. Potentials for nuclear fuel, cladding, glassy materials and waste forms are analysed in detail. The chapter also explores the challenges in parameterising potentials with limited experimental data.

The fourth chapter on classical MD simulation, written by Duffy, continues the discussion of dynamical simulation of interacting atoms or ions by numerically integrating Newton's equations of motion. As an example of the close ties between chapters and the logical progression of scales, this chapter briefly touches upon the interatomic potentials covered in the previous chapter. The early stages of radiation damage can be modelled using MD simulations of displacement cascades. These simulations only account for nuclear stopping. For materials that are sensitive to the effects of electronic excitations, the authors point out the need to go beyond cascade simulations. With recent developments in reactive potentials and the incorporation of electronic effects in MD simulations, there is reason to be optimistic about the predictive modelling of the performance of irradiated materials.

Uberuaga, Perez and Voter take atomistic simulations a step further in the fifth chapter by presenting accelerated dynamics methods that enable the study of the long-time evolution of systems driven far from equilibrium. These methods fall under the categories of accelerated MD or adaptive KMC, and include parallel replica dynamics, hyperdynamics, temperature accelerated dynamics and κ-dynamics. The authors provide illustrative examples, such as non-equilibrium transport of Xe in irradiated UO_2 nuclear fuel, aggregation of defects in MgO and enhanced vacancy-interstitial recombination near grain boundaries. The chapter concludes with a discussion of the limitations of accelerated dynamics methods, and the potential to reach millisecond time scales and beyond with emerging computer architectures.

Chapter 6 also discusses the potential of KMC methods to reach much longer time scales than classical MD. In Chapter 6, Caturla and Stoller describe four KMC methods relevant to this report, namely atomistic KMC, object KMC, event KMC, and first passage KMC. The goal is to reach time scales of the order of seconds or even hours to enable direct comparison with experimental observations. The authors also present advanced methods, such as on-the-fly KMC that is especially relevant when the number of events possible is too large to be tabulated *a priori* or when the reactions strongly depend on the changing local environment. It is also possible to use machine learning to predict energy barriers on the fly in complex systems, where it is challenging to catalogue transition rates. There is considerable potential to enhance understanding of microstructural evolution under irradiation by combining KMC with other methods, including cluster dynamics and dislocation dynamics (DD).

In Chapter 7, Topuz, Luzginova and van der Giessen describe the DD method that lies between atomistic dynamics and crystal plasticity. DD uses atomistic input from DFT calculations, MD and KMC simulations and microscopy experiments, such as microstructural information, defect properties, defect density and defect distribution. In turn, DD provides the critical resolved shear stress and hardening rate as parameters to be used in crystal plasticity calculations. The focus of this chapter is on dislocation interactions with irradiation-induced defects. The authors provide keen insights into the limitations of DD and the need for further improvements, especially in describing fracture of nuclear materials.

The discussion moves further along the scale towards the continuum level in the next chapter. The phase-field model, presented in Chapter 8 by Welland, uses order parameters to model the evolution of the microstructure. Here the atomic scale is bypassed, and the focus is on microstructure and interface evolution at the mesoscale. This chapter presents the Allen-Cahn and Cahn-Hilliard models. Data from thermodynamic databases and parameters from DFT and MD in the form of interfacial energies and diffusion coefficients serve as much-needed input for the phase-field model. There is growing potential for expanding the reach of this method using novel techniques, for example adaptive mesh refinement that reduce the computational cost.

Chapter 9 reviews the rate theory of defect clustering in irradiated materials, with Seif and Ghoniem discussing the clustering of defects when materials are driven far from equilibrium due to bombardment by energetic particles. This theory is valuable for understanding microstructure evolution under irradiation at experimentally-relevant length and time scales. The authors review the physics of defect production under irradiation to provide an understanding of the processes that contribute to the generation and annihilation of defects. As in the case of the methods discussed earlier, DFT and MD play a valuable role in providing key parameters. There is also potential to couple rate theory to DD so as to understand embrittlement and failure of irradiated materials.

Finally, in Chapter 10, Rashid discusses the modelling of nuclear fuel using the finite element method. This macroscale model is used to understand the thermos-mechanical behaviour of irradiated nuclear fuel by solving coupled displacement and temperature equations. The chapter introduces the pellet-clad mechanical interaction problem and the constitutive formulation used to model fuel pellet fracture. Common codes that are used to implement the model and limitations of the method, such as the lack of microstructural representation, are also presented.

Overall, the modelling methods presented in the present report represent a progression of scales from the atomic to the macro scale. There is considerable potential to link scales by passing relevant parameters from the lower scale or by providing microstructural information from a higher scale to lower scale models. Such multi-scale modelling can overcome the limitations imposed by the current use of empirical models in computer codes, and enable predictive understanding of materials used in nuclear reactors.

1. Density functional theory

A. Uldry and M. Krack
Paul Scherrer Institute, Switzerland

1.1. Introduction

Density functional theory (DFT) forms the basis of the multi-scale ladder. It deals with the fundamental interactions in materials at the level of the electrons, and provides the means of solving the resulting quantum many-body problem. DFT is a first-principles method in the sense that, in its raw formulation, no empirical laws or adjustable parameters are needed. Going from the DFT formulation to an actual ground state requires, however, a series of approximations. The use of pseudopotentials may introduce certain types of errors, and the accuracy and applicability of the method rest with the choice of exchange-correlation functionals. The decision on the choice of functionals is usually made by comparison with experimental results or based on past experience with similar materials. Most functionals, including those popular within the nuclear materials community, benefit to some degree from cancellations of errors. Despite such shortcomings, DFT has been an extremely successful approach to understanding properties of matter and also as a predictive tool. Both the theoretical aspects of functionals and the computational implementations of DFT are active fields of research and subject to future developments.

1.2. The first-principles approach and the density functional theorems

This section sketches the essential facts behind DFT. Proofs and in-depth analysis can be found in the two text books upon which this introduction is loosely drawn, Richard Martin's (2004), and Parr and Young's book (1989). We first introduce the fundamental interactions that are taken into account. The central theorem of DFT, that all properties of an interacting system are unique functionals of the ground state density, is briefly stated. Finally, the formulation of the DFT as an independent particles problem in an effective potential is presented.

The starting point for a closed system of electrons and nuclei is to consider the kinetic energy and the electrostatic Coulomb interactions of all particles present. The Hamiltonian would therefore include the kinetic energy operator of the electrons $\hat{\mathrm{T}}_{\mathrm{e}}$ and of the nuclei $\hat{\mathrm{T}}_{\mathrm{n}}$, the interactions between electrons $\hat{\mathrm{V}}_{\mathrm{ee}}$, and between nuclei $\hat{\mathrm{V}}_{\mathrm{nn}}$, and the interactions between the electrons and the nuclei $\hat{\mathrm{V}}_{\mathrm{ne}}$. The wave functions of such a system depend simultaneously on the co-ordinates of all the electrons as well as that of the nuclei.

The very first step taken in the framework of DFT is to apply the *Born-Oppenheimer* approximation (1927), also called the adiabatic approximation. It consists in neglecting the kinetic energy of the nuclei

$$\hat{\mathrm{T}}_{\mathrm{n}} = -\sum_{I}^{N} \frac{\nabla_I^2}{2M_I} \tag{1}$$

This is reasonable in view of the large mass of the nuclei and their relative inertia compared to the electrons. As a consequence, the movements of the electrons are decoupled from that

of the nuclei. In classical molecular dynamics (MD) the nuclei are treated as point particles within Newton's formalism. In DFT, the positions of the nuclei are viewed as parameters, with the electrons moving instantaneously with the nuclei. One thereby reduces the problem to one of *electrons* interacting with one another via the Coulomb interaction, and with an external potential generated by the nuclei. The DFT Hamiltonian for a system consisting of N nuclei and n electrons in atomic units ([a.u.]: $\hbar = m_{\mathrm{e}} = e = \frac{1}{4\pi\epsilon_0} = 1$) is therefore given by

$$\begin{aligned}\widehat{\mathrm{H}} &= \widehat{\mathrm{V}}_{\mathrm{nn}} + \widehat{\mathrm{T}}_{\mathrm{e}} + \widehat{\mathrm{V}}_{\mathrm{ee}} + \widehat{\mathrm{V}}_{\mathrm{ne}} \\ &= \sum_{I}^{N}\sum_{J>I}^{N}\frac{Z_I Z_J}{|\boldsymbol{R}_J - \boldsymbol{R}_I|} - \sum_{i}^{n}\frac{\nabla_i^2}{2} + \sum_{i}^{n}\sum_{j>i}^{n}\frac{1}{|\boldsymbol{r}_j - \boldsymbol{r}_i|} - \sum_{I}^{N}\sum_{i}^{n}\frac{Z_I}{|\boldsymbol{r}_i - \boldsymbol{R}_I|}\end{aligned} \tag{2}$$

The labels i, j and I, J run over all electrons, respectively all nuclei; r_i and R_I denote the position of an electron, respectively of a nucleus of charge Z_I and mass M_I. The electrostatic interactions between nuclei $\widehat{\mathrm{V}}_{\mathrm{nn}}$ cause merely an energy shift, which can be ignored for the present discussion. At this stage the problem we are set to solve

$$\widehat{\mathrm{H}}\ \Psi(\{\boldsymbol{r}_i, \sigma_i\}) = E\ \Psi(\{\boldsymbol{r}_i, \sigma_i\}) \tag{3}$$

is still beyond reach. Although the wave function $\Psi(\{\boldsymbol{r}_i, \sigma_i\})$ now only depends on the co-ordinates $\boldsymbol{r}_i$ and spins σ_i of the electrons, this dependence is non-trivial, as all the co-ordinates of all the electrons are still simultaneously involved.

A standard way to address a many-body problem such as the one above can be invoked at this stage. The *independent electrons approximation* is obtained by considering the electrons as moving independently in an *effective* potential, as yet unspecified. The wave function $\Psi(\{\boldsymbol{r}_i, \sigma_i\})$ for N particles decouples into N one-particle wave functions, and consequently, equation (3) with the Hamiltonian (2) is replaced by N equations of the type

$$\widehat{\mathrm{H}}_{\mathrm{eff}}\ \Psi_i^{\sigma}(\boldsymbol{r}) = \epsilon_i^{\sigma}\ \Psi_i^{\sigma}(\boldsymbol{r}) \quad \text{with } \widehat{\mathrm{H}}_{\mathrm{eff}} = -\frac{\nabla^2}{2} + V_{\mathrm{eff}}^{\sigma}(\boldsymbol{r}) \tag{4}$$

Finding an effective potential that maps as accurately as possible the N independent particles problem (4) into the N many-body problem (2) is a real challenge. A way to tackle the many-body problem via the electron density $n(\boldsymbol{r}) = \sum_i \delta(\boldsymbol{r} - \boldsymbol{r}_i)$, and its exact reformulation as an independent particles problem, were provided respectively by Hohenberg and Kohnin 1964, and Kohn and Sham in 1965.

The Hohenberg and Kohn theorems (1964) apply to any interacting particles in an external potential. This includes the system of electrons with Coulomb interactions $\widehat{\mathrm{V}}_{\mathrm{ee}}$ in the electrostatic field of the nuclei $\widehat{\mathrm{V}}_{\mathrm{ne}}$ of the full many-body problem (2). The first theorem states that *all properties* of the interacting system are completely determined by the *electronic ground state density* $n(\boldsymbol{r})$, and that the total energy of the system E is a unique functional of the density: $E[n(\boldsymbol{r})]$. The second theorem concludes that the density which minimises the total energy $E[n(\boldsymbol{r})]$ is the exact ground state density, and similarly, that the exact ground state is the minimum of the functional.

The Hohenberg-Kohn theorems are wonderfully general but limited to establishing the *existence* of a universal functional $E[n(\boldsymbol{r})]$, which, minimised, gives the exact ground state. The theorems provide no pointers as to the form of $E[n(\boldsymbol{r})]$. The only straightforward case

is that of the external potential, the electrons-nuclei interaction. The functional $E_{\text{ext}}[n(\boldsymbol{r})]$ is directly given by

$$E_{\text{ext}}[n(\boldsymbol{r})] = \int \hat{V}_{\text{ext}}\, n(\boldsymbol{r})\, d\boldsymbol{r} \quad \text{with } \hat{\mathrm{V}}_{\text{ext}} = -\sum_{I}^{N} \frac{Z_I}{|\boldsymbol{r} - \boldsymbol{R}_I|} \tag{5}$$

A particular difficulty is raised on the other hand by the kinetic energy $\hat{\mathrm{T}}_{\mathrm{e}}$. No explicit relationship between the kinetic energy and the density is known. The next progress will be made by using an independent particles formulation, which allows for a formulation of the kinetic energy.

The essential step towards a practical DFT formalism was realised by Kohn and Sham (1965). They write the density in terms of non-interacting particles $n(\boldsymbol{r}) = \sum_{i,\sigma} |\Psi_i^\sigma|^2$. Although this density does represent the electron density, the individual single-particle wave functions Ψ_i^σ are not the wave functions of individual electrons, but that of mathematical objects. A quantity

$$T_s = -\frac{1}{2} \sum_{i,\sigma} \langle \Psi_i^\sigma | \nabla^2 | \Psi_i^\sigma \rangle \tag{6}$$

can be defined as the independent particles kinetic energy, which in general differs from the many-body energy of the operator $\hat{\mathrm{T}}_{\mathrm{e}}$. The Coulomb interaction can be turned into an expression of the density by defining

$$E_{\mathrm{H}}[n] = \frac{1}{2} \int \frac{n(\boldsymbol{r}_1)\, n(\boldsymbol{r}_2)}{|\boldsymbol{r}_1 - \boldsymbol{r}_2|} d\boldsymbol{r}_1 d\boldsymbol{r}_2 \tag{7}$$

the so-called Hartree energy. It too differs in principle from the full many-body version of the Coulomb interaction $E_{\text{ee}}[n]$. An exact formulation of the functional $E[n]$ for the electrons-nuclei problem can however be given as

$$E[n] = T_s[n] + E_{\mathrm{H}}[n] + E_{\text{ext}}[n] + E_{\text{xc}}[n] \tag{8}$$

with $E_{\text{xc}}[n]$, called the exchange-correlation functional, defined as

$$E_{\text{xc}}[n] = (T[n] - T_s[n]) + (E_{\text{ee}}[n] - E_{\mathrm{H}}[n]) \tag{9}$$

So the problem of mapping the non-interacting problem to that of a full many-body problem is delegated to the exchange-correlation functional. Assuming $E_{\text{xc}}[n]$ is known, the transformation (2) to (4) is now *exact*, with the effective Hamiltonian replaced by the Kohn-Sham Hamiltonian

$$\hat{\mathrm{H}}_{\mathrm{KS}} = -\frac{\nabla^2}{2} + \hat{\mathrm{V}}_{\text{ext}}(\boldsymbol{r}) + \int \frac{n(\boldsymbol{r}')}{|\boldsymbol{r} - \boldsymbol{r}'|} d\boldsymbol{r}' + \hat{\mathrm{V}}_{\text{xc}}(\boldsymbol{r}) \quad \text{with } \hat{\mathrm{V}}_{\text{xc}}(\boldsymbol{r}) = \frac{\partial E_{\text{xc}}[n(\boldsymbol{r})]}{\partial n(\boldsymbol{r})} \tag{10}$$

These are the Kohn-Sham equations, which can be solved by iteration. A starting density is calculated from a set of given wave functions. The density is then used to solve the Kohn-Sham equations, which in turn deliver a new set of wave functions. The procedure is repeated until self-consistency is achieved.

1.3. Choice of functionals

The Kohn-Sham equations (10) allowed DFT to become a practical approach for finding the ground state of interacting particles. Provided the functional $E_{\mathrm{xc}}[n]$ is known, the Kohn-Sham equations are an exact formulation of the many-body problem (2). In practice, the exchange-correlation functional is not known and must be approximated. Over the years, substantial developments have added different levels of sophistication to the exchange-correlation energy functional, usually at an added computational cost; the different levels are often quoted as the rungs of the Jacob's ladder (Perdew, Schmidt, 2001).

The notion of exchange and correlation is born from applying the condition of antisymmetry of the many-body wave function, in particular on the Coulomb term $\widehat{\mathrm{V}}_{\mathrm{ee}}$. The alternative to DFT for electronic structure calculations, the Hartree-Fock method, builds in the antisymmetry of the wave function by constructing the many-body wave function with Slater determinants of single-particle wave functions obeying a Hamiltonian of the type (4). The exchange contribution appears in that case as a term lowering the energy,

$$E_{\mathrm{x}} = -\sum_{i,j,\sigma} \frac{1}{2} \int \Psi_i^{\sigma *}(\boldsymbol{r}_1)\Psi_j^{\sigma *}(\boldsymbol{r}_2) \frac{1}{|\boldsymbol{r}_1 - \boldsymbol{r}_2|} \Psi_j^{\sigma}(\boldsymbol{r}_1)\Psi_i^{\sigma}(\boldsymbol{r}_2)\, d\boldsymbol{r}_1 d\boldsymbol{r}_2 \qquad (11)$$

the rest of the Coulomb term being the Hartree energy (7). The exchange term involves electrons of same spins. The concept of "exchange hole" is often introduced at this stage as describing the region around an electron that other electrons of the same spin avoid (Jones, Gunnarsson, 1989). Therefore, in Hartree-Fock the exchange is exact, but the many-body correlations are neglected. In DFT both correlations and exchange have to be taken into account in the exchange-correlation functional.

Local density approximation

Because the Hartree term already explicitly included in the Kohn-Sham equations (10) is a long-range interaction, it seems reasonable to assume that the missing contributions that must come into E_{xc} are of local nature. The local-density approximation (LDA) (Perdew, Zunger, 1981; Vosko, Wilk, Nusair, 1980) constructs $E_{\mathrm{xc}}^{\mathrm{LDA}}[n(\boldsymbol{r})]$ from that of the *homogeneous electron gas* having the same density. In the latter case the correlation and exchange are local and calculable typically by quantum Monte Carlo simulations (Ceperley, Alder, 1980). This seemingly too simple approach has proven surprisingly successful for a wide range of systems. The success of the LDA has been explained (Jones, Gunnarsson, 1989) by three factors: the sum rules related to the "exchange hole" are conserved in the LDA formulation; $\widehat{\mathrm{V}}_{\mathrm{ee}}$ depends only on the spherical average of the "exchange hole"; and the errors introduced in the separate exchange and correlation energy tend to cancel out. LDA is still widely used, although it is known to overbind and underestimate lattice parameters; cohesive energies and bulk modulus are found too large. Applied to pure iron, LDA's extension to spin polarised systems predicts a non-magnetic or anti-ferromagnetic ground state, instead of the ferromagnetic body centred cubic (bcc) structure (Wang, Klein, Krakauer, 1985).

Generalised gradient approximation

A step further is realised by the inclusion of the *gradient* of the density (or spin density) in the expression for the exchange-correlation functional $E_{\mathrm{xc}}^{\mathrm{GGA}}[n(\boldsymbol{r})], \nabla n(\boldsymbol{r})]$. The generalised gradient approximation (GGA), unlike LDA, must be parametrised and

sometimes deviates from strict first principles; a popular parametrisation used for *d*-electron materials like Fe and Fe-alloys as well as nuclear fuel materials like uranium dioxide (UO_2) is that of Perdew, Burke and Ernzerhof (PBE) (1996). An alternative is that of Perdew and Wang (PW91) (1992). Newer parametrisations (Hammer, Hansen, Norskov, 1999; Wu, Cohen, 2006) have been proposed recently but, to our knowledge, have not been widely used for nuclear materials. GGA is usually more accurate than LDA, although it has been found to soften bonds too much in some highly co-ordinate atoms. Both LDA and GGA are failing, however, to make correct predictions for strongly-correlated materials, like UO_2. In particular, for the latter compound, a Mott insulator, LDA and GGA wrongly predict a metallic ground state.

DFT+U

Strongly-correlated systems, like Mott insulators, are manifestations of a many-body physics that is not captured by local or semi-local exchange-correlation functionals. The 5*f* actinide oxides used as nuclear fuels can be tackled within DFT by the inclusion of an on-site Coulomb repulsion in the formalism. Different formulations exist (Dudarev et al., 1998; Lichtenstein, Anisimov, Zaanen, 1995). In essence, the DFT+U functionals (Lichtenstein, Anisimov, Zaanen, 1995) belong to the class of orbital-dependent functionals, whereby the localised orbitals (typically *d* or *f*) are shifted relative to other orbitals. The selective corrective term, called effective Hubbard U parameter, i.e. $U_{\text{eff}} = U - J$, is usually entirely empirical and is often adjusted in order to reproduce an experimentally known property, for instance the band gap of the studied material, although methods exist to calculate U as well (Anisimov, Aryasetiaman, Lichtenstein, 1997; Coconnioni, de Gironcoli, 2005). The DFT+U term introduces basically a penalty function that "discourages" a delocalisation of the *d* or *f* electrons, since it favours energetically either empty or fully occupied orbitals. Unfortunately, this creates a manifold of possible localisation patterns for the *f* electrons, which causes the occurrence of metastable states (Jollet et al., 2009; Dorado et al., 2009, 2013). This introduces technical problems with the convergence to the ground state, since the self-consistently converged solution for actinide materials using DFT+U depends on the initial *5f* orbital occupations. Several recipes have been devised in the literature to tackle this problem, e.g. the occupation matrix control (OMC) (Dorado et al., 2009), the U ramping method (Meredig, 2010), the quasi-annealing method (Geng et al., 2010), the controlled symmetry reduction method (Gryaznov, Heifets, Kotomin, 2012) and combined methods using *5f* occupation smearing with U ramping (FOUR) (Rabone, Krack, 2013).

Self-interaction correction

Another type of orbital-dependent functionals is covered by the label Self-Interaction Corrected (SIC). In Hartree-Fock the spurious $i = j$ term entering into the Hartree expression (7) is automatically exactly compensated by the exchange term (11). In DFT such a cancellation will have to occur via the exchange-correlation functionals. Formalisms whereby the self-interaction is subtracted orbital by orbital for all occupied orbitals have been developed (Svane, Gunnarsson, 1990; Svane et al., 1998). In effect, the self-interaction correction allows the localisation of states, particularly in rare-earth materials, and therefore it leads to an improved treatment of the magnetic ground state.

Dynamical mean field theory

Dynamical mean field theory (DMFT) (Georges, 1996) is often cited as the most successful treatment to date of strongly-correlated systems. Although it does include elements from

band structure, DMFT is not as such a DFT approach. The self-energy is determined self-consistently within the Green's function formalism via the so-called Anderson single-impurity problem, the latter requiring a sophisticated computational apparatus in order to be solved. DMFT is therefore a powerful, but computationally very expensive approach. First applications of DMFT variants like DFT+DMFT to nuclear fuel materials like uranium and plutonium oxides as well as their mixed oxides (MOX) are being performed, but the model system sizes are still limited to a few tenth atoms (Amadon, Condens, 2012).

Hybrid functionals

Hybrid functionals have been proposed as a mean to combine the exact exchange property of Hartree-Fock and the correlations from GGA. This type of approach is useful when the interest lies in defect energies located in the band gap of a semi-conductor. Hybrid functionals are not entirely first-principles, as a mixing parameter has to be defined, and are not easily evaluated for solids, since an explicit (orbital based) calculation of the Hartree-Fock exchange is required which is computationally expensive. The electronic properties of nuclear fuel materials like UO_2 (Kudin, Sciseria, Martin, 2002) as well as of other actinide oxides (Prodan, Scuseria, Matin, 2007) are better described with respect to plain DFT methods, but hybrid functionals share the convergence problems with DFT+U methods due to metastable states while being computational much more demanding, making the exploration of many occupation pattern computationally very expensive.

Meta-generalised gradient approximation functionals

Various other exchange-correlation functionals are available, although not popular within the nuclear materials community. *Meta-GGA* (Tao et al., 2003) comes in the third position after LDA and GGA in the Jacob's ladder as it includes a non-empirical dependence on the orbital kinetic energy density, i.e. terms of order $\nabla^2 n(\boldsymbol{r})$. It has been successfully tested on molecules, hydrogen-bonded complexes, and ionic solids (Tao et al., 2003). Despite the further order and the extra cost, this approximation remains semi-local in nature.

Van der Waals interactions are not included in any of the local or semi-local treatments, but can be considered in an empirical manner using the DFT-D method (Grimme et al., 2010), which provides a dispersion correction for the elements H to Pu for popular exchange and correlations functionals.

Incorporation of a random phase approximation (RPA) approach (Amadon, Applencourt, Bruneval, 2014) could account for such long-range interactions straightforwardly, even though at very high computational cost. Van der Waals forces are usually not relevant in the nuclear context.

As a conclusion, we would like to emphasise that physical properties may depend strongly on the choice of functionals: comparison with experiments for the property under consideration should dictate the choice of functionals.

1.4. Density functional theory implementations

Many different implementations of the resolution of the Kohn-Sham equations (10) have been made available to users. They differ most predominantly in the choice of basis sets for the single-particle wave functions, their different use or not of pseudopotentials, and whether the system is periodic or a cluster of atom.

The one-particle wave functions must be expanded in a basis set in order to proceed with the solution of the Kohn-Sham equation. The prototype choice is between localised sets of orbitals, or a plane waves basis set. In theory, the dimension of the function space is infinite. In practice, however, the number of basis functions must be limited, of course. A localised basis set will have basis functions that already look like the objects they will eventually construct, i.e. have large weights around the atomic positions. These basis sets are biased but computationally cheap, since very few basis functions will be needed to represent the electronic density, if the functions are well chosen. Plane waves, on the other hand, are free of any bias and fully flexible. The convergence with an increasing number of plane waves is also straightforward. However, many plane waves are needed to reconstruct the density around the atom. The situation is worse near to the nuclei, where strong variations occur. It is also not unusual to combine atomic-like functions in some regions of space with a plane waves basis set.

It is a general practice to treat core electrons, semi-core electrons and valence electrons separately. When a plane waves basis is chosen, the potentials of the core electrons are smoothed out, thereby reducing the oscillations and allowing the number of plane waves to remain manageable. These so-called *pseudopotentials* therefore introduce another level of approximation and inaccuracy. They come in different types (norm-conserving, ultra-soft, etc.) and must be tested for transferability in all possible situations.

A powerful approach combining the use of pseudopotentials and all-electron accuracy is the *Projector augmented-wave* (PAW) technique (Blöchl, 1944). By means of projectors, the all-electron (i.e. not pseudised) wave functions of the valence electrons are reconstructed. This is the method of choice nowadays in many cases. It has been applied for Fe and alloys as well as actinide materials, where the computationally cheaper ultra-soft pseudopotentials are not accurate enough (Kresse, Furthmüller, 1996).

DFT codes also differ in their use of periodic boundary conditions or cluster approach. In the latter case, a cluster of atoms, possibly surrounded by compensating charges, is usually treated in a localised basis set. In the former case a unit cell is given, which makes the approach naturally suited for the study of crystal structures. The periodic wave functions in a crystal can be expanded in plane waves whose wave vectors are the reciprocal lattice vectors of the crystal. Plane waves are therefore a natural choice of basis for periodic structures, although many codes very successfully use also local orbital approaches in this case.

Periodic systems have a real space (infinite number of electrons) and a reciprocal space representation (infinite number of *k*-points), and codes rely on both representations. Numerical scheme exists that limit the number of *k*-points that are necessary; metals in particularly require a dense *k*-points coverage.

Some of the most popular codes used by the nuclear materials community are briefly described in the following without any claim for completeness:

WIEN2k (Blaha et al., 1990; Schwarz, Blaha, 2003) uses a combination of atomic-like functions in the regions around the atoms and plane waves for the interstitial regions. The full description of the basis set is “(L)APW+lo”, for (Linearised) Augmented Plane Wave plus Localised Orbitals (Cottenier, 2013). This approach is a “full potential”, computationally demanding all-electron approach that does not call on pseudopotentials, and is thereby one of the most accurate ones available. This can be crucial, if properties sensitive to the wave functions near the core are required, like spin-orbit coupling

(Laskowski et al., 2004). It is often used for benchmarking other DFT codes using pseudopotentials or PAWs (Lejaeghere et al, 2016).

VASP (Kresse, Furthmüller, 1996) is a plane waves code with a choice of ultra-soft pseudopotentials, and also a PAW implementation (Kresse, Joubert, 1999). The latter has been a popular choice for Fe-alloys. VASP is probably the most popular DFT code for the modelling of nuclear materials.

ABINIT (Gonze et al., 2005) is, by contrast to VASP, a freely available code which offers comparable capabilities.

Castep (Clark et al., 2005) is a pseudopotential plane waves code which includes a module for the *ab-initio* determination of Nuclear Magnetic Resonance parameters (Pickard, Mauri, 2001).

SIESTA (Soler et al., 2005) is a pseudopotential method using atom-centred basis sets in periodic systems. Some caution is needed when using the code's ultra-soft pseudopotentials in iron magnetic systems. The efficient basis sets, however, allow for calculations of relatively large systems compared to other approaches.

Dmol^3 (Delley, 1990) uses numerical atomic orbitals as basis sets. The code has both all-electron and pseudopotential capabilities, and an implementation of the scalar relativistic all-electron approach.

CP2K is a freely available code employing a hybrid basis set of Gaussian functions and plane waves basis (GPW) (Lippert, Hutter, Parrinello, 1997). It provides also a Gaussian Augmented Plane Waves (GAPW) implementation (Lippert, Hutter, Parrinello, 1999) which allows for all-electron calculations (Krack, Parrinello, 2000). The hybrid basis set approach enables a Kohn-Sham calculation that scales linearly with the system size (Van de Vondele et al., 2005), like for the SIESTA code. CP2K has been applied for nuclear fuel materials like UO_2 (Krack, 2015), but it provides also, besides the DFT module CP2K/QUICKSTEP, (Van de Vondele et al., 2005) the force field implementation CP2K/FIST (Bertolus et al., 2015).

The accuracy of the results in general will depend on the choice of functionals, and the suitability of the PAW or pseudopotentials used. Moreover, the number of plane waves, or the choice of localised basis functions, must be tested thoroughly, as well as the density of *k*-points, and supercell size when relevant. Unconverged results are totally meaningless, irrespective of the accuracy of the method. One should carefully differentiate between deficiencies of the selected DFT approach and shortcomings of the employed code and the corresponding input parameter settings.

1.5. Scope of density functional theory

DFT determines the ground state of a collection of atoms. As such, it is therefore a zero Kelvin formalism. To this day, the number of atoms considered is of the order of 1 000, with 200-300 typical for supercell alloys system. We describe in the following section only the properties most routinely determined in the contest of multi-scale modelling in nuclear materials.

The self-consistent result of the Kohn-Sham equations gives the density as a function of position and the total energy of the system. So the most direct application of DFT is the *total energy comparison* between different atomic arrangements. Energy differences

converge faster with respect to the basis set quality than absolute energies and in some cases benefit from cancellations of systematic errors.

The determination of forces on the atoms, formally the variation of the energy with respect to the nuclei positions, is implemented in most codes. Most will also have an algorithm whereby positions of the nuclei are modified at constant volume until the forces on the atoms are as small as desired. This *ionic relaxation* will usually drive the system into a local minimum, with no guarantee of having reached a global minimum. When starting from a reasonable geometry and using an appropriate functional, bond lengths comparisons with experiments are nonetheless accurate to the order of the percent.

Full structure optimisations, including *volume relaxation,* are also routinely implemented. These algorithms are based on the expression of the macroscopic stress tensor in the DFT formalism. Volume relaxation, however, has to be tackled carefully. The *k*-points integration accuracy varies with volume changes, so that the density of *k*-points has to be sufficient across the volume range. Incomplete basis sets introduce an error in the stress tensor, called "Pulay stress". In the case of plane waves, varying the volume will effectively change discontinuously the quality of the basis set. These problems can be reduced by using a sufficiently large basis set. It is often recommended to perform first a constant volume cell shape relaxation, then to fit an equation of state through the total energies at different volumes. That fit also gives the *bulk modulus*. Comparisons with experiments show that the lattice parameters can agree with the experimental values to the order of the percent.

Elastic constants can also be deduced from DFT. They are calculated from the stress tensors of several structure optimisations performed under well-chosen strains, which will depend on the symmetry of the crystal. Linear elastic constants can be determined within 10% of the experimental values.

The *interatomic force constants*, formally the derivatives of the force with respect to nuclei positions, are not directly measured quantities. They are often determined in the so-called "frozen phonon" method, whereby the total energy versus displacements is calculated.

Relativistic effects are not explicitly taken into account in Hamiltonian (2), nor are the spin-orbit interactions related to such effects. The relativistic effects are limited to regions near the core, and can be taken into account at the level of the pseudopotentials. In all-electron codes the core electrons are usually treated fully relativistic using scalar-relativistic wave functions, and the spin-orbit coupling is calculated for valence electrons in a second-variational procedure.

1.6. The use of density functional theory in nuclear materials research

The role of DFT in the multi-scale approach is two-fold: to generate data for MD empirical potential parametrisation, and to investigate structures and provide measurable quantities for model validation.

MD empirical potentials are parametrised and adjusted on equilibrium properties obtained by DFT. These are typically equilibrium bond lengths, elastic constants and cohesive energy per atom. For the particular purpose of nuclear materials, point defect formation energies should also be calculated. This is done usually by the {\it supercell} technique. The defects are placed at the centre of a large supercell built from repeated blocks of the original unit cell. The size of the supercell must be large enough that the defects are effectively isolated from their images across the periodic boundaries.

DFT is also the indicated method for the investigations of electronic, structural and magnetic properties. Partial density of states is used to analyse the effect of impurities (Klaver, Drautz, Finnis, 2006). Spin-dependent calculations are also used to identify phase stability in magnetic materials and structural aspects like clustering of impurities (Olsson, Klaver, Domain, 2010).

As hinted in the previous paragraph, some of the information delivered by DFT can be directly compared with experiments, like the lattice parameters. Others, like the interaction energies, can only, if at all, be compared to experiments via related measurable quantities. Structural information, for instance, can be obtained through spectroscopic measurement (Idhil et al., 2012). The migration energy along particular paths can be calculated by DFT (Domain, Becquart, Foct, 2004) and these values can be used as parameters for methods at larger length scales. Resistivity recovery measurements can give access to that information, but there are difficulties in assigning defects to particular binding energies. In many cases DFT and experiments are therefore part of the same model validation loop.

1.7. Conclusions and outlook

DFT allows the calculations of ground state properties from first principles. The applicability of this powerful approach is, however, limited to a relatively small number of atoms. The accuracy of DFT, within its own scope, is determined by that of the exchange-correlation functional. No functional captures correctly all the properties usually required. In particular, the strongly-correlated effects have to be included as an extension to the DFT. Convergence of the results with respect to the controllable parameters of the calculations (size or type of the basis set, *k*-points sampling, and supercell size) should be done carefully. It is also essential to remember that ionic and volume relaxation may not express a global minimum. Moreover, the convergence to the ground state of magnetic systems is often extremely slow depending on the algorithm implemented. Nevertheless, it may still lead to the wrong magnetic state.

Therefore, the accuracy of the DFT-determined values should always be evaluated critically, and should always be related to experimental measurements if possible. It may be worthwhile investigating the benefit versus computational cost of the newer exchange-correlation functionals for some compounds, although the problem of strongly-correlated systems will likely remain restricted to empirical treatments in the near future.

It is likely that the accuracy of molecular dynamics simulations depends more on the form of the potentials than on the absolute convergence and accuracy of the DFT parameters. However, in evaluating the quality of empirical potentials one should remember the limitations of the DFT on which they are based.

References

Amadon, B. (2012), "A self-consistent DFT + DMFT scheme in the projector augmented wave method: applications to cerium, Ce_2O_3 and Pu_2O_3 with the Hubbard I solver and comparison to DFT + *U*, *J. Phys.: Condens. Matter*, Vol. 24, 075604.

Amadon, B., T. Applencourt and F. Bruneval (2014), "Screened Coulomb interaction calculations: cRPA implementation and applications to dynamical screening and self-consistency in uranium dioxide and cerium", *Phys. Rev. B*, Vol. 89, 125110.

Anisimov, V.I., F. Aryasetiawan and A.I. Lichtenstein (1997), "First-principles calculations of the electronic structure and spectra of strongly correlated systems: the LDA+ *U* method", *J. Phys.: Condens. Matter*, Vol. 9, 767.

Bertolus, M. et al. (2015), "Linking atomic and mesoscopic scales for the modelling of the transport properties of uranium dioxide under irradiation", *J. Nucl. Mater.,* Vol. 462, 475.

Blaha, P. et al (1990), "Full-potential, linearized augmented plane wave programs for crystalline systems", *Comput. Phys. Commun.*, Vol. 59, 399.

Blöchl, P.E. (1994), "Projector augmented-wave method", *Phys. Rev. B*, Vol. 50, 17953.

Born, M. and J.R. Oppenheimer (1927), "On the Quantum Theory of Molecules", *Ann. Phys.*, Vol. 389, 457.

Ceperley, D.M. and B.J. Alder (1980), "Ground State of the Electron Gas by a Stochastic Method", *Phys. Rev. Lett.*, Vol. 45, 566.

Clark, S.J. et al. (2005), "First principles methods in using CASTEP", *Z. Kristallogr.* Vol. 220, 567.

Cococcioni, M. and S. de Gironcoli. S. (2005), Linear response approach to the calculation of the effective interactive parameters in the LDA + *U* method", *Phys. Rev. B*, Vol. 71, 035105.

Cottenier, S. (2013), *Density Functional Theory and the family of (L)APW-methods: a step-by-step introduction*, Ghent University, Belgium, ISBN 978-90-807215-1-7

Delley, B. (1990), "An all-electron numerical method for solving the local density functional for polyatomic molecules", *J. Chem. Phys.*, Vol. 92, 508.

Domain, C., C.S. Becquart and J. Foct (2004), "*Ab initio* study of foreign interstitial atom (C,N) interactions with intrinstic point defects in α–Fe", *Phys. Rev. B,* Vol. 69.

Dorado, B. et al. (2009), "DFT + U calculations of the ground state and metastable states of uranium dioxide", *Phys. Rev. B*, Vol. 79, 235125.

Dorado, B. et al (2013), "Advances in first-principles modelling of point defects in UO_2: f electron correlatios and the issue of local energy minima", *J. Phys.: Condens. Matter,* Vol. 25, 333201.

Dudarev, S.L. et al. (1998), "Electron-energy-loss spectra and the structural stability of nickel oxide: an LSDA+U study", *Phys. Rev. B*, Vol. 57, 1505.

Georges, A. et al (1996), "Dynamical mean-field theory of strongly correlated fermion systems and the limit of infinite dimensions", *Rev. Mod. Phys.*, Vol. 68, 13.

Geng, H.Y. et al (2010) "Interplay of defect cluster and the stability of xenon in uranium dioxide from density functional calculations", *Phys. Rev. B*, Vol. 82, 094106.

Gonze, X. et al. (2005), "A brief introduction to the ABINIT software package", *Z. Kristallogr.*, Vol. 220, 558.

Grimme, S. et al (2010), "A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu", *Chem. Phys.*, Vol. 132, 154104.

Gryaznov, D., E. Heifets and E. Kotomin (2012), "The first-principles treatment of the electron-correlation and spin-orbital effects in uranium mononitride nuclear fuels", *Phys. Chem. Chem. Phys.*, Vol. 14, 4482.

Hammer, B., L.B. Hansen and J.K. Norskov (1999), "Improved adsoption energetics within density-functional theory using revised Perdew-Burke-Ernzerhod functional", *Phys. Rev. B*, Vol. 59, 7413.

Hohenberg, P. and W. Kohn (1964), "Inhomogeneous Electron Gas", *Phys. Rev.*, Vol. 136, B864.

Idhil, A. et al. (2012), "The influence of Cr-composition on the local magnetic structure of FeCr alloys", *Nucl. Instrum. Methods Phys. Res. Sect B*, Vol. 284, 1.

Jollet, F. et al (2009), "Hybrid functional for correlated electrons in the prpjector augmented-wave formalism: Study of multiple minima for actinide oxides", *Phys. Rev. B*, Vol. 80, 235109.

Jones, R.O. and O. Gunnarsson (1989), "The density functional formalism, its applications and prospects", *Rev. Mod. Phys.*, Vol. 61, 689.

Klaver, T.P.C, R. Drautz and M.W. Finnis (2006), "Magnetism and thermodynamics of defect-free Fe-Cr alloys", *Phys. Rev. B,* Vol. 74, 094435.

Kohn, W. and L. J. Sham (1965), "Self-Consistent Equations Including Exchange and Correlation Effects", *Phys. Rev.*, Vol. 140, A1133.

Krack, M. (2015), "On the ground state electronic structure of uranium dioxide", *Phys. Scr.* Vol. 90, 094014.

Krack, M. and M. Parrinello (2000), "All-electron ab-initio molecular dynamics", *Phys. Chem. Chem. Phys.* Vol. 2, 2105.

Kresse, G. and D. Joubert (1999), "From ultrasoft pseudopotentials to the projector augmented-wave method", *Phys. Rev. B*, Vol. 59, 1758.

Kresse, G. and J. Furthmüller (1996), "Efficient iterative schemes for *ab initio* total-energy calculations using a plane-wave basis set", *Phys. Rev. B*, Vol. 54, 11169.

Kudin, K.N., G.E. Scuseria and R.L. Martin (2002), "Hybrid Density-Functional Theory and the Insulating Gap of UO_2", *Phys. Rev. Lett.*, Vol. 89, 266402.

Laskowski, R. et al (2004), "Magentic structure and electric-field gradients of uranium dioxide: An *ab initio* study", *Phys. Rev. B*, Vol. 69, 40408.

Lejaeghere, K. et al. (2016), "Reproducibility in density functional theory calculations of solids", *Science*, Vol. 351, aad3000.

Lichtenstein, A.I., V.I. Anisimov and J. Zaanen (1995), "Density-functional theory and strong interaction: Orbital ordering in Mott-Hubbard insulators", *Phys. Rev. B*, Vol. 52, R5467.

Lippert, G., J. Hutter and M. Parrinello (1997), "A hybrid Gaussian and plane wave density function scheme", *Mol. Phys.*, Vol. 92, 477.

Lippert, G. J. Hutter and M. Parrinello (1999), "The Gaussian and augmented-plane-wave density functional method for ab intio molecular dynamics simulation", *Theor. Chem. Acc.*, Vol. 103, 124.

Martin, R.M. (2004), *Electronic Structure: Basis theory and practical methods*, Cambridge University Press, Cambridge.

Meredig, B. et al (2010), "Method for location low-energy solutions within DFT + *U*", *Phys. Rev. B*, Vol. 82, 195128.

Olsson, P., T.P.C. Klaver and C. Domain (2010), "*Ab initio* study of solute transition-metal interactions with point defects in bcc Fe", *Phys. Rev. B,* Vol. 81.

Parr, R.G. and W. Yang (1989), *Density-Functional Theory of Atoms and Molecules*, Oxford University Press, Oxford.

Perdew, J.P. and A. Zunger (1981), "Self-interaction correction to density-functional approximations for many-electon systems", *Phys. Rev. B*, Vol. 23, 5048.

Perdew, J.P., K. Burke and M. Ernzerhof (1996), "Generalized Gradient Approximation Made Simple", *Phys. Rev. Lett.*, Vol. 77, 3865.

Perdew, J.P. and K. Schmidt (2001), *AIP Conf. Proceedings*, 577(1).

Perdew, J.P. and Y. Wang (1992), "Accurate and simple analytic representation of the electron-gas correlation energy", *Phys. Rev. B*, Vol. 45, 13244.

Pickard, C.J. and F. Mauri (2001), "All-electron magnetic response with pseudopotentials: NMR chemicals shifts", *Phys. Rev. B*, Vol. 63.

Prodan, I.D., G.E. Scuseria and R.L. Martin (2007), "Convalency in the actinide dioxides: Systematic study of the electronic properties using screened hybrid density functional theory", *Phys. Rev. B*, Vol. 76, 033101.

Rabone, J. and M. Krack (2013), "A procedure for bypassing metastable states in local basis set DFT+ U calculations and its application to uranium dioxide surfaces" *Comput. Mater. Sci.*, Vol. 71, 157.

Schwarz, K. and P. Blaha (2003), "Solid state calculation using WIEN2K", *Comput. Mat. Sci.,* Vol. 28, 259.

Soler, J.M. et al. (2002), "The SIESTA method for *ab initio* order-*N* materials simulation", *J. Phys.: Condens. Matter*, Vol. 14, 2745.

Svane, A. et al (1998), "Electronic structure of cerium monopnictides under pressure", *J. Phys.: Condens. Matter*, Vol. 10, 5309.

Svane, A. and O. Gunnarsson (1990), "Transition-metal oxides in the self-interaction – correction density-function formalism", *Phys. Rev. Lett.,* Vol. 65, 1148.

Tao, J. et al (2003), "Climbing the Density Functional Ladder: Nonempirical Meta-Generalized Gradient Approximation Designed for Molecules and Solids", *Phys. Rev. Lett.*, Vol. 91, 146401.

VandeVondele, J. et al. (2005), "QUICKSTEP: Fast and accurate density functional calculations using a mixed Gassian and plane waves approach", *Comput. Phys. Commun.*, Vol. 167, 103.

Vosko, S.H., L. Wilk and M. Nusair (1980), "Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis", *Can. J. Phys*., Vol. 58, 1200.

Wang, C.S., B.M. Klein and H. Krakauer (1985), "Theory of Magnetic and Structural Ordering in Iron", *Phys. Rev. Lett.*, Vol. 54, 1852.

Wu, Z. and R.E. Cohen (2006), "More accurate generalized gradient approximation for solids, *Phys. Rev. B*, Vol. 73, 235116.

2. *Ab-Initio* molecular dynamics

Thomas D. Kühne
Department of Chemistry and Paderborn Center for Parallel Computing, University of Paderborn, Germany

2.1. Introduction

Computer simulations and molecular dynamics in particular, is a very powerful method to provide detailed and essentially exact information of classical many-body problems. With the advent of *ab-initio* molecular dynamics (AIMD), where the forces are computed on-the-fly by accurate electronic structure calculations, the scope of either method has been greatly extended. This new approach, which unifies Newton's and Schrödinger's equations, allows for complex simulations without relying on any adjustable parameter. This review is intended to outline the basic principles as well as a survey of the field. Beginning with the derivation of Born-Oppenheimer molecular dynamics, the Car-Parrinello method as well as novel hybrid scheme that unifies best of either approach are discussed. The predictive power is demonstrated by a series of applications ranging from insulators to semiconductors and even metals in condensed phases.

The geometric increase in performance of computers over last few decades, together with advances in applied physics and mathematics, has led to the birth of a new way of doing science that is in the intersection of theory and experiment. As a result, they are referred to as computational sciences and allow for computer experiments under perfectly controllable and reproducible conditions. In this way, computer simulations have been very successful in explaining a large variety of physical phenomena and guiding experimental work. In addition, it is even possible to predict new phenomena by conducting experiments *in silico* that would otherwise be too difficult, too expensive, or simply impossible to perform. However, by far the most rewarding outcome of computer simulations is the invaluable insight they provide into the behaviour and the dynamics of a system. The two most common algorithms for such studies are the Monte Carlo (MC) (Metropolis et al., 1953) and molecular dynamics (MD) (Alder, Wainwright, 1957; Rahman, 1964; Verlet, 1967] algorithm. The latter is simply the numerical solution of Newton's equation of motion, which allows both equilibrium thermodynamic and dynamical properties of a system at finite temperature to be computed. Since it also provides a 'window' onto the atomic real-time evolution of the atoms, another role of MD is that of a computational microscope.

One of the most challenging, but very important aspect of MD simulations is the calculation of the interatomic forces. In classical simulations, they are computed from empirical potential functions, which have been parametrised to reproduce experimental or accurate *ab-initio* data on small model systems. Even though great strides in elaborating these empirical potentials have been made, often the transferability to systems or regions of the phase diagram different from the ones to which they have been fitted is restricted. Furthermore, they are not able to simulate with sufficient predictive power chemical bonding processes that take place in many relevant systems. Eventually, some of the most important and interesting phenomena of modern physics and chemistry are intrinsically non-classical. Therefore, a first principle based approach, such as AIMD (Marx, Hutter, 2009), where the forces are calculated on-the-fly from accurate electronic structure calculations, is very attractive since many of these limitations can in principle be removed.

However, the increased accuracy and predictive power of AIMD simulations comes at significant computational cost. For this reason, density functional theory (DFT) (Hohenberg, Kohn, 1964) is to date by far the most commonly employed electronic structure theory, but it is important to note that AIMD is a general concept that in principle can be used in conjunction with any electronic structure method. Nevertheless, the *ab-initio* approach is not without problems – the relevant energy scale is tiny, well below $k_{\mathrm{B}}T$, and in particular the attainable length and time scales are still one of its major limitations.

2.2. Molecular dynamics

The mathematical task of MD is to evaluate the expectation value $\langle \mathcal{O} \rangle$ of an arbitrary operator $\mathcal{O}(\boldsymbol{R}, \boldsymbol{P})$ with respect to the configurational Boltzmann distribution

$$\langle \mathcal{O} \rangle = \frac{\int d\boldsymbol{R}\, d\boldsymbol{P}\, \mathcal{O}(\boldsymbol{R}, \boldsymbol{P})\, e^{-\beta E(\boldsymbol{R},\boldsymbol{P})}}{\int d\boldsymbol{R}\, d\boldsymbol{P}\, e^{-\beta E(\boldsymbol{R},\boldsymbol{P})}} \tag{1}$$

where $\beta = 1/k_{\mathrm{B}}\,T$ is the inverse temperature. The total energy function

$$E(\boldsymbol{R}, \boldsymbol{P}) = \sum_{I=1}^{N_{\mathrm{ion}}} \frac{P_I^2}{2M_I} + \Phi(\boldsymbol{R}_I) \tag{2}$$

where the first term denotes the nuclear kinetic energy, $\Phi(\boldsymbol{R}_I)$ the potential energy function, N_{ion} the number of ions and M_I the corresponding masses, depends itself on nuclear positions $\boldsymbol{R}$ and momenta $\boldsymbol{P}$.

One way to evaluate equation (1), at least in principle, is to directly solve such a high-dimensional integral, whose integrand is very sharply peaked in many dimensions, by a uniform sampling using the MC algorithm. However, such an MC algorithm is very inefficient, if it would not be for importance sampling (Metropolis et al., 1953), which satisfies the sufficient detailed balance condition by rejections.

On the other hand, assuming the ergodicity hypothesis, the thermal average $\langle \mathcal{O} \rangle$ can not only be determined as the ensemble average of a MC simulation, but using MD equally as a temporal average

$$\langle \mathcal{O} \rangle = \lim_{\tau \to \infty} \frac{1}{\tau} \int dt\, \mathcal{O}(\boldsymbol{R}(t), \boldsymbol{P}(t)) \tag{3}$$

However, by propagating the classical many-body system in time according to Newton's equation of motion, the ions are treated only classically; this approximation is usually negligible, except for very light atoms or low temperature, where nuclear quantum effects may be important and adopting a quantum formalism such as imaginary-time path integrals (Feynman, Hibbs, 1965; Ceperley, 1995) is required.

Similar to MC, within MD some kind of importance sampling is naturally performed by preferentially visiting phase space of low potential energy. Furthermore, as denoted by the additional time dependence in equation (3), MD allows for additional insights from the ionic real-time evolution, at least in a statistical average sense. It is neither the intention, nor possible, to obtain "exact" trajectories by MD due to the infamous Lyapunov instability, which states that slightly perturbed trajectories are intrinsically exponentially diverging with time.

The equipartition theorem

$$\left\langle \frac{1}{2} M_I \dot{\boldsymbol{R}}_I^2 \right\rangle = \frac{3}{2} k_{\mathrm{B}} T \tag{4}$$

where M_I is the atomic mass, k_{B} the Boltzmann constant and T the instantaneous temperature, offers an elegant way to bridge the gap between molecular mechanics and thermodynamics. This opens the door to extract a vast variety of relevant static and dynamic, as well as transport properties from a MD simulation.

Nevertheless, any computational resource is finite, which limits the time and length scales accessible by computer simulations. One way to partially bridge the gap between the microscopic size of the simulated system and the macroscopic reality is to introduce periodic boundary conditions. In this way surface effects are eliminated by effectively simulating an infinite system, albeit with a finite periodicity that is identical with the length L of the simulation cell. As a consequence, only phenomena whose characteristic correlation length is much smaller than L can be simulated. By similar means only processes whose typical relaxation time is significantly smaller than the simulation time τ can be studied. Even though great strides have been made to extend the time and length scales, it is apparent that techniques such as those reviewed here are clearly needed.

2.3. An *ab-initio* potential

In AIMD the forces $\boldsymbol{F}_I = -\nabla_{\boldsymbol{R}_I} \Phi(\boldsymbol{R}_I)$ are determined on-the-fly using first-principles electronic structure methods. That means that AIMD is not relying on any adjustable parameter, but only on $\boldsymbol{R}_I$, which constitutes its predictive power. However, finding the antisymmetric ground state eigenfunctions ψ_0 of the corresponding many-body Hamiltonian at each MD step comes at a significant computational cost, which has to be carefully balanced against the size and sampling requirements of MD.

2.3.1. The many-body Schrödinger equation

Applying the so-called Born-Oppenheimer (BO) approximation (1927), which we have implicitly assumed in the preceding subsection, $\Phi(\boldsymbol{R}_I)$ can written as

$$\Phi(\boldsymbol{R}_I) = \langle \psi_0 | \mathcal{H}_e(\{\boldsymbol{r}_i\}; \boldsymbol{R}_I) | \psi_0 \rangle + E_{II}(\boldsymbol{R}_I) \tag{5}$$

where $\mathcal{H}_e(\{\boldsymbol{r}_i\}; \boldsymbol{R}_i)$ is the electronic many-body Hamiltonian, that depend on the electronic co-ordinates $\{\boldsymbol{r}_i\}$ and parametrically on $\boldsymbol{R}_I$. Essentially, the BO approximation allows for a *product ansatz* of the total wavefunction consisting of the nuclear and electronic wavefunctions. Due to the large separation of the nuclear and electronic masses, the electrons can be expected to be in its instantaneous equilibrium with the much heavier nuclei, so that the electronic subsystem can be treated independently at constant $\boldsymbol{R}_I$. Nevertheless, we are left with a formidable task to solve the electronic, non-relativistic, time-independent, many-body Schrödinger equation

$$\mathcal{H}_e(\{\boldsymbol{r}_i\}; \boldsymbol{R}_i)\, \psi_0(\{\boldsymbol{r}_i\}) = \varepsilon_0(\boldsymbol{R}_I)\, \psi_0(\{\boldsymbol{r}_i\}) \tag{6}$$

which is a high-dimensional, non-linear eigenvalue problem, with eigenfunctions $\psi_0(\{\boldsymbol{r}_i\})$ and eigenvalues $\varepsilon_0(\boldsymbol{R}_I)$, respectively. To visualise the complexity of equation (6) let us consider the following *Gedankenmodell* to represent the solution $\psi_0(\{\boldsymbol{r}_i\})$ on a real-space grid, where each co-ordinate is discretised by 100 mesh points. Ignoring spin and taking

$\psi_0(\{\boldsymbol{r}_i\})$ to be real, for N_e electrons 10^{6N_e} grid points are required, so that the solution of a single Si atom would require more grid points than the number of electrons in the whole universe, not to mention solving such a large non-linear eigenvalue problem.

2.3.2. Density functional theory

Fortunately, this curse of dimensionality can be ingeniously bypassed by the use of DFT, which is based on two celebrated papers of Hohenberg, Kohn and Sham (1964; 1965). The former, the so-called Hohenberg-Kohn (HK) theorem, proves the existence of a one-to-one mapping between the ground state density $\rho_0(\boldsymbol{r})$ and an external potential $v(\boldsymbol{r})$. In this vain, $\rho_0(\boldsymbol{r})$, which depends on just three electronic degrees of freedom, is designated as the principal quantity rather than the $3N_e$-dimensional many-body wavefunction. As a consequence, the nondegenerate ground state wavefunction $\psi_0(\{\boldsymbol{r}_i\}) = \psi[\rho_0(\boldsymbol{r})]$ and likewise $\mathcal{H}_e[\rho_0(\boldsymbol{r})]$ are both unique functionals of $\rho_0(\boldsymbol{r})$, just as the ground state energy

$$E_0 = E^{\mathrm{DFT}}[\rho_0(\boldsymbol{r})] = \langle\psi[\rho_0(\boldsymbol{r})|\mathcal{H}_e[\rho_0(\boldsymbol{r})]|\psi[\rho_0(\boldsymbol{r})\rangle \tag{7}$$

The latter obeys the variational property

$$E^{\mathrm{DFT}}[\rho_0] = \langle\psi_0|\mathcal{H}_e|\psi_0\rangle \leq \langle\psi'|\mathcal{H}_e|\psi'\rangle = E^{\mathrm{DFT}}[\rho'] \tag{8}$$

for which equality holds if and only if $\rho_0 = \rho'$. As a consequence equation (6) can be solved not only by iteratively diagonalising $\mathcal{H}_e[\rho]$ within a self-consistent field (SCF) procedure, but equally by minimising the quantum expectation value of $\mathcal{H}_e[\rho]$:

$$E^{\mathrm{DFT}}[\rho_0] = \min_{\psi} \langle\psi|\mathcal{H}_e|\psi\rangle = \min_{\rho} \langle\psi[\rho]|\mathcal{H}_e[\rho]|\psi[\rho]\rangle = \min_{\rho} E^{\mathrm{DFT}}[\rho] \tag{9}$$

In principle the minimisation has to be performed under the constraint that $\rho(\boldsymbol{r})$ is N-representable, i.e. that it is arising from an antisymmetric N-body wavefunction $\psi(\{\boldsymbol{r}_i\})$. Luckily, this had been solved, and it can be demonstrated that any single-particle density can be written in terms of an antisymmetric many-body wavefunction (Gilbert, 1975; Harriman, 1981). On the contrary, for the v-representability problem, which states that $\rho(\boldsymbol{r})$ is the ground state density of a local potential $v(\boldsymbol{r})$, no such general solution is known. The HK theorem just guarantees that there cannot be more than one potential for each density, but does not exclude the possibility that there is no potential realising that density. It is only known for discretised systems that every density in interacting ensemble is v-representable. Interestingly, the constructive proof of Levy and Lieb (Levy, 1982; Leib, 1983) shows that for an interacting system v-representability is not required for the proof of the HK theorem.

For the sake of simplicity in the following I will throughout assume atomic units and confine myself to the physical relevant Coulomb system, for which

$$\mathcal{H}_e = \frac{1}{2}\sum_{i=1}^{N_e} \nabla_i^2 + \sum_{i<j} \frac{1}{|\boldsymbol{r}_i - \boldsymbol{r}_j|} + \sum_{I,i} \frac{Z_I}{|\boldsymbol{R}_I - \boldsymbol{r}_i|} = \hat{T} + \hat{U} + \hat{V} \tag{10}$$

Where $\hat{T}$ is the kinetic energy operator of the electrons, while $\hat{U}$ is the electron-electron interaction and $\hat{V} = \sum_i v(\boldsymbol{r}_i)$ the electron-ion operator. The former two operators are universal and independent of the system, while the latter is system dependent, or non-universal. DFT explicitly recognises that it is indeed the potential $v(\boldsymbol{r})$, which distinguishes non-relativistic Coulomb systems and offers a prescription how to deal with $\hat{T}$ and $\hat{U}$ once

and for all (Capelle, 2006). Hence, even at this stage based on nothing but the HK theorem, DFT is already of some practical use without having to solve the many-body Schrödinger equation and without having to make a single-particle approximation. In principle it should be even possible to calculate all observables, since the HK theorem guarantees that they are all functionals of $\rho(\boldsymbol{r})$. Presuming the availability of physical sound and accurate approximations one can write

$$E^{\mathrm{DFT}}[\rho(\boldsymbol{r})] = T[\rho(\boldsymbol{r})] + U[\rho(\boldsymbol{r})] + V[\rho(\boldsymbol{r})] \tag{11}$$

In the so-called Thomas-Fermi (TF) approximation (Thomas, 1927; Fermi, 1927) the full electron-electron interaction energy is approximated by the Hartree energy

$$U[\rho(\boldsymbol{r})] \approx U_H[\rho(\boldsymbol{r})] = \frac{1}{2}\int d\boldsymbol{r} \int d\boldsymbol{r}' \, \frac{\rho(\boldsymbol{r})\rho(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|} \tag{12}$$

that is the electrostatic interaction energy of $\rho(\boldsymbol{r})$. In addition, the kinetic energy is approximated as

$$T[\rho(\boldsymbol{r})] \approx \int d\boldsymbol{r}\, t^{\mathrm{hom}}[\rho(\boldsymbol{r})] = T^{\mathrm{LDA}}[\rho(\boldsymbol{r})] \tag{13}$$

where $t^{\mathrm{hom}}[\rho(\boldsymbol{r})]$ is the kinetic energy density of a homogeneous interacting system, which is also known as the local-density approximation (LDA). Due to the fact that the explicit form of $t^{\mathrm{hom}}[\rho(\boldsymbol{r})]$ is only known for a non-interacting system, $t^{\mathrm{hom}}[\rho(\boldsymbol{r})]$ is further estimated by the single-particle approximation $t^{\mathrm{hom}}[\rho(\boldsymbol{r})]$, i.e.

$$T^{\mathrm{LDA}}[\rho(\boldsymbol{r})] \approx \int dr\, t_s^{\mathrm{hom}}\,[\rho(\boldsymbol{r})] = T_s^{\mathrm{LDA}}[\rho(\boldsymbol{r})] \tag{14}$$

where

$$t_s^{\mathrm{hom}}[\rho(\boldsymbol{r})] = \frac{3}{10}(3\pi^2)^{2/3}\rho(\boldsymbol{r})^{5/3} \tag{15}$$

In the end the TF energy functional

$$E^{\mathrm{TF}}[\rho(\boldsymbol{r})] = T_s^{\mathrm{LDA}}[\rho(\boldsymbol{r})] + U_H[\rho(\boldsymbol{r})] + V[\rho(\boldsymbol{r})] \tag{16}$$

implies not only the single-particle approximation to the full electron-electron interaction, but also the single-particle mean-field approximation $T_s^{\mathrm{LDA}}[\rho(\boldsymbol{r})]$ to the exact kinetic energy of the inhomogeneous interacting system. As a consequence, all many-body correlation effects are neglected.

However, the HK theorem predicates that they are again a functional of $\rho(\boldsymbol{r})$. The addition of an approximation to the exact exchange and correlation (XC) energy results in a formally exact theory, which is referred to as orbital-free DFT (Smargiassi, Madden, 1994). It is therefore important to recognise, that the HK theorem is nothing but the formal exactification of the TF approximation. Similarly, the Kohn-Sham (KS) (1965) scheme can be considered as the exactification of the self-consistent Hartree equations (HE) (1928), which differs only in the kinetic energy from the TF approximation. In fact, for the fictitious non-interacting system the kinetic energy is known exactly, even though only in terms of an explicit single-particle orbital functional, i.e. as an implicit density functional

$$T_s[\rho(\boldsymbol{r})] = -\frac{1}{2}\sum_{i=1}^{N_e}\int d\boldsymbol{r}\, \psi_i^*(\boldsymbol{r})\nabla^2\psi_i(\boldsymbol{r}) = T_s[\{\psi_i[\rho(\boldsymbol{r})]\}] \tag{17}$$

Here the fictitious single-particle orbitals, or simply KS orbitals, are denoted as $\psi_i(\boldsymbol{r})$. As we will see immediately they are eigenfunctions of a fictitious system, known as the KS system. It therefore should be noted that they differ from the single-particle orbitals used in wavefunction based methods and have no strict physical meaning, with two notable exceptions: at the presence of the exact XC functional for the special case of an isolated system with $v(\infty) = 0$, (i) the highest occupied eigenvalue ε_N can be shown to be the negative of the exact, many-body, first ionisation potential including relaxation effects, and (ii) that the lowest unoccupied eigenvalue ε_{N+1} is the negative of the electron affinity. Beside these two exceptions, only the density has a real physical meaning and can be written in terms of $\psi_i(\boldsymbol{r})$ as

$$\rho(\boldsymbol{r}) = \sum_{i=1}^{N_{\mathrm{occ}}} f_i \psi_i(\boldsymbol{r}) \psi_i^*(\boldsymbol{r}) \tag{18}$$

where N_{occ} is the number of occupied orbitals and f_i the occupation number of state i, so that

$$\sum_{i=1}^{N_{\mathrm{occ}}} f_i = N_e \tag{19}$$

Therewith, the KS energy functional is simply given by

$$\begin{aligned} E^{\mathrm{KS}}[\rho(\boldsymbol{r})] &= E^{\mathrm{KS}}[\{\psi_i[\rho(\boldsymbol{r})]\}] \\ &= T_s[\{\psi_i[\rho(\boldsymbol{r})]\}] + U_H[\rho(\boldsymbol{r})] + V[\rho(\boldsymbol{r})] + E_{\mathrm{XC}}[\rho(\boldsymbol{r})] \\ &= -\frac{1}{2}\sum_{i=1}^{N} f_i \int d\boldsymbol{r}\, \psi_i^*(\boldsymbol{r}) \nabla^2 \psi_i(\boldsymbol{r}) \\ &+ \frac{1}{2}\int d\boldsymbol{r} d\boldsymbol{r}' \frac{\rho(\boldsymbol{r})\rho(\boldsymbol{r}')}{|\boldsymbol{r} - \boldsymbol{r}'|} \\ &+ \int d\boldsymbol{r}\, v_{\mathrm{ext}}(\boldsymbol{r})\rho(\boldsymbol{r}) + E_{\mathrm{XC}}[\rho(\boldsymbol{r})] \end{aligned} \tag{20}$$

where $E_{\mathrm{XC}}[\rho(\boldsymbol{r})] = (T[\rho(\boldsymbol{r})] - T_s[\{\psi_i[\rho(\boldsymbol{r})]\}]) - (U[\rho(\boldsymbol{r})] - U_H[\rho(\boldsymbol{r})]$ is the already mentioned and apparently unknown XC energy functional, whereas $v_{\mathrm{XC}}(\boldsymbol{r})$ is the corresponding XC potential. This definition also shows that a significant of part $E_{\mathrm{XC}}[\rho(\boldsymbol{r})]$ is due to correlation effects of the kinetic energy that is known explicitly only in terms of the reduced two-particle density matrix (Dreizler, Gross, 1990).

Since up to the exactifying term $E_{\mathrm{XC}}[\rho(\boldsymbol{r})]$ equation (20) is identical to the HE, it is not surprising that the corresponding Euler-Lagrangian equation

$$\left(-\frac{1}{2}\nabla^2 + v_s^{\mathrm{KS}}(\boldsymbol{r})\right)\psi_i(\boldsymbol{r}) = \varepsilon_i \psi_i(\boldsymbol{r}) \tag{21}$$

also results in a similar fictitious single-particle equation. Since $v_s^{\mathrm{KS}}(\boldsymbol{r}) = v_{\mathrm{H}}(\boldsymbol{r}) + v_{\mathrm{XC}}(\boldsymbol{r}) + v(\boldsymbol{r})$ is the effective potential of an artificial system, such that the ground state density and therewith the energy equals those of the true interacting many-body system.

This particular system is therefore called KS system, its effective potential KS potential and the resulting set of self-consistent equations are referred to as KS equations:

$$\begin{gathered}\left(-\frac{1}{2}\nabla^2 + v_{\mathrm{H}}(\boldsymbol{r}) + v_{\mathrm{XC}}(\boldsymbol{r}) + v(\boldsymbol{r})\right)\psi_i(\boldsymbol{r}) = \varepsilon_i\psi_i(\boldsymbol{r}) \\ \sum_{i=1}^{N_{\mathrm{occ}}} f_i\psi_i(\boldsymbol{r})\psi_i^*(\boldsymbol{r}) = \rho(\boldsymbol{r}) \\ \frac{\delta E_{\mathrm{XC}}[\rho(\boldsymbol{r})]}{\delta\rho(\boldsymbol{r})} = v_{\mathrm{XC}}(\boldsymbol{r})\end{gathered} \tag{22}$$

At self-consistency it is possible to express $E_0^{\mathrm{KS}}[\rho(\boldsymbol{r})]$ in terms of the single-particle KS eigenvalues ε_i. Due to the fact that they are not the eigenvalues of the interacting many-body system, but of fictitious non-interacting KS system, $E_0^{\mathrm{KS}}[\rho(\boldsymbol{r})]$ is merely the sum of ε_i, but

$$E_0^{\mathrm{KS}}[\rho(\boldsymbol{r})] = \sum_i^{N_{\mathrm{occ}}} f_i\varepsilon_i - \frac{1}{2}\int d\boldsymbol{r}d\boldsymbol{r}'\frac{\rho(\boldsymbol{r})\rho(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|} - \int d\boldsymbol{r}\, v_{\mathrm{XC}}(\boldsymbol{r})\rho(\boldsymbol{r}) + E_{\mathrm{XC}}[\rho(\boldsymbol{r})] \tag{23}$$

That is to say that in order to make genuine calculations the KS scheme systematically maps the full interacting many-body problem, with $\hat{U}$, onto an equivalent fictitious single-body problem, with an effective potential operator $\hat{V}_{\mathrm{KS}} = \hat{U}_s + \hat{V}_{\mathrm{H}} + \hat{V}_{\mathrm{XC}}$, but without $\hat{U}$ (Capelle, 2006):

$$\begin{array}{ccc} \mathrm{TF} & \xrightarrow{E_{\mathrm{XC}}[\rho(\boldsymbol{r})]} & \mathrm{HK} \\ T_s[\{\psi_i[\rho(\boldsymbol{r})]\}]\downarrow & & \downarrow T_s[\{\psi_i[\rho(\boldsymbol{r})]\}] \\ \mathrm{HE} & \xrightarrow[E_{\mathrm{XC}}[\rho(\boldsymbol{r})]]{} & \mathrm{KS} \end{array} \tag{24}$$

2.3.3. The exchange and correlation functional

In the previous subsection DFT has been outlined as an exact theory, presuming that the exact XC functional is known. Unfortunately, except for the uniform electron gas (Ceperley, Alder, 1980), this is not the case and one has to resort to more or less accurate approximations. For the sake of brevity, only the main physical principles rather than the various rungs of "Jacob's ladder to heaven" (Tao et al., 2003) will be discussed here.

On this account the following break-up is particularly convenient:

$$E_{\mathrm{XC}}[\rho(\boldsymbol{r})] = \frac{1}{2}\int d\boldsymbol{r}\int d\boldsymbol{r}'\,\frac{\rho(\boldsymbol{r})\rho_{\mathrm{XC}}(\boldsymbol{r},\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|} = E_{\mathrm{X}}[\rho(\boldsymbol{r})] + E_{\mathrm{C}}[\rho(\boldsymbol{r})] \tag{25}$$

where $E_{\mathrm{X}}[\rho(\boldsymbol{r})]$ is the exchange energy due to Pauli repulsion, $E_{\mathrm{C}}[\rho(\boldsymbol{r})]$ the electron correlation energy and $\rho_{\mathrm{XC}}(\boldsymbol{r},\boldsymbol{r}') = \rho_{\mathrm{X}}(\boldsymbol{r},\boldsymbol{r}') + \rho_{\mathrm{C}}(\boldsymbol{r},\boldsymbol{r}')$ the XC hole. The former is therefore the energy lowering due to the antisymmetry requirement on the wavefunction of a fermionic system and can be exactly calculated in terms of an explicit orbital or implicit density functional

$$E_{\mathrm{X}}[\rho(\boldsymbol{r})] = -\frac{1}{2}\int d\boldsymbol{r}\int d\boldsymbol{r}'\,\frac{\psi_i^*(\boldsymbol{r})\psi_j^*(\boldsymbol{r}')\psi_i(\boldsymbol{r})\psi_j(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|} = E_{\mathrm{X}}[\{\psi_i[\rho(\boldsymbol{r})]\}] \tag{26}$$

Equation (26) is the so-called Hartree-Fock exchange energy, but with KS orbitals. However, the nonlocal form of the exact exchange energy comes at a considerable computational burden to solve four-centre integrals, which is about two orders of magnitude more expensive than is the case for local or semi-local approximations to the exact XC functional. The correlation energy accounts for the additional energy lowering, since electrons with opposite spins also avoid each other. However, contrary to the exchange part, no exact expression for $E_{\mathrm{C}}[\rho(\boldsymbol{r})]$ is known, neither in terms of orbitals nor densities.

Obviously, DFT would be of little use if one had to know $E_{\mathrm{XC}}[\rho(\boldsymbol{r})]$ exactly, but luckily it is usually energetically substantially smaller than each of the remaining terms, which are known. One can thus hope that reasonable simple approximations to $E_{\mathrm{XC}}[\rho(\boldsymbol{r})]$, will still allows for qualitatively correct estimates of $E_0[\rho(\boldsymbol{r})]$, without relying on additional adjustable parameters.

2.4. *Ab-initio* molecular dynamics

In the following let us assume that the potential energy function is calculated on-the-fly using DFT, so that $\Phi(\boldsymbol{R}_I) = E[\{\psi_i\}; \boldsymbol{R}_I] = E^{\mathrm{KS}}[\{\psi_i[\rho(\boldsymbol{r})]\}] + E_{II}(\boldsymbol{R}_I)$. In any case, AIMD (Marx, Hutter, 2009; Car, Parrinello, 1985; Payne et al., 1992; Parinello, 1997; Car, 2002; Tuckerman, 2002) comes in two fundamental flavours, which are outlined in this section.

2.4.1. Born-Oppenheimer molecular dynamics

In Born-Oppenheimer MD (BOMD) the potential energy $E[\{\psi_i\}; \boldsymbol{R}_I]$ is minimised at every MD step with respect to $\{\psi_i(\boldsymbol{r})\}$ under the holonomic orthonormality constraint $\langle\psi_i(\boldsymbol{r})|\psi_j(\boldsymbol{r})\rangle = \delta_{ij}$. This leads to the following Lagrangian

$$\mathcal{L}_{\mathrm{BO}}(\{\psi_i\}; \boldsymbol{R}_I, \dot{\boldsymbol{R}}_I) = \frac{1}{2}\sum_{i=1}^{N} M_I \dot{\boldsymbol{R}}_I^2 - \min_{\{\psi_i\}} E[\{\psi_i\}; \boldsymbol{R}_I] + \sum_{i,j} \Lambda_{ij}\left(\langle\psi_i|\psi_j\rangle - \delta_{ij}\right) \qquad (27)$$

where Λ_{ij} is a Hermitian Lagrangian multiplier matrix. By solving the according Euler-Lagrangian equations

$$\begin{aligned} \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{\boldsymbol{R}}_I} &= \frac{\partial \mathcal{L}}{\partial \boldsymbol{R}_I} \\ \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \langle\dot{\psi}_i|} &= \frac{\partial \mathcal{L}}{\partial \langle\psi_i|} \end{aligned} \qquad (28)$$

one obtains the associated equations of motion (EOM)

$$\begin{aligned} M_I \ddot{\boldsymbol{R}}_I^2 &= -\nabla_{\boldsymbol{R}_I}\left[\min_{\{\psi_i\}} E[\{\psi_i\};\boldsymbol{R}_I]\,\middle|_{-\{\langle\psi_i|\psi_j\rangle-\delta_{ij}\}}\right] \\ &= -\frac{\partial E}{\partial \boldsymbol{R}_I} + \sum_{i,j}\Lambda_{ij}\frac{\partial}{\partial \boldsymbol{R}_I}\langle\psi_i|\psi_j\rangle - 2\sum_i \frac{\partial\langle\psi_i|}{\partial \boldsymbol{R}_I}\left[\frac{\delta E}{\delta\langle\psi_i|} - \sum_j \Lambda_{ij}|\psi_j\rangle\right] \\ 0 &\lesssim -\frac{\delta E}{\delta\langle\psi_i|} + \sum_j \Lambda_{ij}|\psi_j\rangle \\ &= -\hat{H}_e\langle\psi_i| + \sum_j \Lambda_{ij}|\psi_j\rangle \end{aligned} \tag{29}$$

The first term on the right hand side (RHS) of equation (29) is the so-called Hellmann-Feynman force. The second term, which is denoted as Pulay (1969), or wavefunction force F_{WF}, is a constraint force due to the holonomic orthonormality constraint, and is non-vanishing if and only if the basis functions ϕ_j explicitly depend on $\boldsymbol{R}_I$. The final term stems from the fact that, independent of the particular basis set, there is always an implicit dependence on the atomic positions through the expansion coefficient $c_{ij}(\boldsymbol{R}_I)$ within the common linear combination of atomic orbitals ϕ_i:

$$\psi_i(\boldsymbol{R}_I) = \sum_j c_{ij}(\boldsymbol{R}_I)\phi_j \tag{30}$$

The second factor stems from the assumption that the KS orbitals are real, which is an inessential simplification. Nevertheless, the whole term vanishes whenever $\psi_i(\boldsymbol{R}_I)$ is an eigenfunction of the Hamiltonian within the subspace spanned by the not necessarily complete basis set (Almlöf, Helgaker, 1981; Scheffler, Vigneron, Bachelet, 1985). Note, that this is a much weaker condition than the original Hellmann-Feynman theorem (Hellman, 1937; Feynman, 1939), which we hence have not availed throughout the derivation, except as an eponym for the first RHS term of equation (29). However, as the KS functional is non-linear, eigenfunctions of its Hamiltonian $\hat{H}_e$ are only obtained at exact self-consistency, which is why the last term of equation (29) is also referred to as non-self-consistent force F_{NSC}. Unfortunately, in any numerical calculation this cannot be assumed, which results in immanent inconsistent forces and to the inequality of equation (29). Neglecting either F_{WF} or F_{NSC}, i.e. applying the Hellmann-Feynman theorem to a non-eigenfunction leads merely to a perturbative estimate of the generalised forces (Bendt, Zunger, 1983)

$$F = F_{\mathrm{HF}} + F_{\mathrm{WF}} + F_{\mathrm{NSC}} \tag{31}$$

which, contrary to the energies, depends just linearly on the error in the electronic charge density. As a consequence, it is much more exacting to calculate accurate forces than total energies. However, as a corollary of the BO approximation, the electronic, as well as the ionic subsystems are adiabatically strictly separated from each other, and therefore does not entail any restrictions on the maximum possible integration time step, so that time steps up to the ionic resonance limit are feasible. This actually holds irrespective of the band gap, so, at least in principle, even metals can be straightforwardly treated.

2.4.2. Car-Parrinello molecular dynamics

In Car-Parrinello MD (CPMD) (1985), a coupled electron-ion dynamics is performed, in which the electronic degrees of freedom are added to the Lagrangian as classical ones:

$$\begin{aligned}\mathcal{L}_{\mathrm{CP}}\left(\{\psi_i\};\boldsymbol{R}_I,\dot{\boldsymbol{R}}_I\right) &= \frac{1}{2}\mu\sum_{i=1}^{M}\langle\dot{\psi}_i|\dot{\psi}_i\rangle+\frac{1}{2}\sum_{i=1}^{N}M_I\dot{\boldsymbol{R}}_I^2-E[\{\psi_i\};\boldsymbol{R}_I]\\ &+\sum_{i,j}\Lambda_{ij}\left(\langle\psi_i|\psi_j\rangle-\delta_{ij}\right)\end{aligned} \tag{32}$$

Once again, applying the Euler-Lagrangian equations equation (28) entails an EOM, where the electronic degrees of freedom inhere an artificial inertia μ and are propagated within a fictitious Newtonian dynamics, such that the electrons follows the ions adiabatically:

$$\begin{aligned}M_I\ddot{\boldsymbol{R}}_I^2 &= -\nabla_{\boldsymbol{R}_I}\left[E[\{\psi_i\};\boldsymbol{R}_I]\Big|_{-\{\langle\psi_i|\psi_j\rangle-\delta_{ij}\}}\right]\\ &= -\frac{\partial E}{\partial\boldsymbol{R}_I}+\sum_{i,j}\Lambda_{ij}\frac{\partial}{\partial\boldsymbol{R}_I}\langle\psi_i|\psi_j\rangle\\ \mu\ddot{\psi}_i(\boldsymbol{r},t) &= -\frac{\delta E}{\delta\langle\psi_i|}+\sum_j\Lambda_{ij}|\psi_j\rangle\\ &= -\widehat{H}_e\langle\psi_i|+\sum_j\Lambda_{ij}|\psi_j\rangle\end{aligned} \tag{33}$$

As a consequence of the BO approximation, the high frequency oscillations of equation (33) vanishes on ionic time scales, so that $\ddot{\psi}_i\simeq 0$. Hence, similar to Ehrenfest dynamics (1927), the total derivative of the instantaneous, rather than the fully minimised, expectation value $\langle\Psi_0|\widehat{H}_e|\Psi_0\rangle$ of the Hamiltonian yields the forces that are consistent with the corresponding energies. This means, that owing to the absence of necessity to fully minimise the energy functional but rather to simply evaluate it at the instantaneous time step, F_{NSC} is identical to zero by its very definition. Given a sufficiently small fictitious mass, the constant of motion is strictly conserved and errors in the forces are negligible, in particular if the ionic masses are renormalised by a constant mass tensor (Blöchl, Parrinello, 1992; Blöchl, 1994; Tangney, Scandolo, 2002). In this respect, CPMD combines most of the advantages of BO and Ehrenfest MD in the sense that the KS functional are only evaluated. There is no need to repeatedly solve it either by diagonalisation or, equivalently, iterative minimisation. However, due to the finite accuracy of any integrator, the holonomic orthonormality constraint of the orbitals has to be explicitly enforced. In order to ensure an adiabatic energy-scale separation of the nuclear and the electronic degrees of freedom and to prevent energy transfer between them, the ionic phonon frequency ω_I has to be much smaller than the electronic analogue ω_e. As had been suggested in an eminent phenomenological study of Pastore and Buda (1991)

$$\omega_e\propto\sqrt{\frac{\Delta E_{\mathrm{gap}}}{\mu}} \tag{34}$$

As a consequence, the maximum integration time step $\Delta t_{\max}$ depends on the inertia like $\sqrt{\mu}$. The same also holds for the deviation from the BO surface (Borbemann, Schütte, 1998)

$$\left|\psi_\mu(r,t) - \psi_0(r,t)\right| \leq C\sqrt{\mu} \tag{35}$$

Therefore, the fictitious mass, although physically completely meaningless, acts as a continuous slider which allows to adjust any desired degree of accuracy, in terms of deviation from the BO surface, reciprocal to the computational efficiency in a well controlled manner. But if a metallic system is treated, due to the fact that Car-Parrinello (CP) states are strictly not KS eigenstates, equation (34) is identical to zero and either a thermostat for the electronic degrees of freedom (Blöchl, Parrinello, 1992; Sprik, 1991; Blöchl, 2002) to counterbalance the exchange of energy, or an extended functional with fractional occupation numbers (Mermin, 1965; Gillan, 1989; Alavi et al., 1994; Marzari, Vanderbilt, Payne, 1997) is necessary. In the end drawing a proper conclusion, if either BOMD or CPMD is to be favoured, turns out to be very subtle (Tangney, 2006) and depends largely on the definition of accuracy and on the particular application.

2.5. An efficient and accurate Car-Parrinello-like approach to Born-Oppenheimer molecular dynamics

Even though DFT-based AIMD has been very successful in describing a large variety of physical phenomena, its computational cost has limited the attainable length and time scales in spite of substantial progress. For a while it was believed that linear scaling methods (Yang, 1991; Galli, Parrinello; 1992; Goedecker, 1999) could have offered a solution. Unfortunately, the crossover point at which linear scaling methods become advantageous has remained fairly large, especially if high accuracy is needed (Ceriotti, Kühne, Parrinello, 2008, 2009). Therefore, it would be very desirable to accelerate *ab-initio* simulations with up to thousands of atoms, such that simulations as long as a few nanoseconds can be routinely performed, thus making completely new phenomena accessible to AIMD simulations. BOMD, in which the DFT functional is fully minimised at each MD time step, does not seem to offer much room for further improvement. For this reason, recently another direction has been followed to improve the efficiency at current system sizes. In the spirit of CPMD (Car, Parrinello, 1985), some form of dynamics for the electronic degrees of freedom is implemented, which automatically keeps the system close to the instantaneous BO surface, but at variance to the original proposal in a localised orbital representation (Iyengar et al., 2001; Sharma, Wu, Car, 2003; Herbert, Head-Gordon, 2004). The acceleration stems on the one hand from this more compact description of the electronic wavefunctions, but is mainly due to the ability to reduce or even fully bypass the aforementioned SCF cycle. Nevertheless, just like in CPMD, these methods suffer from rather short integration time steps. However, rather recently a novel Car-Parrinello-like approach to BOMD has been proposed, which overcomes this limitation and combines the accuracy and long-time steps of BOMD with the efficiency of CPMD (Kühne et al., 2007).

From now on the general case will be considered, where the DFT KS orbitals are expanded in a non-orthogonal basis set. Let M be the dimension of the Hilbert space, i.e. the number of basis functions, and $\boldsymbol{S}$ the $M \times M$ overlap matrix. As usual the expansion coefficients of the N lowest occupied orbitals are arranged in a rectangular $M \times N$ matrix $\boldsymbol{C}$. The density matrix can then be written as $\boldsymbol{P} = \boldsymbol{C}\boldsymbol{C}^T$ and must obey the idempotency condition $\boldsymbol{P} = \boldsymbol{P}\boldsymbol{S}\boldsymbol{P}$ that is due to the fermionic nature of electrons, which compels the wavefunction to be antisymmetric in order to meet the Pauli Exclusion Principle. The potential energy surface

on which the ions move is defined by the minimum of an appropriately chosen energy functional $E_{\mathrm{DFT}}[\boldsymbol{C}, \boldsymbol{R}_I]$, which is expressed as a functional of $\boldsymbol{C}$ and a function of the ionic co-ordinates $\boldsymbol{R}_I$. In this notation the BO EOM reads as follows:

$$M_I \ddot{\boldsymbol{R}}_I^2 = -\nabla_I \min_{\mathbf{C}} E_{\mathrm{DFT}}[\boldsymbol{C}, \boldsymbol{R}_I] \tag{36}$$

where the search for the minimum is restricted to the $\boldsymbol{C}$'s that satisfy the orthonormality condition $\boldsymbol{C}^T \boldsymbol{S} \boldsymbol{C} = \boldsymbol{I}$, which is equivalent to imposing the idempotency condition on $\boldsymbol{P}$. As before, the forces of equation (36) can be divided into three contributions, (i) the Hellmann-Feynman forces (Hellmann, 1937; Feynman, 1939), (ii) the Pulay forces (1969), which are present whenever the basis set depends on the ionic positions, and (iii) a residual term (Bendt, Zunger, 1983), which is non-zero except when full self-consistency is reached. The last term leads inevitably to poor energy conservation in BOMD unless a very tight convergence criterion is imposed. In Car-Parrinello-like approaches this is circumvented by the design of a coupled electron-ion dynamics, which maintains the system very close to the BO surface, but at the cost of small integration time steps.

2.5.1. Density matrix propagation

Based on ideas of the original CP approach it is possible to design an improved dynamics for the coupled system of electrons and ions (Kühne et al., 2007). However, contrary to the original scheme, this novel method is not expressed as an explicit EOM for the $\boldsymbol{C}$'s, but rather as an integration scheme for the electronic degrees of freedom. The knowledge of the previous K values of $\boldsymbol{C}(t_{n-l})$, where $l \in [1, K]$, determines the value of $\boldsymbol{C}(t_n)$, such that at any instant of time the $\boldsymbol{C}$'s are as close as possible to the instantaneous ground state. As for the short-term integration of the electronic degrees of freedom, accuracy is crucial so a highly accurate and efficient algorithm is required. Therefore, here the always stable predictor-corrector (ASPC) method of Kolafa (2004, 2005) has been selected. This scheme was originally devised to deal with classical polarisation, so that care must be taken that during the evolution the idempotency condition is always satisfied. The modified predictor

$$\boldsymbol{C}^p(t_n) \cong \sum_{m=1}^{K} (-1)^{m+1} m \frac{\binom{2K}{K-m}}{\binom{2K-2}{K-1}} \underbrace{\boldsymbol{C}(t_{n-m})\boldsymbol{C}^T(t_{n-m})}_{\boldsymbol{P}(t_{n-m})} \boldsymbol{S}(t_{n-m})\boldsymbol{C}(t_{n-1}) \tag{37}$$

uses the extrapolated contra-covariant density matrix $\boldsymbol{PS}$ as an approximate projector on to the occupied subspace $\boldsymbol{C}(t_{n-1})$. In this way, the fact that the physically relevant contra-covariant density matrix $\boldsymbol{PS}$ evolves much more smoothly and is therefore substantially easier to predict than $\boldsymbol{C}$ is ideally utilised. The modified predictor is followed by a corrector step to minimise the error and to further reduce the deviation from the instantaneous ground state. The corrector

$$\boldsymbol{C}(t_n) = \omega \min[\boldsymbol{C}^p(t_n)] + (1-\omega)\boldsymbol{C}^p(t_n) \quad \text{with } \omega = \frac{K}{2K-1} \tag{38}$$

consists only of a single preconditioned minimisation step $\min[\boldsymbol{C}^p(t_n)]$ of a properly selected minimisation procedure. Apparently, the predictor can also be repeatedly applied, in which case the ground state is even more closely approached, but at the cost of additional electronic gradient calculations. However, as will be shown immediately in general this is not necessary. The numerical coefficients of equation (37) were selected in order to ensure time-reversibility up to $O(h^{K+2})$ while ω was chosen to guarantee a stable relaxation towards the minimum. Due to the fact that the energy is invariant under unitary transformations within the subspace of occupied orbitals $\boldsymbol{C}$, it must be ensured that this

time-dependent gauge transformation is not strongly changed by $\min[\boldsymbol{C}^p(t_n)]$, as in this case continuity between the $\boldsymbol{C}$'s may be lost.

2.5.2. Electronic forces by orbital transformations

Moreover, the minimisation scheme must be very efficient in bringing the system as close as possible to the instantaneous ground state and at the same time preserves the idempotency condition of the density matrix. For these reasons, the orbital transformation (OT) method of VandeVondele and Hutter (2003) has been chosen. Inspired by the form of the exponential transformation (Hutter, Parrinello, Vogel, 1994) an auxiliary variable X is introduced, to parameterise the occupied orbitals

$$\boldsymbol{C}(\boldsymbol{X}) = \boldsymbol{C}^p(t_n)\cos(\boldsymbol{U}) + \boldsymbol{X}\boldsymbol{U}^{-1}\sin(\boldsymbol{U}) \tag{39}$$

where $\boldsymbol{U} = (\boldsymbol{X}^T\boldsymbol{S}\boldsymbol{X})^{1/2}$ and the variable $\boldsymbol{X}$ has to obey the linear constraint $\boldsymbol{X}^T\boldsymbol{S}\boldsymbol{C}^p(t_n) = 0$. Under this condition $\boldsymbol{C}(\boldsymbol{X})$ leads to an idempotent density matrix for any choice of $\boldsymbol{X}$, provided that the reference orbitals $\boldsymbol{C}^p(t_n)$ are orthonormal. Thus, any finite step along the preconditioned gradient direction will exactly fulfil the idempotency constraint by construction. Due to the linear constraint the minimisation with respect to $\boldsymbol{X}$ is performed in an auxiliary tangent space. Since this space is linear, no curved geodesics must be followed, as is the case for variables such as $\boldsymbol{C}$ that are nonlinearly constrained. In this way, large minimisation steps can be taken, especially if a good preconditioner is used (Gan, Haynes, Payne, 2000). In fact, using an efficient, idempotency conserving direct minimiser such as OT is decisive for the success of this approach. Since the ASPC integrator only approximately preserves the idempotency constraint, it sporadically has to be explicitly enforced, either by Cholesky decomposition or by single purification iterations (McWeeny, 1960).

2.5.3. Total energies and forces

Having obtained the new wavefunction it is now possible to evaluate the energy and the nuclear forces, which are derived from the following approximate energy functional:

$$\begin{aligned} E_{\mathrm{PC}}[\rho^p] = \mathrm{Tr}\,[\boldsymbol{C}^T H[\rho^p]\boldsymbol{C}] - \frac{1}{2}\int d\boldsymbol{r}\int d\boldsymbol{r}'\,\frac{\rho^p(\boldsymbol{r})\rho^p(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|} \\ - \int d\boldsymbol{r}\, V_{\mathrm{XC}}[\rho^p]\rho^p + E_{\mathrm{XC}}[\rho^p] + E_{II} \end{aligned} \tag{40}$$

where ρ^p is the density associated with $\boldsymbol{C}^p(t_n)$. $E_{\mathrm{PC}}[\rho]$ can be thought of as an approximation to the Harris-Foulkes functional (Harris, 1985; Foulkes, Haydock, 1989) and maintains the predictor-corrector flavour of this method. The validity of $E_{\mathrm{PC}}[\rho]$ depends only on the efficiency of the minimiser and on the quality of the propagation scheme. The ionic forces are calculated by evaluating the analytic gradient of $E_{\mathrm{PC}}[\rho]$ with respect to the nuclear co-ordinates. However, as $\Delta\rho = \rho - \rho^p \neq 0$, besides the usual Hellmann-Feynman and Pulay forces an extra term appears:

$$-\int d\boldsymbol{r}\left\{\left[\left(\frac{\partial V_{\mathrm{XC}}[\rho^p]}{\partial\rho^p}\right)\Delta\rho + V_H[\Delta\rho]\right](\nabla_I\rho^p)\right\} \tag{41}$$

where ρ is the corrected density evaluated using $\boldsymbol{C}(t_n)$ and ρ^p is the predicted density calculated from $\boldsymbol{C}^p(t_n)$. Using variational density functional perturbation theory (Putrino, Sebastiani, Parrinello, 2000; Benoit, Sebastiani, Parrinello, 2001), equation (41) can be

efficiently computed very similar to employing the coupled-perturbed KS scheme. However, due to the fact that usually only a single preconditioned minimisation step is performed, $\boldsymbol{C}(t_n)$ is just an approximate eigenfunction of $H[\rho^p]$ within the subspace spanned by the finite basis set used. This leads to an insignificant error in the forces, provided that $\boldsymbol{C}(t_n)$ is very close to the ground state.

2.5.4. Modified Langevin equation

The ability of this dynamics to maintain the system on the BO surface may vary considerably. It is essentially ideal in systems like water, but potentially somewhat less perfect in liquid Si at high temperature, where swift bonding and rebonding processes continuously take place. However, in all cases the dynamics is dissipative, most likely because the employed propagation scheme is not symplectic. Nevertheless, it is possible to rigorously remedy this downward drift if we assume that the forces arising from our dynamics $\boldsymbol{F}_{\mathrm{PC}}$ can be modelled as $\boldsymbol{F}_{\mathrm{PC}} = \boldsymbol{F}_{\mathrm{BO}} - \gamma_D \dot{\boldsymbol{R}}_I$, which, as we shall see immediately, is an excellent assumption. The value of the intrinsic friction coefficient γ_D does not need to be known but it can be bootstrapped by taking a cue from the work of Krajewski and Parrinello (2001). The canonical distribution is sampled by using the following Langevin-type equation

$$M_I \ddot{\boldsymbol{R}}_I = \boldsymbol{F}_{\mathrm{PC}} - \gamma_L \dot{\boldsymbol{R}}_I + \Xi_I \tag{42}$$

where M_I is the ionic mass, γ_L is a Langevin friction coefficient and $\Xi_I = \Xi_I^D + \Xi_I^L$ an additive white noise. Using the above assumption equation (42) is identically to:

$$M_I \ddot{\boldsymbol{R}}_I = \boldsymbol{F}_{\mathrm{BO}} - (\gamma_D + \gamma_L)\dot{\boldsymbol{R}}_I + \Xi_I \tag{43}$$

In order to guarantee an accurate sampling of the Boltzmann distribution, the noise has to obey the fluctuation dissipation theorem:

$$\langle \Xi_I(0)\Xi_I(t)\rangle = 6(\gamma_D + \gamma_L) M_I k_{\mathrm{B}} T \delta(t) \tag{44}$$

The choice of γ_L is arbitrary, while the unknown γ_D has to be determined by requiring that the aggregate noise term generate the correct average temperature, i.e. fulfils the equipartition theorem $\langle \frac{1}{2} M_I \dot{\boldsymbol{R}}_I^2 \rangle = \frac{3}{2} k_{\mathrm{B}} T$. As we see in a moment, this leads to correct a sampling of the Boltzmann distribution. In addition, since the initial dynamics is quite accurate, γ_D is rather small and even dynamical properties can be very well reproduced.

2.5.5. Illustrative examples: liquid silicon, silica and water

For the purpose of demonstrating this new approach, it has been implemented in the mixed Gaussian Plane Wave (GPW) (Lippert, Hutter, Parrinello, 1997) code QUICKSTEP (Krack, Parrinello, 2004; VandeVondele et al., 2005), which is part of the publicly available suite of programmes CP2K.[2] In order to illustrate that this method works well irrespective of band gap, system size and type, calculations on metallic liquid silicon and liquid silica are presented. Both systems are known to be very difficult, and are examples of liquid metals (Si) as well as of complex, highly polarisable, ionic liquids (SiO_2). Furthermore, the simulations have been performed at 3 000 K and 3 500 K respectively, which leads to rapidly varying density matrix elements, thus making the propagation of the electronic

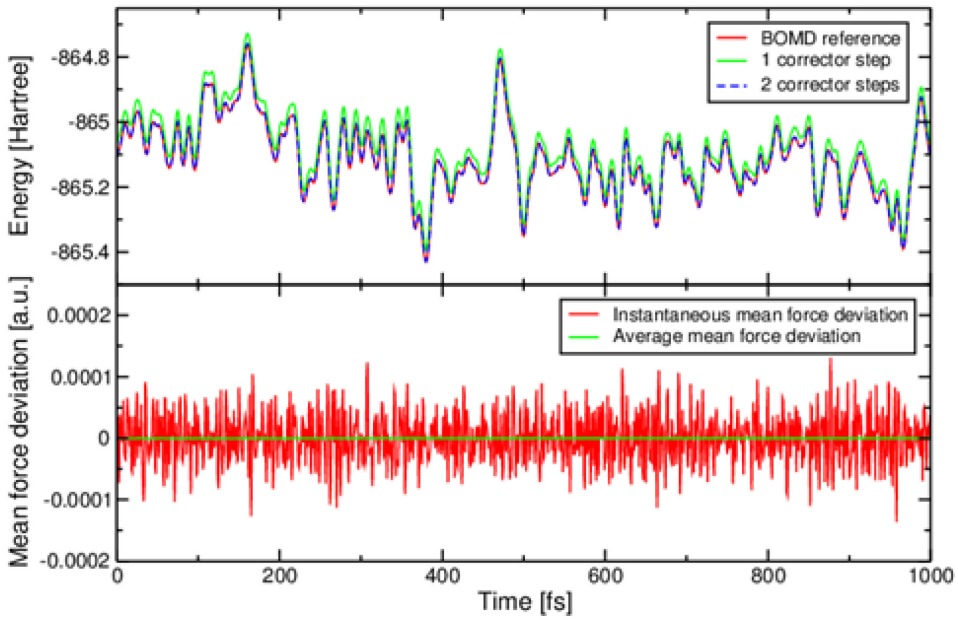
2. www.cp2k.org

degrees of freedom particularly challenging. Hence, the selected test cases can be considered as worst-case scenarios for any method.

All simulations have been performed at their experimental liquid densities using double-zeta valence polarisation (DZVP) basis sets, adequate density cutoffs, Goedecker-Teter-Hutter pseudopotentials (Goedecker, Teter, Hutter, 1996; Hartwigsen, Goedecker, Hutter, 1998; Krack, 2003) and the local density approximation to the exact exchange and correlation functional. For simplicity the Brillouin zone is sampled at the Γ-point only, while equation (43) is integrated using the algorithm of Ricci and Ciccotti (2003), with a time step of $h = 1$ fs. The friction coefficient γ_L was set equal to zero, while the values for γ_D turned out to be in the range of 10^{-4} fs^{-1}. The new $\boldsymbol{C}$'s are predicted using $K = 4$ in equation (37), which ensures time-reversibility up to $O(h^6)$.

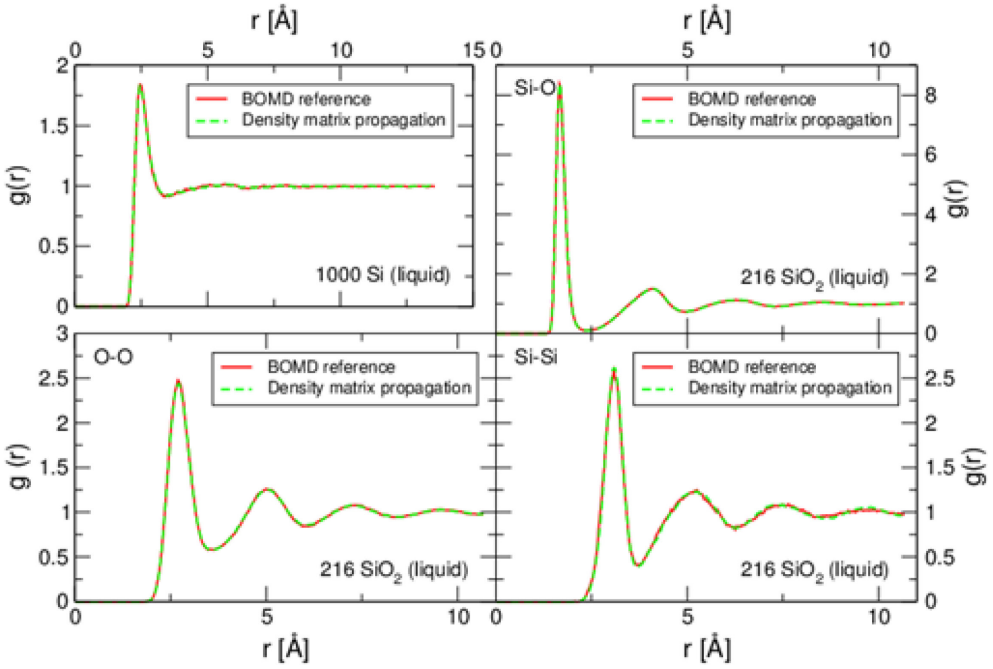
First, the accuracy in terms of the energetic deviation from the BO surface is considered. As can be seen in Figure 2.1. the energies are an upper bound to the ground state and are displaced by a very small and approximately constant amount. It is also shown that, as already mentioned, the deviation from the BO surface can be even further reduced by increasing the number of corrector steps. In fact, it is actually possible to control the deviation from the BO surface by varying the number of corrector steps in order to achieve a preassigned accuracy level. However, in the following only simulations based on a single corrector step will be reported.

Figure 2.1. Deviations from the BO surface of liquid SiO_2 with respect to total energies (upper panel) and mean force deviations (lower panel)



Note: The deviation in the energies corresponds to a constant shift of 4.16 · 10−4 Hartree per atom for one corrector step and 3.5 · 10−5 Hartree per atom for two corrector steps. The average mean force deviation is unbiased.
Source: Kühne, 2019.

Figure 2.2. Partial pair-correlation functions g(r) of liquid Si (upper left panel) and liquid SiO_2 at 3 000 K and 3 500 K respectively, using a DZVP Gaussian basis set

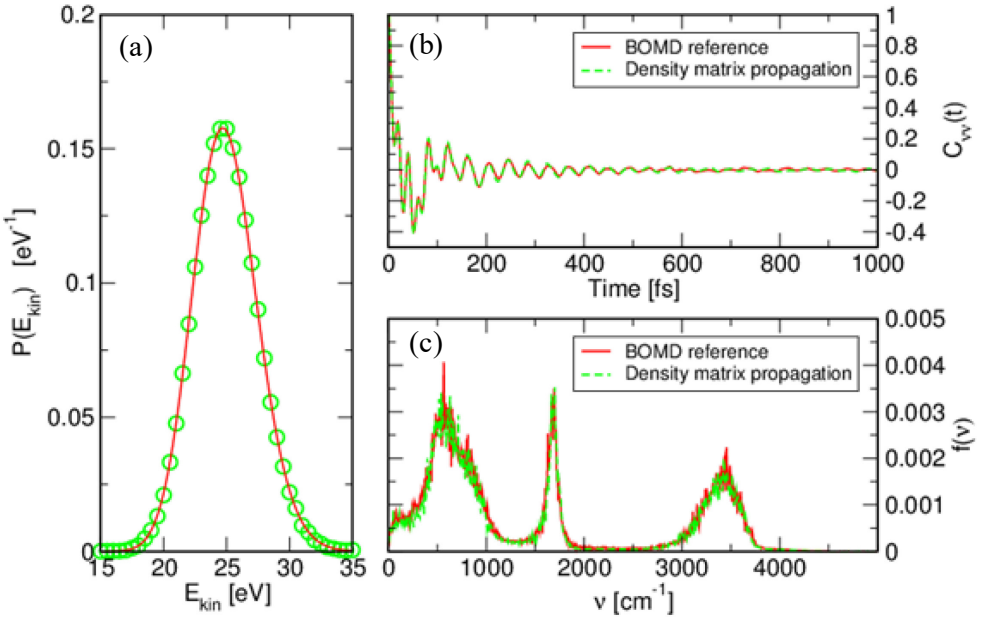


Source: Kühne, 2019.

Nevertheless, let us now to turn to more realistic problems such as those shown in Figure 2.2. Although these simulations have been performed with only a single corrector step, they are still amazingly close to the BOMD reference results. It should be emphasised that even in liquid Si, which is metallic and poses problems when using an ordinary CP scheme, a single corrector step is sufficient. This establishes the efficiency of this method, since only a single preconditioned gradient calculation with no additional minimisation step has to be performed. The possible acceleration, in comparison with regular BOMD calculations, depends crucially on the system studied. In the undoubtedly difficult cases just presented a speedup of two orders of magnitude compared to using a pure extrapolation scheme has been observed. For simpler problems still an increase in efficiency of at least one order of magnitude can be expected.

In Figure 2.3. displays results, which prove that also dynamical properties can be evaluated with accuracy. To that extent the velocity autocorrelation function and its Fourier transform at 325 K is presented. The results are in good agreement with accurate reference calculations and are consistent with experiment, as well as *ab-initio* all-electron calculations (Krack, Parrinello, 2000), showing that in spite of the stochastic nature of equation (43) dynamical properties can also be simulated. This implies that also chemical reactions and even non-equilibrium processes can be treated. In the same picture it is explicitly verified that the previous assumptions are justified, and indeed a canonical sampling is performed, by showing that the kinetic energy distribution is Maxwellian distributed. To this end, a 64 atom liquid Si simulation is carried out for as long as 1 ns, to reduce the noise and to ensure a proper sampling of the relevant kinetic energy distribution tails.

Figure 2.3. (a) The kinetic energy distribution of metallic liquid Si_{64} (b) Velocity autocorrelation function. (c) Fourier transform of the velocity autocorrelation function of 32 water at 325 K.



***Note*:** Calculation details: (a) 1 ns trajectory, time step: 3.25 fs, basis set: DZVP Gaussian basis set, density cutoff: 100 Ry (c) basis set: TZV2P Gaussian basis set, density cutoff: 280 Ry, exchange-correlation functional: BLYP,. Langevin friction coefficients: $\boldsymbol{\gamma_L = 0}$ and $\boldsymbol{\gamma_D \sim 10^{-8}}$ fs^{-1}.

Source: Kühne, 2019.

Due to space considerations only a fraction of the systems studied is reported here. Nevertheless, in all cases this method has proven to be accurate and the gain in speed has always been remarkable (Kühne, Krack, Parrinello, 2009; Camellone, Kühne, Passerone, 2009; Cuinotta et al., 2009; Kühne et al., 2011; Luduena, Kühne, Sebastiani, 2011a, 2011b). Structure relaxations via dynamic annealing and geometry optimisation have also been successfully performed (Carvati et al, 2007, 2009a, 2009b, 2009c, 2010, 2011). Contrary to CPMD and related methods integration time steps up to the ionic resonance can be used. Thanks to this development it is now possible to perform AIMD simulations on medium-sized systems up to a few nanoseconds, thus making a new class of problems accessible to *ab-initio* simulations.

2.6. Conclusion

To conclude it should be noted, that with increasing length and time scales CPMD-based approaches are expected to become more advantageous than BOMD, since otherwise meeting the more and more stringent accuracy requirements of longer simulations and larger system sizes would entail an ever tighter wavefunction convergence. The Car-Parrinello-like approach to BOMD (Kühne et al, 2007) just described extends the scope of *ab-initio* simulations by combining the best of either method and allows for AIMD simulations previously thought not feasible.

Acknowledgements

The author would like to thank Matthias Krack and Michele Parrinello for various fruitful discussions, as well as the whole CP2K-Team. In addition, he is indebted to the Graduate School of Excellence MAINZ and the IDEE project of the Carl-Zeiss Foundation for financial support. The generous allocation of computer time and support from CSCS, the ICT Services of ETH Zurich and the DEISA consortium is kindly acknowledged.

References

Alavi, A. et al. (1994), "Ab initio molecular dynamics with excited electrons", *Phys. Rev. Lett.*, Vol. 73, 2599.

Alder, B.J. and T.E. Wainwright (1957), "Phase transition for a hard sphere system", *J. Chem. Phys.*, Vol. 27, p. 1 208.

Almlöf, J. and T. Helgaker (1981), "Basis set considerations for the calculations of gradients in the lcao formalism", *Chem. Phys. Lett.*, Vol. 83, p. 125.

Bendt, P. and A. Zunger (1893), "Simultaneous relaxation of nuclear geometries and electric charge densities in electronic structure theories", *Phys. Rev. Lett.*, Vol. 50, p. 1 684.

Benoit, D.M., D. Sebastiani and M. Parrinello (2001), "Accurate total energies without self-cosistency", *Phys. Rev. Lett.*, Vol. 87, 226401.

Blöchl, P.E. and M. Parrinello (1992), "Adiabaticity in first-principles molecular dynamics", *Phys. Rev. B*, Vol. 45, 9413.

Blöchl, P.E. (1994), "Projector augmented-wave method", *Phys. Rev. B*, Vol. 50, 17953.

Blöchl, P. (2002), "Second-generation wave function thermostat for *ab initio* molecular dynamics", *Phys. Rev. B*, Vol. 65, 104303.

Born, M. and R. Oppenheimer (1927), "Zur Quantentheorie der Molekeln", *Ann. Phys. (Leipzig)*, Vol. 389/20, p. 457.

Bornemann, F.A. and C. Schütte (1998), "A mathematical investigation of the Car-Parrinello method", *Numer. Math.*, Vol. 78, p. 359.

Camellone, M.F., T. D. Kühne and D. Passerone (2009), "Density functional theory study of self-trapped holes in disordered SiO_2", *Phys. Rev. B*, Vol. 80, 033203.

Capelle, K. (2006), "A bird's-eye view of density-functional theory", *Braz. J. Phys.*, Vol. 36, 1318.

Car, R. (2002), "Introduction to density-functional theory and ab-initio molecular dynamics", *Quant. Struct. Act. Rel.,* Vol. 21, p. 97.

Car, R. and M. Parrinello (1985), "Unified approach for molecular dynamics and density-functional theory", *Phys. Rev. Lett.*, Vol. 55, 2471.

Caravati, S., M. Bernasconi, T.D. Kühne, M. Krack, M. Parrinello (2007), "Coexistence of tetrahedral and octahedral-like sites in amorphous phase change materials", *Appl. Phys. Lett.*, Vol. 91, 171906.

Caravati, S., M. Bernasconi, T.D. Kühne, M. Krack, M. Parrinello (2009a), "Unravelling the mechanism pf pressure induced amorphization of phase change materials", *Phys. Rev. Lett.*, Vol. 102, 205502.

Caravati, S., M. Bernasconi, T.D. Kühne, M. Krack, M. Parrinello (2009b), "First-principles study of crystalline and amorphous $Ge_2Sb_2Te_5$ and the effects of stoichiometric defects", *J. Phys.: Condens. Matter*, Vol. 21, 255501.

Caravati, S., M. Bernasconi, T.D. Kühne, M. Krack, M. Parrinello (2009c), "Corrigendum: First-principles study of crystalline and amorphous $Ge_2Sb_2Te_5$ and the effects of stoichiometric defects", *J. Phys.: Condens. Matter*, Vol. 21, 499803.

Caravati, S., M. Bernasconi, T.D. Kühne, M. Krack, M. Parrinello (2010), "Corrigendum: First-principles study of crystalline and amorphous $Ge_2Sb_2Te_5$ and the effects of stoichiometric defects", *J. Phys.: Condens. Matter*, Vol. 22, 399801.

Caravati, S., D. Colleoni, R. Mazzarello, T.D. Kühne, M. Krack, M. Bernasconi and M. Parrinnello (2011), "First-principles study of nitrogen doping in cubic and amorphous $Ge_2Sb_2Te_5$", *J. Phys.: Condens. Matter*, Vol. 23, 265801.

Ceperley, D.M. and B.J. Alder (1980), "Ground state of the electron gas by stochastic method", *Phys. Rev. Lett.*, Vol. 45, p. 566.

Ceperley, D.M. (1995), "Path integrals in the theory of condenses helium", *Rev. Mod. Phys.*, Vol. 67, p. 279.

Ceriotti, M., T.D. Kühne, and M. Parrinello (2008), "An efficient and accurate decomposition of the Fermi operator", *J. Chem. Phys.*, Vol. 129, 024707.

Ceriotti, M., T.D. Kühne and M. Parrinello (2009), "A hybrid approach to Fermi operator expansion", *AIP Conf. Proc.*, Vol. 1148, p. 658.

Cucinotta, C.S. et al. (2009), "Superionic conduction in substoichiometric LiAl Alloy: An *Ab Initio* study", *Phys. Rev. Lett.*, Vol. 103**,** 125901.

Dreizler, R.M. and E.K.U. Gross (1990), *Density Functional Theory*, Springer-Verlag, Berlin.Fermi, E. (1927), "Un Metodo Statistico per la Determinazione di alcune Prioprietà dell'Atomo", *Rend. Accad. Naz. Lincei*, Vol. 6, p. 602.

Ehrenfest, P. (1927), "Bemerkung über die angenäherte Gültigkeit der klassischen Mechanik innerhalb der Quantenmechanik", *Z. Phys*., Vol. 45, p. 455.

Foulkes, W.M.C. and R. Haydock (1989), "Tight-binding models and density-functional theory", *Phys. Rev. B*, Vol 39, 12520.

Feynman, R.P. (1939), "Forces in molecules", *Phys. Rev.*, Vol. 56, p. 340.

Feynman, R.P. and A.R. Hibbs (1965), *Quantum Mechanics and Path Integrals*, McGraw-Hill, New York.

Galli, G. and M. Parrinello (1992), "Large scale electronic structure calculations", *Phys. Rev. Lett.*, Vol. 69, 3547.

Gan, C.K., P. D. Haynes, and M.C. Payne (2000), "Preconditioned conjugate gradient method for the sparse generalized eigenvalue problem in electronic structure calculations", *Comput. Phys. Commun.*, Vol. 134, p. 33.

Gilbert, T.L. (1975), "Hohenberg-Kohn theorem for nonlocal external potentials", *Phys. Rev. B*, Vol. 12, 2111.

Gillan, M.J. (1989), "Calculations of the vacancy formation energy in alunium", *J. Phys.: Condens. Matter*, Vol. 1, p. 689.

Goedecker, S. (1999), "Linear scaling electronic structure methods", *Rev. Mod. Phys.*, Vol. 71, 1085.

Goedecker, S., M. Teter, and J. Hutter (1996), "Separable dual-space Gaussian pseudopotentials", *Phys. Rev. B*, Vol. 54, p. 1 703.

Harriman, J.E. (1982), "Orthonormal orbitals for the presentation of an arbitrary density", *Phys. Rev. A*, Vol. 24, p. 680.

Harris, J. (1985), "Simplified method for calculating the energy of weakly interacting fragments", *Phys. Rev. B*, Vol. 31, 1770.Hartree, D.R. (1928), "The wave mechanics of an atom with a non-coulomb central field. Part I. Theory and methods", *Proc. Cambridge Philos. Soc.*, Vol. 24, p. 89.

Hartwigsen, C., S. Goedecker, and J. Hutter (1998), "Relativistic separables dual-space Gaussian pseudopotentials from H to Rn", *Phys. Rev. B*, Vol. 58, 3641.

Hohenberg, P. and W. Kohn (1964), "Inhomogeneous electron gas", *Phys. Rev.*, Vol. 136, B864.

Hellmann, H. (1937), *Einführung in die Quantenchemie*, Deuticke, Leipzig.

Herbert, J.M. and M. Head-Gordon (2004), "Curvy-steps approach to constraint-free extended-Lagrangian *ab initio* molecular dynamics, using atom-centered basis functions: Convergences toward Born-Oppenheimer trajectors", *J. Chem. Phys.*, Vol. 121, 11542.

Hutter, H., M. Parrinello, and S. Vogel (1994), "Exponential transformation of molecular orbitals", *J. Chem. Phys.*, Vol. 101, 3862.

Iyengar, S.S. et al (2001), "*Ab initio* molecular dynamics: Propagating the density matrix with Gaussian orbitals", *J. Chem. Phys.*, Vol. 114, 9758.

Kohn, W. and L.-J. Sham (1965), "Self-consistent equations including exchange and correlation effects", *Phys. Rev.*, Vol. 140, A1133.

Kolafa, J. (2004), "Time-reversible always stable predictor-corrector method for molecular dynamics of polarizable molecules", *J. Comput. Chem.*, Vol. 25, p. 335.

Kolafa, J. (2005), "Gear formalism of the always stable predictor-corrector method for molecular dynamics of polarizable molecules", *J. Chem. Phys.*, Vol. 122, 164105.

Krack, M. (2005), "Pseudopotentials for H to Kr optimized for gradient-corrected exchange-correlation functional", *Theor. Chem. Acc.*, Vol. 114, p. 145.

Krack, M. and M. Parrinello (2000), "All-electron *ab-initio* molecular dynamics", *Phys. Chem. Chem. Phys.*, Vol. 2, 2105.

Krack, M. and M. Parrinello (2004), *Forschungszentrum Jülich*, NIC Series, Vol. 25, p. 29.

Krajewski, F.R. and M. Parrinello (2006), "Linear scaling for quasi-one-dimentional systems", *Phys. Rev. B*, Vol. 73, 041105(R).

Kühne, T.D. et al. (2007), "Efficient and accurate Car-Parrinello-like approach to Born-Oppenheimer Molecular Dynamics", *Phys. Rev. Lett.*, Vol. 98, 066401.

Kühne, T.D., M. Krack and M. Parrinello (2009), "Static and dynamical properties of liquid water from first principles by a novel Car-Parrinello-like approach", *J. Chem. Theory Comput.*, Vol. 5, p. 235

Kühne, T.D. et al. (2011), "New insights into the structure of the vapour/water interface from large-scale first-principles simulations", *J. Phys. Chem. Lett.*, Vol. 2, p. 105.

Levy, M. (1982), "Electron densities in search of Hamiltonians", *Phys. Rev. A*, Vol. 26, 1200.

Lieb, E.H. (1983), "Density functionals for coulomb systems", *Int. J. Quantum Chem.*, Vol. 24, p. 243.

Lippert, G., J. Hutter and M. Parrinello (1997), "A hybrid Gaussian and plane wave density functional scheme", *Mol. Phys.*, Vol. 92, p. 477.

Luduena, G.A., T.D. Kühne and D. Sebastiani (2011a), "Mixed Grotthuss and vehicle transport mechanism in proton conducting polymers from Ab initio molecular dynamics simulations", *Chem. Mater.*, Vol. 23, p. 1 424.

Luduena, G.A., T.D. Kühne and D. Sebastiani (2011b), "Reply to Mixed Grotthuss and vehicle transport mechanism in proton conducting polymers from Ab initio molecular dynamics simulations", *Chem. Mater.*, Vol. 23, p. 3 379.

Marx, D, and J. Hutter (2009), *Ab Initio Molecular Dynamics: Basic Theory and Advance Methods*, Cambridge University Press, Cambridge.

Marzari, N., D. Vanderbilt and M.C. Payne (1997), "Ensemble density-functional theory for ab initio molecular dynamics of metals and finite-temperature insulators", *Phys. Rev. Lett.*, Vol. 79, p. 1 337.

McWeeny, R. (1960), "Some recent advances in density matrix theory", *Rev. Mod. Phys.*, Vol. 32, p. 335.

Mermin, N.D. (1965), "Themal properties of the inhomogenous electron gas", *Phys. Rev.* Vol. 137, A1441.

Metropolis, N. et al. (1953), "Equation of state calculations by Frast Computer Machines", *J. Chem. Phys.*, Vol. 21, p. 1 087.

Parrinello, M. (1997), "From silicon to RNA: The coming of age of *ab initio* molecular dynamics", *Solid State Commun.*, Vol. 102, p. 107.

Pastore, G., E. Smargiassi, and F. Buda (1991), "Theory of ab initio molecular-dynamics calculations", *Phys. Rev. A*, Vol. 44, 6334.

Payne, M.C. et al. (1992), "Iterative minimization techniques for *ab initio* total-energy calculations: molecular dynamics and conjugate gradients", *Rev. Mod. Phys.*, Vol. 64, 1045.

Pulay, P. (1969), "*Ab initio* calculation of force constants and equilibrium geometries in polyatomic molecules", *Mol. Phys.*, Vol. 17, p. 197.

Putrino, A., D. Sebastiani and M. Parrinello (2000), "Generalized variational density functional perturbation theory", *J. Chem. Phys.*, Vol. 113, 7102.

Rahman, A. (1964), "Correlations in the motion of atoms in liquid Argon", *Phys. Rev.*, Vol. 136, A405.

Ricci, A. and G. Ciccotti (2003), "Algorithms for Brownian dynamics", *Mol. Phys.*, Vol. 101, p. 1927.

Scheffler, M., J.P. Vigneron, and G.B. Bachelet (1985), "Total-energy gradients and lattice distortions at point defects in semiconductors", *Phys. Rev. B*, Vol. 31, 6541.

Sharma, M, Y. Wu, and R. Car (2003), "Ab initio molecular dynamics with maximally localized Wannier functions", *Int. J. Quantum Chem.*, Vol. 95, p. 821.

Smargiassi, E. and P.A. Madden (1994), "Orbital-free kinetic-energy functionals for first-principles molecular dynamics", *Phys. Rev. B*, Vol. 49, 5220.

Sprik, M. (1991), "Computer simulation of the dynamics of induced polarization fluctuations in water", *J. Phys. Chem.*, Vol. 95, p. 2 283.

Tangney, P. (2006), "On the theory underlying the Car-Parrinello method and the role of the fictitious mass parameter", *J. Chem. Phys.*, Vol. 124, 044111.

Tangney P. and S. Scandolo (2002), "How well do Car-Parrinello simulations reproduce the Born-Oppenheimer surface?", *J. Chem. Phys.*, Vol. 116, p. 14.

Tao, J. et al. (2003), "Climbing the density functional ladder: Nonempirical meta-generalized approximation designed for molecules and solids", *Phys. Rev. Lett.*, Vol. 91, 146401.

Tuckerman, M.E. (2002), "*Ab initio* molecular dynamics: basic concepts, current trends and novel applications", *J. Phys. Condens. Matter*, Vol. 14, R1297.

Thomas, J.W., R. Iftimie, and M.E. Tuckerman (2004), "Field theory approach to dynamical orbital localization in *ab initio* molecular dynamics", *Phys. Rev. B*, Vol. 69, 125105.

Thomas, L.H. (1927), "The calculation of atomic fields", *Proc. Cambridge Philos. Soc.*, Vol. 23, p. 542.

VandeVondele, J. and J. Hutter (2003), "An efficient orbital transformation method for electronic structure calculations", *J. Chem. Phys.*, Vol. 118, 4365.

VandeVondele, J. et al. (2005), "The influence of temperature and density functional models in ab initio molecular dynamics simulations of liquid water", *J. Chem. Phys.*, Vol. 122, 014515.

Verlet, L. (1967), "Computer "Experiments" on classical lluids. I. Thermodynamical propeties of Lennard-Jones molecules", *Phys. Rev.*, Vol. 159, p. 98.

Yang, W. (1992), "Direct calculations of electron density in density-functional theory", *Phys. Rev. Lett.*, Vol. 66, p. 1 438.

3. Classical interatomic potentials

T. R. Zeitler and L. J. Criscenti
Sandia National Laboratories, United States

3.1. Introduction

The behaviour of nuclear materials under operating conditions is difficult, expensive, unsafe and in some cases impossible to observe experimentally. Computer simulation provides an attractive alternative to direct experimental observation and has the distinct advantage of allowing for systematic study of individual effects on materials properties. Classical simulations have a rich history of providing atomic-scale information based on simple energy minimisation (e.g. predicting stable defect configurations), molecular dynamics (MD) (i.e. the time evolution of a system of atoms), and Monte Carlo (MC) (i.e. stochastic) simulations that can then be used to drive the design of future experiments. These techniques require a description of the potential energy of a system, which is usually described via a set of classical potentials (or force field).

A set of interatomic potentials consists of parameterised equations and parameter values developed for specific interactions that allow for calculation of potential energy as a function of a geometric variable (usually interatomic separation, but angles are sometimes used) (Cygan, 2001). For example, a simple harmonic potential has the form $E(r_{ij})=k(r_{ij}-r_o)^2$, where the parameters k and r_o are defined for the interaction of two types of atoms i and j — e.g. O and H in an O-H bond—and the interatomic separation is given by r_{ij}. A harmonic potential for the interaction of Si and O in a Si-O bond will have the same form, but different parameter values. The potential energy of a system is therefore defined by the potentials and the geometry of the system. In this way, the potential energy of a complex system can be calculated by summing the contributions of each interaction according to the prescribed analytical expressions defined by the potentials (Cygan, 2001).

Interatomic potentials are parameterised and verified using experimental results and electronic-scale *ab-initio* calculations. These potentials are then used as input for atomic-scale simulations. The materials properties calculated from atomic-scale simulations (e.g. thermal conductivity or defect populations) can then be used for higher-level (e.g. kinetic Monte Carlo (KMC)) calculations. Of course, the reliability of the properties calculated by higher-level methods depend heavily on the quality of the underlying potentials used for the atomic-scale simulations. Clearly, developing reliable potentials is of great necessity to multi-scale modelling endeavours and the ability of a higher-level method to accurately describe physical processes. For example, a classical potential should be able to properly reproduce point defect populations before a higher-level method could be expected to accurately calculate the effects of defects on microstructure for a material undergoing neutron irradiation (Becquart, Domain, 2011).

If the ultimate goals of simulation are to predict material lifetimes or the long-term microstructure evolution due to exposure to severe conditions experienced in nuclear reactors (e.g. elevated temperatures, corrosive environments, cyclic loading [Devananthan et al., 2001]), then an accurate description of atomic-scale interactions for nuclear materials is necessary. Increased reliability and confidence in potential models will lead to improved multi-scale modelling results. When an accurate multi-scale modelling approach is

implemented, a feedback loop will then be available which will allow for increasing efficiency in exploiting the properties of current materials (e.g. enhancing enrichment and fuel burnup in nuclear fuels) and a true engineering of novel nuclear materials. One of the indispensable keys to a proper multi-scale model is the development of quality interatomic potentials.

3.2. Key issues in developing classical potentials

3.2.1. Fitting/verification

After assuming a form of the classical potential (e.g. harmonic, Buckingham), parameters are fit in order to replicate experimental data (typically static elastic and dielectric constants and lattice parameters, but also defect properties) or theoretical calculations. Parameters are changed systematically, manually or using a computer program such as GULP (Gale, Gulp, 1997), to minimise the difference between given data and those calculated with the new potential. The number of parameters needed for a complete potential depends on the form of the potential. Potential parameters can be improved by including a greater number of experimental comparisons.

A key test to ensure an accurate potential is model verification. This can be done by comparison with experimental data that has not been used to initially fit the parameters, though calculated properties (e.g. from *ab-initio* MD simulations) also provide additional points of comparison. In most cases, some experimental data is used as fitting parameters for development of the model, so it is not surprising when there is strong agreement with these same calculated properties. However, more complicated properties (i.e. second-order properties) are increasingly used as both fitting parameters and points of verification. One example is parameterising to vibrational power spectra.

3.2.2. Applicability/transferability

One of the key issues in potential development has been determining the extent to which a set of interatomic potentials can reliably be applied to situations far removed from the fitting domain. This is true in terms of interatomic separations, temperature/energy range, and composition (e.g. the transferability from pure to mixed systems). For example, potential parameters that are fit to zero K lattice properties may not be able to fully reproduce dynamical properties at high temperatures (e.g. predicting melting point). And fitting parameters to a distinct range of temperature data does not guarantee a proper fit for temperatures outside of that range. In this way, a lack of quality experimental data can limit the development of a quality potential.

A lack of initial fitting parameters can lead to limitations in its range of applicability. A single model that is applicable over a large range of temperature and composition is rare (Tiwary van de Walle, Gronbech-Jensen., 2009). By fitting to a large number of initial structures, transferability of a potential set to a new system (which can be fundamentally different, but consist of the same components) is helped (Cygan, Liang, Kalinichev, 2004). General force fields (such as the universal force field (UFF) (Rappe et al., 1992) or Dreiding (Mayo et al., 1990) force field) have been used for materials (Duren, Bae, Snurr, 2009; Nalaparaju et al., 2009) that did not exist when the force fields were created.

3.3. Classification of classical potentials

A set of classical potentials can be classified in a number of ways, including its description of individual ions, what types of interactions are included, and the specific forms of those interactions.

3.3.1. Description of ion

Two basic models of interatomic potentials have been used with nuclear materials, the core-shell model and the rigid ion model. The core-shell model accounts for polarisation effects by describing ions as massive cores connected to massless charged shells by a spring (Dick, Overhauser, 1958). In this case, the shells interact with one another via short-range potentials, while long-range Coulomb interactions take place among all cores and shells. Some examples of core-shell models are those of Jackson (1986) and Gotte (2007). In the rigid ion model, all ions are described as massive point charges. MD simulations with rigid ion models are faster because there are fewer interactions to calculate. Rigid ion models include those of Morelon (2003), Sindzingre and Gillan (1988), and Karakasidis (1994).

Defect energy calculation benefits from a description of electronic polarisation of the lattice ions. This is usually accounted for by use of core-shell models in which an ionic dipole can be represented by the core-shell separation. However, due to the extremely high energy considered in displacement cascade simulations (discussed below), core-shell models are not suitable for this type of simulation (Morelon et al., 2003).

3.3.2. Interatomic potential types

Besides a model type, each set of potentials has general forms of energy description. The electrostatic contribution to energy is long-range in nature and given by the well-known form of the Coulomb energy:

$$-E_{ij} = \frac{1}{4\pi\epsilon_0}\frac{q_i q_j}{r_{ij}} \tag{1}$$

where q_i is the charge assigned to atom i, and r_{ij} is the interatomic distance between atoms i and j and ε_0 is the permittivity of free space. The total contribution of electrostatics can be obtained by summing over all atoms and their periodic images. The Ewald summation (1921) or Wolf's direct method (1999) are typically used to efficiently calculate this sum. In some cases, partial charges are used instead of formal charges.

Other potentials are restricted to apply only for a smaller range of interatomic separations (i.e. a cutoff separation is enforced) and are thus termed "short-range" contributions. The simplest short-range description of forces is that of a two-body ("pair") potential that is defined only between two atoms. A number of pair potential forms exist, including Buckingham (1938), Born-Mayer-Huggins (BMH) (Fumi, Tosi, 1964a, 1964b), Lennard-Jones (L-J) (Frenekl, Smit, 2002), and Morse (1929).

Many-body potentials are used in cases where a simple two-body treatment is insufficient to properly replicate fitting data. In a simple case, the potential energy of a three-body term is based on the positions of two neighbours. This formulation is able to represent some directionality in bonding. Some examples of three-body potential forms are Stillinger-Weber (1985) and the embedded-atom method (EAM) (Daw, Baskes, 1983).

Finally, complex reactive force fields, where explicit charge transfer and even the breaking of bonds are permitted, are becoming more popular as computational power continues to

explode. Some examples of reactive force fields include REAXFF (van Duin et al. 2008) and that of Garofalini (2001).

3.3.3. Specific forms of classical potentials

Buckingham

The Buckingham form of pair potentials is common for oxide materials and has the form:

$$E_{ij} = A_{ij} exp\left(\frac{-r_{ij}}{\rho_{ij}}\right) - \frac{C_{ij}}{{r_{ij}}^6} \tag{2}$$

where A_{ij}, ρ_{ij}, and C_{ij} are parameters for the interactions of atoms of type i and j. The first term represents Pauli electron repulsion due to overlapping electron densities while the second describes van der Waals attraction. At the very short interatomic separations experienced in high-energy displacement cascade simulations, the net Buckingham potential is unphysically attractive, which has led to the development of splined potentials that result in repulsion even at very small interatomic distances (Ziegler, Biersack, Littmark et al., 1985; Govers et al., 2007).

The BMH potential energy description is similar to the Buckingham potential (Lewis, Catlow, 1985). The Bushing-Ida (Yoshiaki, 1976) form of potential energy has been used in a number of models (Basak, Sengupta, Kamath, 2003; Yamada et al., 2000a) for UO_2 and can be written as the summation of Buckingham and Morse potentials (Watanabe et al., 2008).

Morse

Covalent bonding between anions and cations is sometimes accounted for via use of a Morse potential of the generic form:

$$E_{ij} = D_0\left[1 - e^{-\alpha(r-r_0)}\right]^2 \tag{3}$$

Lennard-Jones

A simple and commonly used potential form is that of L-J:

$$E_{ij} = D_0\left[\left(\frac{R_0}{r_{ij}}\right)^{12} - 2\left(\frac{R_0}{r_{ij}}\right)^6\right] \tag{4}$$

where D_0 is the depth of the potential energy well and R_0 is the equilibrium atomic separation. The first term describes electron repulsion and the second reflects London dispersion (Cygan, 2001).

Vashishta-Rahman

A Vashishta-Rahman (VR) potential has the form:

$$E_{ij} = \frac{q_i q_j}{r_{ij}} + \left(\frac{A_{ij}(\sigma_i + \sigma_j)^{\eta_{ij}}}{r_{ij}^{\eta_{ij}}}\right) - \left(\frac{P}{r_{ij}^4}\right) - \frac{C_{ij}}{{r_{ij}}^6} \tag{5}$$

where the first term is the Coulomb term, the second term represents electron repulsion, the third term represents a monopole-induced dipole interaction, and the final term describes van der Waals attraction (Gunay et al., 2011).

Ziegler-Biersack-Littmark

The Ziegler-Biersack-Littmark (ZBL) universal pair potential (1985) is widely used for small interatomic separations where a traditional Buckingham or BMH potential is unphysically attractive (Tiwary et al. 2009, 2001; Ackland, 1997). The ZBL potential is then smoothly connected to the other form using some sort of (arbitrary) spline, such that the overall potential, as well as its derivative is always continuous. The morphology of a displacement cascade is heavily influenced by this repulsion at small separations (Becquart et al., 2000). An alternative to the ZBL potential is a Buckingham (4-range) potential, sometimes used to avoid the attraction at unphysically small O-O separations (Jackson et al., 1986; Read, Jackson, 2010; Björkas, Nordlund, 2007).

Embedded-atom method

The EAM (Daw, Baskes, 1983, 1984) and Finnis-Sinclair (FS) (1984) are potential forms for describing metallic bonding and represent significant advances in metallic potentials. EAM consists of pair potentials plus an embedding energy to represent many-body interactions due to the local environment. This embedding energy is analogous to the bonding term in FS. In addition to being a function of interatomic separation, the interaction between two atoms is also a function of local environment. These potentials are density-dependent, which makes validation more challenging, especially when the potentials are extended to alloys.

The form of the EAM potential is:

$$E_i = F_\alpha\left(\sum_{j\neq i}\rho_\beta(r_{ij})\right) + \frac{1}{2}\sum_{j\neq i}\phi_\beta(r_{ij}) \tag{6}$$

where F is the embedding energy and is a function of atom i's environment via the density ρ of surrounding atoms. The potential is defined for atom types i and j and applies to all neighbours j of atom i within a prescribed cutoff distance. A pair potential ϕ is also defined for the i-j interaction. The EAM formulism has been shown to be versatile for a number of different Fe-alloys (Jang, Lee, Hong, 2008).

3.4. State of the art in classical potentials for nuclear fuel cycle materials

3.4.1. Nuclear fuels

UO_2

The most common nuclear fuel in use is UO_2; likewise, the greatest number of computational studies of nuclear fuels is for UO_2. A number of different potentials have been used over the past 50 years to model UO_2 structure, defect behaviour, and diffusion, as well as the effect of neutron irradiation on those properties. In many cases, incremental advances in the UO_2 potentials have led to a greater agreement between simulation results and experimental data (or theoretical calculations).

The first UO_2 potential was introduced in 1962 (Benson, Freeman, Dempsey, 1963). Catlow et al. later developed a series of improved UO_2 potentials (1978, 1977) in the 1970s – these were core-shell models fit to elastic and dielectric constants and lattice parameter and were designed to calculate defect formation energies. A set of rigid ion potentials (Walker, Catlow, 1981) of the BMH form was also developed by Catlow et al to be able to run MD at high temperatures in order to investigate superionic properties. In addition, Grimes and Catlow also developed an extensive set of self-consistent core-shell potentials to examine migration and solution energies of Xe, I, Ba, Kr, Rb, Br, He, Sr, Y, Te, Cs, Ba, La, Ce, Zr, Ru in UO_2. (See references 26-34 in Govers et al., 2007.)

Jackson later modified (Jackson et al., 1986, 1985) Catlow's potentials to eliminate the unphysical attraction of Buckingham energy at small separations by adding a new functional form of the potential and combining it with the Buckingham form using polynomial interpolations (splines). Sindzingre and Gillan (1988) later modified O-U interactions for the same potential and added a splined (4-range) Buckingham for O-O interactions. Karakasidis and Lindan (1994) revised these parameters after discovering that the O-O spline potential was not stable.

Another core-shell model was developed for MD simulations of UO_2 by Lindan and Gillan (1994) to look at oxygen diffusion at high temperature and the superionic phase. Yamada et al. (2000a) developed an independent potential with non-integer charges (q_U=+2.4) and a covalent U-O (Morse) term.

Morelon et al. (2003) used point defect energies (static and migration) for fitting to a BMH-type potential, eliminating U-U short-range interactions, zeroing the attractive term for U-O interactions, and revising the partial charges of U and O species (Morelon et al., 2003). These defect properties are important in displacement cascade simulations, though the migration energies are more difficult to calculate (than defect energy formation) because they require a determination of the transition structure of the migration process (Morelon et al., 2003). The Morelon potential showed improved agreement with thermal expansion data, especially above 2000 K (whereas the Yamada potential (2000a) is only good up to about 1 500 K) where anharmonicity effects have been taken into account by the use of defects and transition states were included in the parameterisation of potential parameters. It also shows improved defect energies, but mediocre agreement with elastic properties.

An extensive two-part investigation of interatomic potentials by Govers et al. (2007, 2008) for UO_2 includes a thorough review of the history of UO_2 potentials from 1962 to 2007. In addition to detailing significant modifications of a number of independently-developed potentials, the first part of the investigation compared 19 separate versions of potential models on their ability to reproduce experimental and *ab-initio* properties including lattice parameters; elastic constants; phonon dispersion modes; and the formation, migration and binding energies for various defects using static calculations. The potentials reproduced the lattice parameters and elastic constants, though this is not surprising because most parameters were originally fit to these properties. A wide range of values was calculated for the other properties, leading the authors to conclude that no best potential existed in the group. However, they could conclude that neither the class (i.e. core-shell or rigid ion) nor type (e.g. Buckingham, Morse) of model had a great impact on the accuracy of the potentials; rather, it was the parameterisation of the model that was critical.

In the second part of the investigation, Govers et al. (2008) used MD simulations to test the same potentials for their abilities to reproduce the temperature dependence of various properties including lattice parameters, specific heat and bulk modulus, as well as melting temperature. This was a test of the anharmonicity of the potentials, i.e. their abilities to

provide information on non-equilibrium properties even though they were fit to static lattice properties. While Govers et al. again conclude that no single best potential was found, they make the generalisation that rigid ion models with partial-charge assignments and some core-shell potentials were best able to reproduce experimental data (likely because thermal expansion curves were part of the fitting processes in many of these cases). Core-shell potentials are generally developed to be able to reproduce static polarisation effects and thus do not accurately reproduce anharmonicity effects (Govers et al., 2008). However, they did note that they expect challenges with partial-charge rigid ion models when dealing with charged defects (e.g. how to properly determine the ionicity of the defect)—and suggest that a potential developed with formal charges would be more appropriate in that case.

Since Govers' extensive reviews, additional UO_2 models have been published, as well as a comparison of various potentials for their ability to produce defects via energetic recoils (Devananthan, Yu, Weber, 2009). This study showed a dramatic effect of a chosen potential on the calculated defect population. Because the size of defect clusters and distribution of defects play a significant role in the later evolution of the defect site, a proper description of the potential energy is crucial to be able to have confidence in higher-level calculations. Also, for an electronic insulator such as UO_2, defects play a major role in determining thermal transport properties. This study also found that the potential of Yakub (2007) best replicated equilibrium properties.

Tiwary et al. (2009) developed a UO_2 potential that is applicable across all interatomic distances by fitting to the results of *ab-initio* calculations of defect energies. They also underlined the importance of a proper splining procedure to join multiple potentials that are individually valid over separate spatial regions.

Using a novel methodology for potential development, Read and Jackson (2010) introduced a new potential to more accurately represent the UO_2 lattice following Frenkel defect creation due to neutron irradiation and to correct the overestimated contribution of the attractive term in the cation-anion Buckingham potentials previously published. This model was developed to enhance applicability over a wide range of structural data and a wide range of other physical properties. In this core-shell model, the anion-anion Buckingham potential was derived and the cation-anion short-range interactions were given a BMH form. Though not included in the original parameterisation, calculated phonon dispersion curves and defect energies compare well with literature values. This model has not yet been applied to MD simulations.

A new potential by Gunay (2001) has been shown to be excellent at describing the transition to the superionic state and melting in UO_2. The potential has a VR form and is a rigid ion model with partial-charge assignments.

Other UO_2 models of note include those of Arima (rigid ion) (2005), Meis (core-shell) (1998), and Busker (core-shell) (Abramowski, Grimes, Bradford, 2000; Busker, Grimes, Bradford, 2000, 2003). The Busker model is notable because it has good transferability among a number of elements and has ability to study nonstoichiometry since it is parameterised for U^{3+} and U^{5+}.

Because it is radioactive, UO_2 is experimentally difficult to use for studying the effects of radiation damage. However, CeO_2 is similar to UO_2 in terms of fundamental properties (e.g. lattice parameter, density, thermal diffusivity), so it has been proposed as a surrogate material for investigations of radiation damage in UO_2 (Aidhy, Wolf, El-Azab, 2001). A number of pair potentials have been developed for CeO_2, including rigid ion BMH and

core-shell Buckingham forms (Xu et al., 2010). Particularly of interest is that some have been parameterised for nonstoichiometry (Gotte et al., 2007).

PuO_2/MOX fuel

Yamada et al. (2000b) have developed rigid ion potentials for mixed oxide (MOX) and nitride nuclear fuels, adding a Morse potential to the Buckingham potential. The potentials were developed to be transferable among solid solutions of UO_2 and PuO_2 for MOX fuels. The form of the rigid ion potential is Buckingham with an additional Morse term for cation-anion pairs. Thermal and mechanical properties of MOX fuel were calculated via MD simulations. This potential was later modified by Basak (2003) to better reproduce thermal expansion data. The partial-charge potentials developed by Arima (2005) are also of the BMH-type and have been used to examine thermal properties of (U, Pu)O_2 MOX fuels via MD simulation.

Tiwary et al. (2011) developed an approach for creating a consistent form of the potential over all distances and applied it to UO_2 (Tiwary et al., 2009) and then extended its use to PuO_2 and NpO_2 (Tiwary et al., 2011) in a self-consistent manner such that MOX fuels (U, Pu, Np)O_2 can be simulated over a wide range of experimental conditions. They fit the potentials to an extensive database of experimental and *ab-initio* results and thoroughly showed the potentials reliably reproduce properties outside of the fitting parameters. This represents a major step forward in the production of transferable potentials.

Nitride fuel

A mixed nitride (U, Pu)N fuel potential was developed by fitting to thermal expansion and pressure expansion data to a BMH + Morse form (Kurosaki et al., 2000); and, calculating thermal and mechanical properties with MD simulations. Radiation defect generation in UN was studied via *ab - initio* calculations, the results of which can be used to develop better quality classical potentials (Kotomin et al., 2008).

3.4.2. Cladding/structural elements

Fe and Fe-alloys

Ferritic and austentitic steels in nuclear reactors (e.g. pressure vessel ferritic steel and internal structure austentitic steel) are subject to neutron irradiation that can lead to stress corrosion cracking and eventually failure. Metal cladding encapsulates nuclear fuel and is exposed to neutron irradiation, which can lead to swelling (due to void formation) and affect diffusion properties in the material. The effects of irradiation on a number of steel/iron alloys are now routinely studied using MD displacement cascade simulations where primary knock-on atoms (PKA) are given a high initial kinetic energy and based on the time evolution of the system of interest, defects are introduced in a perfect crystal as momentum is transferred to a PKA's neighbour atoms.

The quality of the potentials describing interatomic interactions influences the ability of the simulation to predict neighbour atom trajectories, diminished mechanical properties, and defect generation and clustering (Devananthan, Yu, Weber, 2009; Malerba et al., 2010; Malerba, 2006). Primary damage defects in iron include self-interstitial atoms (SIAs), vacancies, and Frenkel pairs, and their clustering can give an indication of diffusion mechanisms and nucleation sites for larger defects (Björkas, Nordlund, 2007). The results from MD cascade simulations can be used as input into higher-level (e.g. KMC or continuum) models, so accuracy in calculated properties is of great importance (Becquart,

Domain, 2011). With this consideration, many recent studies rely on electronic-scale *ab - initio* DFT calculations for development and verification of the interatomic potentials used in MD simulations. This is particularly important because experimental observation of irradiation effects is nearly impossible due to the time and length scales at work (Bequart, Domain, 2011).

Primarily, Fe has been studied as a model system for steel, though development of new potentials now allows for study of a number of Fe-alloys. Bequat and Domain (2011) give an excellent review of the over 45-year history of damage formation investigations for Fe and Fe-alloys (e.g. Fe-Cr and Fe-Cu). The effects of P (Ackland et al., 2004), Cu (Pasianot, Malerba, 2007), and He (Seletskaia et al., 2007; Gao et al., 2011), which are known to embrittle Fe, have also been considered. Little research has been done on the Fe-C system (Becquart et al, 2007) due to difficulty of simulating *ab - initio* supercell sizes that are large enough to depict the low C concentration present in ferritic steels. The effects of interatomic potentials on irradiation damage in Fe are reviewed by Becquart et al. (2000).

Despite the seemingly simplistic structure of Fe, a number of potentials have been developed, none of which are able to reproduce experimental data entirely. Of particular concern are the well-known deficiencies in many Fe models for predicting defect energies (Malerba, 2006). Understanding the accumulation of damage due to neutron irradiation is crucial to the understanding of the effects on Fe, but there are still discrepancies in predicted defect cluster populations at the atomic scale (Björkas, Nordlund, 2007; Malerba, 2006) that make predicting longer time scale damage problematic. Although *ab-initio* calculations of defect structures and energies continue to be increasingly important for interatomic potential development and verification (Becquart, Domain, 2009), there is still a lack of fundamental understanding of how and why defect clustering occurs and how that clustering affects defect mobility (Samaras, Victoria, Hoffelner, 2009).

A number of studies exist which compare radiation damage effects among multiple Fe potentials (Becquart et al., 2000; Malerba et al., 2010). One of these studies (Becquart et al, 2000) found significant differences in primary damage creation due to subtle differences in the potentials. While many Fe potentials have been derived for equilibrium-type calculations, they are later “hardened” (to eliminate the possibility of close contacts of atoms) to be suitable for cascade simulations. Although all potentials in the study had reasonable threshold displacement energies, they each have had different hardening procedures, which leads to differences in repulsive potential for small separations. These differences result in significant differences in cascade shape, as well as defect energies and populations. Even thermal properties are affected. It was concluded that the different hardening methods play a role in these properties, underlining the importance of proper potential development and the applicability of different potentials over different interatomic separation ranges (Becquart et al, 2000).

The Mendelev set of interatomic potentials for body centred cubic-Fe (bcc-Fe) (Mendelev et al, 2003) is of the EAM type and was derived based on a careful fitting procedure that took into account a large set of experimental data and DFT-derived point defect energies. A more recent analysis of the state of the art in Fe potentials indicates that Mendelev-type potentials are the best choice for developing Fe-alloy potentials, as well as applying the calculated properties to larger scale models (Malerba et al., 2020). This study also concludes that the ability of classical interatomic potentials to capture SIA clusters and dislocations is still elusive. The study provides an extensive database of static and dynamic properties (including defect formation and migration) for Fe based on the results of simulations for five published potential models.

Some more recent models of note include those of Muller (2007) and Dudarev and Derlet (2005, 2011). Muller's potential for Fe is able to reproduce the temperature-dependent body centred cubic to face centred cubic (bcc-to-fcc) transition of Fe. The "magnetic" potential of Dudarev and Derlet contains an explicit magnetic contribution to the interaction energy.

3.4.3. Waste forms

The containment of high-level waste (HLW) has been an active area of research and development for over 50 years (Weber et al., 2009). Borosilicate glasses and complex ceramic composites have been developed to meet HLW containment issues. Research to date has focused on immobilising existing HLW that consists of all the liquid effluents from the reprocessing of commercial and defence spent nuclear fuel. New research is focused on creating new separation waste streams, recycling spent nuclear fuel, and minimising the volume and heat load of waste forms. New waste forms tailored to specific waste streams and high waste loadings may be needed. Two useful review papers on nuclear waste glasses and ceramic waste forms for actinides are from Grambow (2006) and Lumpkin (2006), respectively.

Waste streams include: (1) the long-lived fission product ^{99}Tc, (2) the principal heat-generating isotopes ^{137}Cs and ^{90}Sr, (3) lanthanides, (4) minor actinides, (5) remaining fission products, (6) volatile radionuclides, (7) the undissolved solids from fuel dissolution (Weber et al., 2009). Waste-form types include (1) glass, (2) ceramic, and (3) glass-ceramic. Classical MD has been used to look at the chemical durability of glass and various single-phase ceramic waste forms. The main vulnerabilities of polyphase ceramic or glass-ceramic waste forms for radionuclide immobilisation are along grain boundaries and can be better investigated using mesoscale models.

Glass waste forms

Glass waste forms are the most commonly used HLW form today. The amorphous and relatively disordered structures of glasses can incorporate a wide range of chemical elements. Borosilicate glass compositions may include up to 30-40 different elements. Although borosilicate glass is used by most countries, Russia has chosen to use aluminophosphate glass for HLW. One of the advantages of glass waste forms is that the particles emitted in radioactive decay can be readily accommodated in their amorphous structure, and the effects of α-decay are small at the ambient temperatures expected over the decay times for actinides. Volume changes and changes in chemical durability due to α-decay are minimal. Potential problems with glass waste forms relative to crystalline waste forms include (1) more rapid dissolution in groundwater, and (2) radioisotope diffusion rates in the waste form could increase the local surface concentration of radioisotopes exposed to water.

Classical MD simulations have been used to study glass structure and the interaction of glass surfaces with water for many years. The choice of interatomic potentials to use in classical MD simulations is of fundamental importance for the accuracy of simulated physical and chemical properties such as structure, elastic constants, heat capacities and other thermodynamic properties (Pedone et al., 2006). Numerous interatomic potentials are available in the literature for silica glass and silica polymorphs (van Beest, Kramer, van Santen, 1990; Duffrene, Kieffer, 1998; Demiralp, Cagin, Goddard, 1999; Vashista et al., 1990; Takada et al., 2004). These force fields incorporate a variety of functions and terms to describe different silica glass properties.

However, fewer potentials are available for multicomponent glasses like those proposed for HLW immobilisation, and there have not been many efforts to provide generalised self-consistent force fields that are capable of modelling the structures, mechanical, and chemical properties of silica-based glasses with different compositions (Pedone et al., 2006). Table 3.1. lists classical potentials found in the literature that have been developed and used to simulate multicomponent silica-based glasses. Table 3.2. lists classical and reactive potentials that have been used to study the interaction of multicomponent glass with water.

Table 3.1. Force field models for glass

Composition	Form of potential	Authors
Na, K, Ca, Si, B, O	Modified BMH used by Busing w/o dispersion terms	Soules (1979)
SiO_2, $AlPO_4$	Buckingham form known as BKS	Van Beest, Kramer, and van Santen (1990)
Al, O	Modified BMH	Blonski and Garofalini (1993)
Li, Cs, B, O	BMH	Verhoef and den Hartog (1995)
Mg, Ca, Al, Si, O	Modified BMH used by Busing w/o dispersion terms	Okuno and Kawamura (1995)
Na, Ca, Al, Si, B, Zr, O	BMH	Delaye and Ghaleb (1996)
Na, Ca, Al, Si, B, O	BMH – improvement on D&G (1996)	Abbas et al. (2003) Cormier et al. (2003)
Na, Si, O	Buckingham	Du and Cormack (2004)
Al, O	Buckingham	Adiga et al. (2006)
Na, Ca, Si, Al, O, Zr	Morse function with partial ionic charge model	Pedone et al. (2006)

Table 3.2. Force field model for glass-water interaction

Composition	FF for glass	FF for –OH and H_2O	Authors
Si, Na, H_2O	BMH	Shell model for H_2O -OH: Morse potential	Leed and Pantano (2003)
Si, O, Si-OH, H_2O	Buckingham (Teter's)	3-body term for Si-O-H; Coulomb subtracted Morse potential for –OH group	Du and Cormack (2005)
Si, O, Si-OH, H_2O	BMH + 3 bond angle terms	Rahman, Stillinger & Lemberg form for Si-H, O-H, and H-H	Feuston and Garofalini (1990)
Si, O, Si-OH, H_2O	Modified BMH + 3-body potential	Modified BMH + 3-body potential	Litton and Garofalini (2001)
Si, O, Si-OH, H_2O	Modified BMH + 3-body potential	Modified BMH + 3-body potential	Garofalini (2001)

From our review of available force fields, three stand out as potential candidates for future research on the design of new HLW glass waste forms and their dissolution properties. The first force field was specifically developed by Delaye and Ghaleb (1996) to simulate a nuclear waste glass matrix. As in previous simulation studies of the structure of simple glasses (e.g. Soules, 1979; Soules, Varshneya, 1981), potentials of the BMH form were used to represent a glass comprising the major components of a nuclear waste glass (SiO_2 + B_2O_3 + ZrO_2 + Al_2O_3 + Na_2O (+ CaO)). Three-body potentials were applied to O-Si-O, O-B-O and Si-O-Si triplets. The local environmental structures showed overall agreement with experimental results and the simulated densities, thermal expansion coefficients and viscosities of the simulated glasses were on the same order of magnitude as experimental data. This force field has been used to examine a suite of nuclear waste glass compositions (Delaye, Louis-Achille, Ghaleb, 1997). Experimentally determined changes in glass structure with increasing boron concentration were reproduced by the calculations.

Additional work by this French group includes combined multiple-quantum magic angle spinning nuclear magnetic resonance (MQ-NMR) spectroscopic and MD studies to examine the Na environment in several two- and three-component glasses (Angeli, Delaye, Charpentier, Petit, Ghaleb and Faucon, 2000a) and changes in Al-O-Si angle with increasing Ca content (Angeli, Delaye, Charpentier, Petit, Ghaleb and Faucon, 2000b). Abbas et al. (2003) used this same force field model to compare the structural properties of bulk glass with a glass surface. The simulations predict the migration of alkali cations towards the surface, lower coordination numbers for trivalent elements in the subsurface layer, oxygen enrichment in the outer layer, and larger tetrahedral ring structures on the surface.

The second force field model is that of Pedone et al. (2006) who developed a generalised self-consistent force field that is able to model both the structures and mechanical properties of silica-based glasses with different compositions. To produce a coherent set of potential functions, these researchers performed empirical fitting to structural and mechanical properties of a large set of crystalline oxides. The potentials developed in this work consist of three terms: a long-range Coulomb potential, a short-range Morse function, and a repulsive contribution (C/r^{12}). The resulting potential is given by:

$$U(r) = \frac{z_i z_j e^2}{r} + D_{ij}\left[\left\{1 - e^{-a_{ij}(r-r_0)^2}\right\} - 1\right] + \frac{C_{ij}}{r^{12}} \quad (7)$$

A rigid ionic model with partial charges is used to handle the partial covalency of silicate systems in order to model the quenching of melts and structures and mechanical properties of glasses. A core-shell model might provide more accurate results for surfaces, but would require a shorter-time step and lead to computationally more expensive quench simulations.

This force field has been parameterised to include numerous glass components including many of importance in nuclear waste glasses (Na_2O, CaO, Al_2O_3, SiO_2 and ZrO_2). However, to our knowledge, this force field has neither been used to study glass surface structures nor combined with a water model to examine glass surface hydroxylation or other interactions between glass surfaces and water. In earlier papers by the same group (Du, Cormack, 2005; Zeitler, Cormack, 2006), surface hydroxyl-silicon and hydroxyl-phosphorous potentials were developed to study the hydroxylation of silica and phosphosilicate glasses. These potentials include a short-range interaction described by a Buckingham equation, a potential function for the hydroxyl group that has the Coulomb-subtracted Morse form, and a three-body term introduced to reproduce the Si(P)-O-H bond angles on glass surfaces. The surface structures were used successfully to investigate the energy of chemisorption of water on different surface sites and to determine the OH coverage of glasses.

The third force field that shows promise for studying multicomponent nuclear waste glass dissolution is the reactive force field of Garofalini (2001). From the beginning, this force field was developed to investigate glass surface structures and the interaction of water with glass surfaces. It now includes parameters for Si, Al, and B, and allows for breaking glass network bonds between Si-O, Al-O and B-O by reaction with water molecules (Garofalini, 2011). This force field has been demonstrated to successfully reproduce bulk glass and hydroxylated surface structures, but is not the best for calculating the mechanics or energetics of these systems. Mahadevan and Garofalini (2007, 2008) have developed a new dissociative water potential to improve upon the previous force field model. However, this new potential is still being developed for multicomponent systems.

Ceramic waste forms

In crystalline ceramic phases, radionuclides occupy specific atomic positions in the periodic structures of constituent crystalline phases allowing high loadings of specific radionuclides (Weber et al., 2009). Waste-form phases tend to have complex structures, with a number of different coordination polyhedra of various sizes and shapes and with multiple schemes to allow for charge balance with radionuclide substitutions. Ceramic waste forms are designed to incorporate either actinide or Cs/Sr waste streams. These waste forms are usually polyphase, however, in some cases single-phase ceramics such as zirconolite, monazite, apatite, or NaZr-phosphate can incorporate nearly all the radionuclides (i.e. actinide or Cs/Sr) into a single structure. Unlike glass, crystalline phases can be shown to survive several hundred million years or more in wet, thermal geologic environments.

A major concern with the use of ceramic waste forms is the modification of crystalline structures by irradiation (Weber et al., 2009; Grambow, 2006; Lumpkin, 2006). At high ionisation doses and T, many materials undergo decomposition, phase separation and bubble formation under electron-beam irradiation on laboratory time scales. Single-phase ceramics generally exhibit a crystalline-to-amorphous transformation with volume changes ranging from 5-18%. Because this is a critical concern for the use of ceramic waste forms, classical MD studies have focused on the impact of irradiation on candidate crystalline materials.

Classical potentials used to study the effects of the α-decay process, the collision of α-recoil nuclei with surrounding atoms, and localised displacement cascades (DCs) typically are BMH potentials or potentials that include Buckingham parameters to describe short-range interactions. An example of the latter type of potential was developed for zirconolite ($CaZrTi_2O_7$), a matrix being considered for the long-term confinement of actinides, by Veiller et al. (2001):

$$U_{ij}(r_{ij}) = A_{ij}\exp\left(\frac{-r_{ij}}{p_{ij}}\right) - \frac{C_{ij}}{r_{ij}^6} + \frac{1}{4\pi\varepsilon_0}\frac{Z_i Z_j e^2}{r_{ij}}erfc\left(\frac{r_{ij}}{\eta}\right) \quad (8)$$

In this equation, r_{ij} is the interionic distance, $Z_i e$ and $Z_j e$ are the formal charges of the ions, η=6.7 Å and A_{ij}, p_{ij}, and C_{ij} are adjustable parameters to be determined. The first two terms correspond to the short-range Buckingham potentials and represent both the repulsive interactions due to ion overlap and the attractive van der Waals interaction between them. The last term is a screened Coulombic term where the complementary error function reduces the ionic charges $Z_i e$ and $Z_j e$ as a function of r_{ij}.

The best mineral structure potentials include ion polarisation effects. For example, a shell model includes the electronic cloud linked to the core ion by a spring of constant stiffness. However, DCs can only be modelled using rigid ion model potentials (Veiller, Crocombette, Ghaleb, 2002). An important parameter that affects radiation damage is the threshold displacement energy (E_d) which is the minimum kinetic energy necessary to displace an atom from its equilibrium crystallographic site. It is possible to calculate E_d through classical simulations; however, there are not very many experimental values for E_d except for zircon ($ZrSiO_4$). Different atoms in each crystal phase will have different E_d values based on ionic mass, charge and nearest neighbours. To initiate the DC, one cation in the crystal structure is replaced by U^{4+} or U^{3+}.

Veiller et al. (2002) compared simulation results for zirconolite with those for zircon (Crocombette, Ghaleb, 2001) and UO_2 (Jackson, Huntington, Ball, 1995). A trend that is

consistent with experimental observation is observed by this suite of studies. For zircon, an amorphous core is created around the DC track with no surrounding defects. For UO_2, vacancies and interstitials are created throughout the crystal with no specific defects along the DC trajectory. The behaviour of zirconolite lies between that of zircon and UO_2. The simulated irradiation of zirconolite does not result in as much amorphisation as observed in experiment; nor is the phase transition from monoclinic to fluorite structure reproduced by the simulations. These differences could be due to weaknesses in the empirical potentials, but there are other possible explanations for them as well (Veiller, Crocombette, Ghaleb, 2002).

One disadvantage of using a rigid ion potential, is that the energy or heat generated by the (PKA, cannot be transferred to the electrons in the crystal structure. To compensate for this Phillips and Crozier (2009) developed a two-temperature model to represent the interaction between atoms and free electrons during thermal transients such as radiation damage and cascade simulations. This model communicates energy between electronic and atomic subsystems using a Langevin thermostat, and does not impact the form of the classical potentials. This approach was used to study the effects of cascade damage in a gadolinium pyrochlore ($Gd_2Zr_2O_7$). The simulations showed that the electronic subsystem can act as a heat sink for the PKA and the inclusion of this system reduces the predicted cascade damage in the crystal structure (Ismail et al., 2010). More research is required to more accurately depict the electron-ion interactions during a cascade event.

3.5. Challenges and future direction for development of classical potentials for nuclear materials

Range of application

A well-known limitation of classical potentials is that they are generally only applicable over a small range of temperatures and/or interatomic separations. Even when potentials that are good over different ranges are combined, the way that they are combined also affects the resulting properties (Tiwary et al., 2009). Recent work by Tiwary et al. (2009) has taken advantage of DFT calculations in order to develop a systematic method for producing UO_2 potentials that are good over a large range of interatomic separations. By applying this procedure to MOX fuels and Fe-alloy claddings will lead to better potentials for radiation damage studies.

While ever-increasing computer power has led to routine electronic-scale DFT calculations, these *ab-initio* calculations are usually limited to predicting structures at 0 K. For the case of defects, these configurations may be expected to be different at the high local temperatures due to irradiation. The use of *ab-initio* MD (AIMD) to address this problem, while computationally costly, may uncover some new information about defect structure and therefore allow for better fitting of classical potential parameters.

Great strides have been made to make potentials that are applicable from room temperature to above the melting point of a material, but based on recent reviews, clearly there are different potentials that are best over different temperature ranges (Govers et al., 2007, 2008).

Charge transfer/reactive models

While reactive models are a relatively new invention (even in the relatively short history of classical simulation), their development for nuclear materials will represent a major step forward in the ability to consider defect migration and charge transfer between ions. This

may be particularly useful for the case of UO_2, where U exists in different oxidation states. It will also allow for more accurate surface reaction simulations. Some initial work has been done for a charge transfer model for Fe (Samaras, Victoria, Hoffelner, 2009), which could be used to predict oxygen mobility and oxidation mechanisms. However, additional work is needed to develop new functional forms for interatomic potentials in mixed ionic-covalent systems of interest.

Density functional theory calculations

To an increasing extent, the results of DFT calculations are being used for developing and verifying interatomic potentials. This is especially true for defect configurations and migration energies, where experimental data is lacking or impossible to obtain. However, there are still some limitations to DFT calculations. Because DFT-calculated properties represent the lowest-level in multi-scale simulation schemes, it is extremely important that these limitations be addressed.

The choice of functional has been shown to influence the resulting structure – e.g. local density approximation (LDA) predicts incorrect structure of low temperature Fe and does not predict its magnetism; the use of generalised gradient approximation (GGA) resolves this problem, though it is not immediately clear why this should be so (Becquart, Domain, 2001). The choice of basis set or pseudo potential can also have an effect on some energetic properties of Fe-alloys since core atoms may overlap in cases where SIAs are predicted to be close to each other – in these cases, a more expensive "all electron" method may be required (Becquart, Domain, 2009). Actinides are especially difficult in this case because they contain *f* electrons (Devanathan et al., 2010).

DFT calculations also may suffer from the problem of applicable range. The stability of defect configurations is still difficult to predict—some configurations not included in the original parameterisation cannot be reproduced by a set of potentials (Geysermans, 2008). The applicability of these potentials for cases where DFT results do not exist is still not clear (Malerba et al., 2010), though a recent study of polymer structures far from equilibrium is promising (Mattsson et al., 2010).

Also, inclusion of the magnetic contribution to structure (e.g. in Fe-alloys) could play a role in calculated defect energies since spin-orbital coupling plays a major role in the structure of some nuclear materials. This needs to be taken into account at the quantum mechanical level (Samaras, Victoria, Hoffelner, 2009). Some initial work has already been done (Dudarev, Derlet, 2005; Chiesa et al., 2001).

Lack of appropriate experimental data

Because of the lack of experimental results for primary damage due to displacement cascades, it is difficult to test the reliability of new potentials for Fe and Fe-alloys. A review of some popular Fe potentials shows that the defect population and distribution can be sensitive to the potential (Becquart et al., 2000). Additional *ab-initio* calculations can help to resolve these issues, which have initially been addressed by Tiwary et al. (2009)

Indeed, standard interatomic potentials have been called into question since increases in computing power have allowed for routine *ab-initio* calculations that can better predict defect configurations. Increased use of *ab-initio* calculation results as fitting parameters has led to greater versatility and applicability in EAM potentials (Mendelev, Ackland, 2007).

In many cases, experimental data does not exist and atomic-scale results must be fed into higher-level models in order to be able to make a reasonable comparison. However, this introduction of an additional model makes any discrepancy difficult to reconcile (Becquart, Domain, 2009).

An additional consideration is that experimental data, even when available, may not be directly comparable to simulated results. Simulations are almost always done for systems without impurities, whereas real materials invariably have impurities that can lead to discrepancies in properties even among experimental data (Becquart, Domain, 2009).

3.6. Conclusions

Computer simulation provides an attractive alternative to direct experimental observation for nuclear materials that are difficult to observe experimentally under operating conditions. Classical simulations are used to predict material lifetimes or the long-term microstructure evolution due to exposure to severe conditions experienced in nuclear reactors. In addition, they are used to shed light on the use of glass and ceramic waste forms to contain radioactive waste over extended periods of time. Classical potentials are expressions that describe interatomic interactions. These potentials are fit to experimental data and quantum-based DFT calculations. The capability of each potential to predict material behaviour is often limited by the parameterisation of the potential; i.e. the range of interatomic distances and conditions (T, P) used to fit the potential. Although there is more work to be done to improve these potentials, their use has already served to illuminate the properties and behaviours of nuclear materials, and will continue to contribute to our development of new approaches to enhance enrichment and fuel burnup in nuclear fuels, predict waste-form lifetimes and engineer novel nuclear materials.

References

Abbas, A. et al. (2003), "Molecular Dynamics Study of the Structure and Dynamic Behavior at the Surface of a Silicate Glass", *J. Non-Cryst. Solids*, Vol. 315, pp. 187-196.

Abramowski, M., R.W. Grimes and S. Owens (1999), "Morphology of UO_2", *J. Nucl. Mater.*, Vol. 275, pp. 12-18.

Ackland, G.J. et al. (1997), "Computer Simulation of Point Defect Properties in Dilute Fe-Cu Alloy Using a Many-Body Interatomic Potential", *Philosophical*

Magazine a-Physics of Condensed Matter Structure Defects and Mechanical Properties, Vol. 75, pp. 713-732.

Ackland, G.J. et al. (2004), "Development of an Interatomic Potential for Phosphorus Impurities in Alpha-Iron", *J. Phys.-Condes. Matter*, Vol. 16, S2629-S2642.

Adiga, S.P., P. Zapol, P. and L.A. Curtiss (2006), "Atomistic Simulations of Amorphous Alumina Surfaces", *Phys. Rev. B*, Vol. 74.

Aidhy, D.S., D. Wolf and A. El-Azab (2011), "Comparison of Point-Defect Clustering in Irradiated CeO_2 and UO_2: A Unified View from Molecular Dynamics Simulations and Experiments", *Scripta Materialia*, Vol. 65, pp. 867-870.

Angeli, F., J.-M. Delaye, T. Charpentier, J.-C. Petit, D. Ghaleb, P. Faucon (2000a) "Influence of Glass Chemical Composition on the Na-O Bond Distance: A ^{23}na 3q-Mas Nmr and Molecular Dynamics Study", *J. Non-Cryst. Solids*, Vol. 276, pp. 132-144.

Angeli, F., J.-M. Delaye, T. Charpentier, J.-C. Petit, D. Ghaleb, P. Faucon (2000b), "Investigation of A1-O-Si Bond Angle in Glass by Al-27 3q-Mas Nmr and Molecular Dynamics", *Chem. Phys. Lett.*, Vol. 320, pp. 681-687.

Arima, T. et al. (2005), "Evaluation of Thermal Properties of UO_2 and PuO_2 by Equilibrium Molecular Dynamics Simulations from 300 to 2000 K.", *J. Alloy. Compd.*, Vol. 400, pp. 43-50.

Basak, C B. and A.K. Sengupta and H.S. Kamath (2003), "Classical Molecular Dynamics Simulation of UO_2 to Predict Thermophysical Properties", *J. Alloy. Compd.*, Vol. 360, pp. 210-216.

Becquart, C.S. et al. (2000), "Influence of the Interatomic Potentials on Molecular Dynamics Simulations of Displacement Cascades", *J. Nucl. Mater.*, Vol. 280, pp. 73-85.

Becquart, C.S. et al. (2007), "Atomistic Modeling of an Fe System with a Small Concentration of C", *Computational Materials Science*, Vol. 40, pp. 119-129.

Becquart, C.S. and C. Domain (2001), "Modeling Microstructure and Irradiation Effects", *Metall. Mater. Trans. A-Phys. Metall. Mater. Sci.*, Vol. 42A, pp. 852-870.

Becquart, C.S. and C. Domain (2009), "Molecular Dynamics Simulations of Damage and Plasticity: The Role of Ab Initio Calculations in the Development of Interatomic Potentials", *Philos. Mag.*, Vol. 89, pp. 3215-3234.

Benson, G.C., P.J. Freeman and E. Dempsey (1963), "Calculation of Cohesive and Surface Energies of Thorium and Uranium Dioxides", *J. Am. Ceram. Soc.*, Vol. 46, pp. 43-47.

Björkas, C. and K. Nordlund (2007), "Comparative Study of Cascade Damage in Fe Simulated with Recent Potentials", *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*, Vol. 259, pp. 853-860.

Blonski, S. and S.H. Garofalini (1993), "Molecular Dynamics Simulations of a-Alumina and G-Alumina Surfaces", *Surface Science*, Vol. 295, pp. 263-274.

Buckingham, R.A. (1938), "The Classical Equation of State of Gaseous Helium, Neon and Argon", *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, Vol. 168, p. 264.

Busker, G., R.W. Grimes and M.R. Bradford (2000), "The Diffusion of Iodine and Caesium in the $UO_{2 +/- x}$ Lattice", *J. Nucl. Mater.*, Vol. 279, pp. 46-50.

Busker, G., R.W. Grimes and M.R. Bradford (2003) R., The Solution and Diffusion of Ruthenium in $UO_{2 +/- x}$", *J. Nucl. Mater.*, Vol. 312, pp. 156-162.

Catlow, C.R.A. (1977), "Point-Defect and Electronic Properties of Uranium-Dioxide", *Proc. R. Soc. London Ser. A-Math. Phys. Eng. Sci.*, Vol. 353, pp. 533-561.

Catlow, C.R.A. (1978), "Fission-Gas Diffusion in Uranium-Dioxide", *Proc. R. Soc. London Ser. A-Math. Phys. Eng. Sci.*, Vol. 364, pp. 473-497.

Chiesa, S. et al. (2011), "Optimization of the Magnetic Potential for Alpha-Fe", *J. Phys.-Condes. Matte*r, Vol. 23.

Cormier, L. et al. (2003), "Chemical Dependence on Networktopolgy of Calcium Aluminosilicate Glasses: A Computer Simulation Study", *Journal of Non-Crystalline Solids*, Vol. 332, pp. 255-270.

Crocombette, J.P. and D. Ghaleb (2001), "Molecular Dynamics Modeling of Irradiation Damage in Pure and Uranium-Doped Zircon", *Journal of Nuclear Materials*, Vol. 295, pp. 167-178.

4. Molecular dynamics

Dorothy M. Duffy
Department of Physics and Astronomy, University College London, United Kingdom

4.1. Introduction to molecular dynamics

Molecular dynamics (MD) employs the numerical integration of Newton's equations of motion to calculate the time evolution of a system of interacting atoms. The time and length scales that can be studied using molecular dynamics make it particularly suitable for the calculation of the primary damage produced by an irradiation event. In this paper we describe the MD simulations, known as cascade simulations, used to model radiation damage and the extent to which these simulations have contributed to our understanding of radiation processes and the resulting defect structures. We discuss the limitations of the current models, in particular the accuracy of the interatomic potentials, the limits to the radiation energy association with simulation cell size and the inadequate description of the role of electrons. We discuss how further development in these areas will improve the predictive nature of the models.

4.1.1. Methodology

MD is an atomistic simulation methodology that is based on the numerical integration of Newton's equations of motion. A simulation is initiated by assigning co-ordinates and velocities to a set of interacting atoms. The force on each atom is calculated from the known positions of all other atoms therefore, in principle at least, the positions and velocities of the atoms at a future time can be determined by numerical integration of the equations of motion. A balance between accuracy and simulation time is achieved by choosing an appropriate time step for integration, which is generally a few (1-3) femtoseconds. Efficient integration algorithms, such as the Verlet velocity algorithm, ensure that numerical errors are minimised. The Verlet algoritm for the calculation of the co-ordinate **r** at time t is derived by summing the Taylor expansions of the cordinate **r** at times $t + \Delta t$ and t-Δt, that is:

$$\mathbf{r}(t+\Delta t) = \mathbf{r}(t) + \mathbf{v}(t)\Delta t + \frac{\mathbf{f}(t)}{2m}\Delta t^2 + \frac{\Delta t^3}{3!}\frac{\partial^3 \mathbf{r}}{\partial t^3} + ... \quad (1)$$

$$\mathbf{r}(t-\Delta t) = \mathbf{r}(t) - \mathbf{v}(t)\Delta t + \frac{\mathbf{f}(t)}{2m}\Delta t^2 - \frac{\Delta t^3}{3!}\frac{\partial^3 \mathbf{r}}{\partial t^3} + ... \quad (2)$$

$$\mathbf{r}(t+\Delta t) \approx 2\mathbf{r}(t) - \mathbf{r}(t-\Delta t) + \frac{\mathbf{f}(t)}{m}\Delta t^2 \quad (3)$$

Therefore, the co-ordinates of all the atoms at a time $t + \Delta t$ can be calculated from the co-ordinates at times t and $t - \Delta t$ and the forces at time t ($\mathbf{f}(t)$). The force on each atom (labelled i) is calculated from the derivatives of the potential energy (U) of the system of atoms with respect to the co-ordinate of atom i:

$$\mathbf{f}_i = -\frac{\partial U}{\partial \mathbf{r}_i} \tag{4}$$

Therefore, if the potential energy of the system of atoms, as a function of the co-ordinates, and the initial co-ordinates are known, then the co-ordinates at a future time can be evaluated using equation (3).

4.1.2. Interatomic potentials

The most important parameters that limit the accuracy of MD simulations are the interatomic potentials that describe the interactions between the atoms. In general, the potential energy of a set of interacting atoms can be written:

$$U = \sum_i V_i + \sum_{i,j \neq i} V_{ij} + \sum_{i,j \neq i,k \neq i,j} V_{ijk} + \ldots\ldots\ldots \tag{5}$$

Therefore, in theory at least, the forces on all atoms can be determined if the positions are known. In practice, however, determining the various interactions in equation (5) is challenging. An alternative approach is to calculate U directly using *ab-intio* techniques, such as density functional theory, and to employ the Born-Oppenheimer approximation to calculate forces. Such *ab-intio* MD methods are under continuous development and it is now possible to model a few thousand atoms for a few picoseconds. These system sizes are well below those required for radiation damage therefore we focus here on classical MD in which an approximate potential energy function must be determined *a priori*.

In classical MD the challenge is to find approximation for the potential energy (U in equation (5)) that will retain the important physics but will not be prohibitively time consuming to calculate. Two body interactions, such as the Lennard Jones potential, truncate equation (5) after the first term. This allows for very large systems to be modelled but it only gives a reasonable description for very simple systems, such as the noble gases. The Coulomb interactions between charged atoms are calculated using efficient summations techniques, such as the Ewald summation, to take account of the long-range nature of this interaction. Covalent bonds in polymers and biomolecules are described well by the first three terms in the equation (bond stretching, bond bending and torsion). Interatomic potentials for materials that have wide ranging technological applications have attracted particular attention, with over 30 different potentials being published for silicon (Bazant et al., 1997). Carbon has presented particular challenges and a number of potentials have been developed that attempt to describe the change in the bond hybridisation with environment (e.g. REBO (Brenner, 1990) and the REAXFF (van Duin et al., 2001) potentials). The effective charge on atoms in ionic crystals may also depend on the environment and variable charge potentials have been developed in order account for this effect to some extent (Rappel, Goddard, 1991). The many-body interactions necessary to capture the elastic properties of metallic materials are included by embedded-atom type potentials (Daw, Baskes, 1992) such as the successful Finnis-Sinclair model (Finnis, Sinclair, 1984) which includes a term related to the local density of the atoms.

In spite of dramatic improvements in potentials over the years, it is not generally possible to ensure that any potential will describe all properties of a material equally well. Potentials are derived by fitting to experimental data or *ab-initio* results and the ability of a potential to predict a property that was not used in the fitting procedure (known as the transferability) is the signal of a good potential. The potential chosen for a particular simulation will depend on which properties make the dominant contribution to the features to be studied.

4.1.3. Time/length scales

The time and length scales that can be modelled using MD are strongly dependent on the complexity of the interatomic potentials. A record breaking one trillion atom simulation has been performed on the Livermore's Bluegene/L computer but this was to demonstrate a capability rather than to extract results from the simulation (Germann, Kadau, 2008). Typical large-scale simulations range from a few hundred thousand atoms to a few million and typical time scales range from a few tens of picoseconds to a few nanoseconds. Many-body potentials for metals are efficient to calculate but the long-range Coulomb interactions in ionic materials are more time consuming. Potentials that attempt to model the effects of chemical environment (such as REAXFF and REBO potentials) result in relatively slow simulation times therefore, as always, a balance needs to be found between accuracy and efficiency.

4.2. Modelling radiation damage using molecular dynamics

4.2.1. Advantages

Early models of radiation damage used the binary collision approximation (BCA), in various forms to estimate the number of Frenkel pairs (N_{KP}) created by a radiation event with a particular energy (E). The first of these, the Kinchin Pease (KP) model, was able to derive a simple formula (equation (2)) by making a number of assumptions (Kinchin, Pease, 1955).

$$N_{KP} = \frac{E}{2E_d} \tag{6}$$

Here E_d is the displacement energy, which is the minimum energy required to displace an atom from its lattice site to a neighbouring interstitial site to produce a stable Frenkel pair. It is clear that the KP model overestimates the number of defects because it neglects the recombination events that inevitably occur. An improved form (the Norgett−Robinson−Torrens model) was suggested by Norgett (1975) which considers only the fraction of the energy not lost to inelastic collisions with electrons and takes some account of recombination. More sophisticated BCA models, based on numerical simulations, have been developed that include accurate scattering potentials and electronic energy loss (Robinson, Torrens 1974; Ziegler 2004). However, binary collision models are only accurate for situations in which the collisions between the moving atoms and the atoms of the host crystal are spatially separated, which is only the case for light energetic atoms. In the general case, the moving atoms interact with many host atoms simultaneously and it is difficult to predict the number of defect pairs created by the resulting correlated atomic motion.

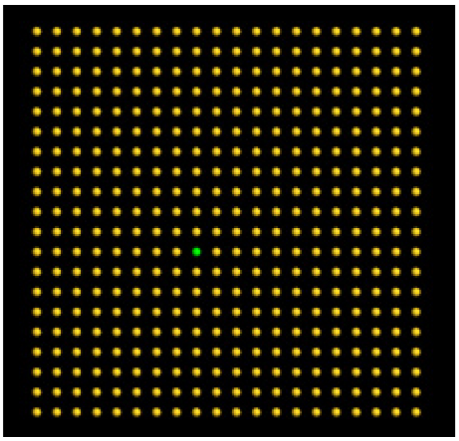
The correlated atomic motion resulting from radiation damage events can be successfully modelled using MD. MD is significantly more demanding, from a computational point of

view, than BCA models, nevertheless, the length and time scales of the primary damage events of radiation damage can readily be modelled. Primary damage refers to the defects remaining after the excess energy imparted by the radiation event has dissipated and the local temperature has returned to the ambient value, which typically occurs around a few tens of picoseconds after the radiation event. The spatial extent of the primary damage is strongly dependent on the mass and energy of the moving atom. Heavy atoms rapidly lose energy to the host crystal and produce localised damage whereas light atoms, such as helium, have a low cross section for nuclear collisions and travel long distances before stopping.

4.2.2. Methodology

A cascade simulation is a particular type of MD simulation that is used to model radiation damage in materials. The initial configuration is an arrangement of atoms with the equilibrium crystal structure and random velocities, corresponding to the simulation temperature, assigned to the atoms. One atom (the primary knock-on atom (PKA)) in the simulation cell is assigned a velocity corresponding to the kinetic energy (the PKA energy) of radiation event to be modelled (Figure 4.1.). The MD simulation is then evolved using a standard integration algorithm until the PKA energy has dissipated and the simulation temperature converges to a constant value.

Figure 4.1. A schematic representation of a cascade simulation.



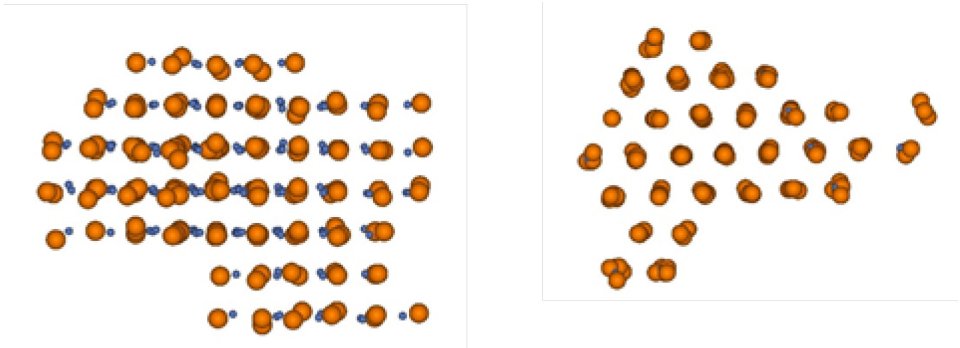
Note: The PKA (green) is assigned a velocity (white arrow) corresponding to the required PKA energy at the start of the simulation.
Source: Duffy, 2019.

Atoms are displaced from their lattice positions during the simulation but many will return to an alternative lattice site as the energy dissipates. Not all atoms return to lattice sites however. Some remain in interstitial positions, leaving a corresponding number of vacant lattice sites. The identification of defects can be carried out by counting the number of atoms in each Wigner Seitz cell of the lattice. An empty cell signals a vacancy and a cell with two atoms signals an interstitial. Alternatively, a lattice site can be labelled as vacant if there are no atoms within a predefined radius (r_o) and an atom is labelled as an interstitial if it is further than r_o from any lattice site. The former (Wigner Seitz) method provides a unique method for defect counting whereas the second method is sensitive to the value of r_o chosen as the cutoff. The second method does, however, capture details about the defect structure that are missed by the Wigner Seitz method. Split interstitials (dumb-bells),

crowdion clusters and stacking fault tetrahedra are examples of common defect clusters that are created by cascade simulations of radiation damage in metals. A crowdion is an extended interstitial, in which an additional atom is inserted in a close packed direction in a metal. This has the effect of displacing several atoms, located along the close packed direction, from their lattice sites. An example of a 30 crowdion cluster is shown in Figure 4.2.

Cascade simulations require a number of features which are not always available in standard MD programmes. One useful feature is a variable time step. The highly non-equilibrium nature of a cascade means that in a typical MD time step of 1 fs a 10 keV Fe atom would move a distance of 0.35 Å, which will result in a large change in the force and possibly a failed simulation. To avoid these effects it is necessary to use a very short simulation time step (possibly as low as 0.001 fs) in the initial stages of a high-energy cascade simulation. As the simulation progresses the velocity of the PKA atoms falls rapidly therefore progressively longer time steps can be employed. A variable time step algorithm imposes to limit the distance moved during one-time step by identifying the maximum velocity of the ensemble of atoms and choosing an appropriate time step.

Figure 4.2. Top view (a) and side view (b) of a crowdion cluster oriented in the <011> direction in a body centred cubic metal.



(a) (b)

Note: The cluster has 170 interstitials (large orange spheres) and 140 vacancies (small blue spheres). Each crowdion has one excess interstitial.
Source: Duffy, 2019.

A further consideration for cascade simulations is how the excess energy of the PKA atom is dissipated during the simulation. Standard thermostats are inappropriate because they calculate the temperature from the average velocity of the atoms. The average velocity can be extremely high in the early stages of a simulation therefore the excess energy will be rapidly removed by the thermostat and the cascade will be quenched. There are two alternatives: either the total energy is conserved (using a constant energy NVE simulation), which will result in a temperature rise as the PKA energy dissipates through the cell, or the excess energy is removed via a boundary thermostat, which rescales the velocities of the atoms that reside within a few Å of the cell boundary. Both methods are commonly used.

As with all MD simulations, the ability of the interatomic potentials to reproduce the effect of atomic interactions in real materials is crucial to the accuracy of the results. The challenges associated with interatomic potentials will be discussed in the next section but here we discuss an issue that concerns cascade simulations in particular – that of close encounters between atoms due to the high velocity of some atoms. Conventional

interatomic potentials are highly inaccurate at close interatomic separations as they use near equilibrium conditions for fitting procedures. For cascade simulations it is necessary to modify the short-range interaction potential to give a more realistic description of forces at close separations and this is achieved by replacing the chosen potential with a screened Coulomb potential, known as the Ziegler-Biersack-Littmark (ZBL) potential ($V(r)$; equations (7)-(9), at close separations (Biersack and Ziegler 1982).

$$V(r) = \frac{Z_1 Z_2 e^2}{4\pi\varepsilon_0 r}\phi(\frac{r}{a}) \tag{7}$$

$$\phi(\upsilon) = 0.1818e^{-3.2\,y} + 0.5099e^{-0.9423\,y} + 0.2802e^{-0.4029\,y} + 0.2817e^{-0.2016\,y} \tag{8}$$

$$a = \frac{0.8853a_0}{Z_1^{0.23} + Z_2^{0.23}} \tag{9}$$

Here r is the interatomic separation, a_0 is the Bohr radius and Z_1 and Z_2 are the atomic numbers of the interacting atoms. The ZBL potential is joined to the standard potential by a spline function to ensure that the potential, and its derivatives, are continuous.

4.2.3. Results

Metals

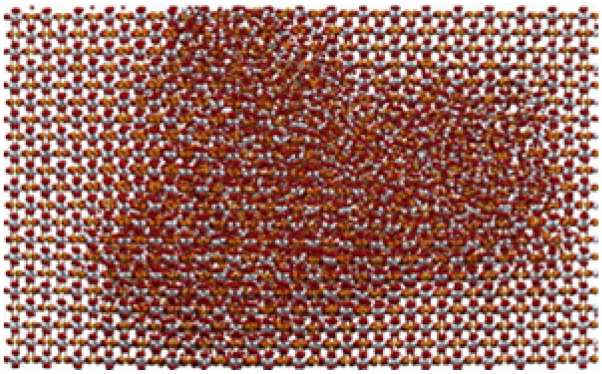
MD has been an invaluable tool for studying radiation damage in metals since the publication of the classic paper by Gibson et al. (1960). Early simulations were restricted to small simulation cells and, consequently, low PKA energies. Nevertheless, interesting and important processes such as channelling, where atoms travel long distances between lattice planes, and replacement collision sequences (RCS) were identified. In an RCS, a moving atom makes a direct collision with a neighbouring atom along a close packed crystallographic direction. The initial atom occupies the neighbouring site and sets the neighbour in motion along the same direction so the sequence is repeated until the energy is dissipated. Such sequences are effective mechanisms for generating well separated vacancy-interstitial pairs with relatively low energy radiation events.

As computer power increased and the accuracy of interatomic potentials improved, the power of cascade simulations increased dramatically. The relationship between the number of Frenkel pairs and the PKA energy was established for several metals and compared to the NRT model (see, for example, de la Rubia, 1996; Stoller et al., 1996, 1997; Stoller, 2012). The structures of defect clusters created by high-energy radiation damage were identified. Dislocation loops, formed by the collapse of interstitial clusters, were observed and stacking fault tetrahedra were observed in face centred cubic (fcc) metals. The formation of defect clusters by radiation damage makes a significant contribution to the change in mechanical properties as these clusters may be more or less mobile than isolated defects. Vacancy clusters may also be nucleation sites for bubbles, as they act as trapping sites for gases created by transmutation reactions. Defect clusters may contribute to radiation embrittlement because they act as obstacles to dislocation motion. Indeed, the barrier strength of different types of defect cluster, a parameter that is required for dislocation dynamics simulations, can be calculated using molecular dynamics. Swelling, which occurs due to the differential volume relaxation between vacancies, can also be calculated using MD simulations.

Ceramics

Ceramic materials are employed for a wide range of nuclear applications, including nuclear fuel and encapsulation materials for nuclear waste disposal. MD has made significant contributions to the understanding of the fundamental processes in these materials and the radiation effects and issues are generally quite different than those in metals. Embrittlement is not a concern, as ceramics are intrinsically brittle, but swelling can be problem as the resulting internal stresses may cause cracking. Amorphisation occurs readily in some ceramic materials and this has been investigated using cascade simulations. The degree of covalency has been identified as having a strong influence on the susceptibility of ceramic materials to amorphisation (Trachenko, 2006). An example of an amorphised region of zircon, formed by a 30 keV U PKA cascade simulation is shown in Figure 4.3. (Devanathan, 2009). The complex crystal structures of many nuclear ceramics means that a rich variety of possible defects structures are possible and MD simulations can help to characterise these defects (Zinkle et al., 2002).

Figure 4.3. An amorphised region of zircon, formed by a 30 keV U PKA cascade simulation



Source: Devanthan, 2009.

4.3. Limitations/challenges

4.3.1. Simulation size

The PKA energy that can be modelled by MD depends on the number of atoms in the simulation cell as a sufficiently large cell must be employed to avoid self-interaction effects. The PKA, and any subcascades it creates, must be confined within the simulation cell. The energy deposited in the cell by the cascade event thermalises by atomic collisions and diffuses to the boundaries of the simulation cell, where it can re-enter the cell through the periodic boundary and affect the results. A boundary thermostat, which dissipates energy in the outer layers of the simulation cell, inhibits energy diffusion across the periodic boundary and permits smaller simulation cells to be used. As a rough guide, around 5 -10 $\times 10^4$ atoms are required per keV of PKA energy but this number depends, to some extent, on the material being modelled. Recent simulations of 0.5 MeV cascades in Fe used 5×10^8 atoms in the simulation cell and were performed on 6×10^4 parallel processors (Zarkadoula et al., 2013). These simulations are particularly relevant to fusion technology as 0.5 MeV is close to the average recoil energy from a 14.1 MeV neutron, produced by deuterium-tritium fusion, in Fe.

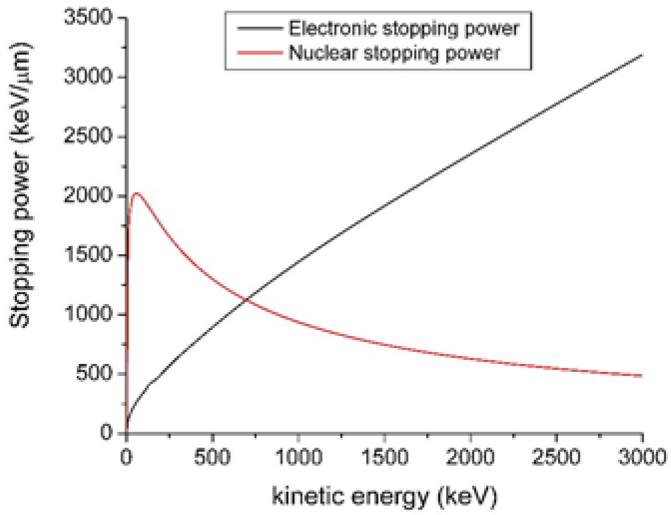
4.3.2. Interatomic potentials

As discussed in the introduction, the results obtained from any MD simulation are only as reliable as the interatomic potential employed. For cascade simulations it is important that the potentials reproduce the melting temperature and thermal expansion coefficient in order to obtain a realistic response of the material to the large energy density associated with the radiation event. The displacement energies and defect properties (energies and mobilities) are also very important as these will influence the number of defects produced and the degree of defect clustering and annihilation. The importance of steel, both ferritic and austenitic, as a structural material for both fission and fusion reactors has led to the development of many interatomic potentials for iron and detailed comparisons between the results obtained from cascades using different potentials (Becquart et al., 2000; Malerba 2006; Terentev et al., 2006; Malerba 2010). The conclusions from these studies were that the number of residual defects produced was largely potential independent, at least for the most recent potentials that focused on fitting parameters that are relevant to radiation damage. However, details of the damage formation and, in particular, the fraction of defects that formed part of clusters did depend on the choice of potentials. Even the method by which the equilibrium potential was joined to the high-energy ZBL potential seemed to affect the details of the cascade evolution.

4.3.3. Electronic effects

Another concern about cascade simulations is that the electrons, and the effects they have on damage production, are largely neglected. In effect, the electronic degrees of freedom are integrated out of the problem and included only at the level of the interatomic potentials. However, an atom moving through a material loses energy via both nuclear collisions, which are described well by MD, and inelastic collisions with electrons. A particular example of the relative stopping power (the energy loss per unit distance) for an Fe atom moving in an Fe lattice, calculated using the SRIM code, is shown in Figure 4.4. It is clear from the figure that nuclear loss dominates at low energies but inelastic (electronic) losses increase as the energy of the moving atom increases.

Figure 4.4. Plots of the variation of electronic and nuclear stopping powers for an Fe ion moving in Fe



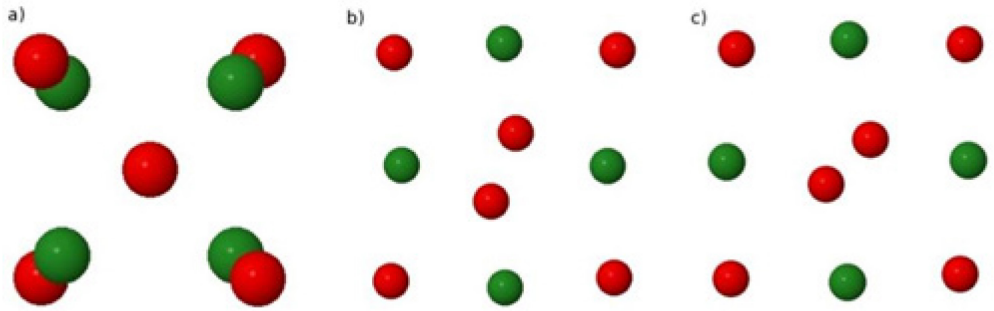
Note: At low energies the nuclear losses dominate and at high energies the electronic losses dominate.
Source: Duffy, 2019.

The energy loss resulting from inelastic collisions in metals is often included as a friction term that operates above some cutoff velocity (Nordlund et al., 1998). This, however, does not take account of what happens to the energy that is deposited in the electrons. For weak coupling between the lattice and the electrons, and a high electronic thermal conductivity, the energy is transported rapidly from the cascade region and it should not affect the residual damage (Flynn, Averbach, 1988, Finnis et al., 1991). However, if the coupling between the lattice and the electrons is strong and the electronic thermal conductivity is low, the electronic energy can be confined close to the cascade for sufficient time to be redistributed back to the lattice. In this case, the residual defect number can either be enhanced or reduced depending on whether the electronic effects quench the cascade or contribute to defect annealing. This effect has been demonstrated (Rutherford, Duffy, 2007) using a methodology that couples the effective electronic temperature to an MD simulation of the lattice (Duffy, Rutherford, 2007). A comprehensive review of the electronic effects in radiation damage has been produced by Race et al., (2010) and the introduction of these effects in radiation damage simulations is reviewed by Darkins and Duffy (2018).

The effects of excited electrons are even more significant in band gap materials than they are in metals. Cations in some metal oxides, including UO_2, exist in a number of different charge states. Electrons excited to the conduction band by radiation, and the corresponding holes, become mobile and thus they may get trapped at lattice defects, possibly created by other damage events. Trapped electrons will substantially modify the defect's conformation and mobility. Figure 4.5. shows the calculated relaxed conformation of an O interstitial in MgO with a net charge of -2e, -1e and 0. The O^- interstitial was found to have a very low (~0.06 eV) migration barrier, which suggests that the interstitial with one trapped hole will be extremely mobile (Mulroue, Duffy 2011). As the defect mobility dominates the microstructure evolution, the electronic defects will have a strong influence on the evolving material properties. Electrons, holes and excitons can also distort the lattice to such an extent that they become trapped by their own distortion field (self-trapped) and the decay

of these self-trapped carriers can result in energy redeposition to the lattice, or even defect formation (Shluger, Stoneham 1993). Thus, it is clear that modelling radiation damage in band gap materials using cascade simulations and classical potentials neglects several effects that could strongly influence the results (Klaumünzer, 2006; Duffy et al., 2012).

Figure 4.5. The relaxed defect configuration for oxygen interstitials in MgO.



Note: The interstitials have a net charge of a) -2e (cell centre configuration) b) -1e (<011> oriented dumb-bell) and c) 0 (<111> oriented dumb-bell).
Source: Duffy, 2019.

4.4. Future developments

The future requirements for MD simulations of radiation damage follow from the limitations discussed in the previous section. Advances in computer power and the development of more efficient algorithms, particularly for reading/writing large datasets, will extend the range of PKA energies that can be modelled, although simulations with PKA energies higher than 500 keV are unlikely to become routine in the near future. Interatomic potential development and evaluation is an area where much effort is required. The KIM (Knowledge base of Interatomic Models) project (Tadmor et al., 2011) should make an important contribution to interatomic potential development and evaluation in the next few years. Much more needs to be done on developing potentials for real engineering alloys although past experience in the development of potentials for Fe/Cr alloys (Olsson et al., 2005; Caro et al., 2005), that aim to describe the observed phase behaviour, reveals just how challenging this can be. It is important that the phase behaviour is correctly described as radiation enhanced diffusion can accelerate the precipitation of stable intermetallic particles. Radiation enhanced segregation and depletion at grain boundaries are important effects that rely on accurate mixed potentials. Oxide dispersion strengthened (ODS) steel offer a potential solution to the low high temperature strength of conventional ferritic steel but modelling radiation effects in ODS steel will require potentials for the oxide ($YTiO_3$) nanoparticles and, even more challenging, the interaction between the oxide and the host metal. It is clear there is much yet to be done in the field of potential development before the full potential of modelling radiation effects using MD can be realised.

A further area where MD could make significant contributions is in the examination of the effect of microstructure on radiation damage. There is evidence from cascade simulations in nanocrystalline metals (Samaras et al., 2002) and from temperature accelerated dynamics (Bai et al., 2010) that grain boundaries enhance defect recombination and there is speculation that the trapping of defects at oxide metal interfaces in ODS steel will also

enhance defect recombination and reduce radiation damage. However, there has been only a limited number of simulations to date and a more comprehensive study of the effect of grain boundaries and other interfaces would help towards the extrapolation to engineering materials.

A related, but just as significant, issue is the effect of radiation damage on the microstructure of materials. This goes beyond the production of defect clusters and bubbles that contribute to embrittlement by interacting with dislocations. It includes effects such as precipitation of stable phase particles from supersaturated solution and the dissolution of metastable phase particles. Alloy steels contain a wide range of elements to enhance properties and aid processing but these additives can form a number of intermetallic phases that may precipitate with the help of radiation enhanced diffusion. The microstructure evolution in nuclear fuel is an extreme example of how radiation, fission gas production and high thermal gradients can transform the microstructure of a ceramic material.

As discussed above, the electronic effects associated with radiation damage are not included in classical MD simulations. The topic attracted attention around 20 years ago but it has largely been neglected since, although there has recently been revived interest in the topic. The development of *ab-initio* techniques, such as time-dependent density functional theory and time-dependent tight binding (Mason et al., 2007) , has helped in the understanding of excited states, but how such excitations affect the radiation damge on the timescales associated with MD has largely been unexplored. This field is now ripe for development.

Validation of a computational model by comparison with experimental measurements is essential for maximum confidence in the results but this is generally challenging due to the widely different length and time scales involved. The situation is improving due to the development of in situ transmission electron microscopy where samples can be observed by TEM during ion bombardment and this enables the development of the microstructure to be directly observed. The JANNUS multibeam experimental facility (Serruys et al., 2009) is one example of a facility that is devoted to promoting interaction between modelling and experiment.

In summary, the time and length scales that can be studied using classical MD simulations make it suitable for the calculation of the primary damage produced by an irradiation event. However, more research and development are required, particularly in areas of interatomic potentials and the description of electronic effects, to improve the predictive nature of the models.

References

Bai, X.M. et al. (2010), “Efficient Annealing of Radiation Damage Near Grain Boundaries via Interstitial Emission”, *Science*, Vol. 327, pp. 1631-1634.

Bazant, M.Z., E. Kaxiras and J.F. Justo (1997), “Environment-dependent interatomic potential for bulk silicon”, *Phys. Rev. B*, Vol. 56, pp. 8 542.

Becquart, C.S., et al. (2000), “Influence of the interatomic potentials on molecular dynamics simulations of displacement cascades”, *J. Nucl. Mater*, Vol. 280, pp. 7-85.

Biersack J.P. and J.F. Ziegler (1982), “Refined universal potentials in atomic-collisions”, *Nucl. Instrum. Meth. B,* Vol. 141, p. 93.

Brenner, D.W. (1990), "Empirical potential for hydrocarbons for use in simulating the chemical vapor-deposition of diamond films", *Phys. Rev. B*, Vol. B 42, p. 9 458.

Caro, A., D.A. Crowson and M. Caro (2005), "Classical many-body potential for concentrated alloys and the inversion of order in iron-chromium alloys", *Phys. Rev. Lett*., Vol. 95, 075702.

Darkins, R. and D.M. Duffy (2018), "Modelling Radiation Effects in Solids with Two-Temperature Molecular Dynamics", *Computational Materials Science*, Vol. 147, pp. 145–153.

Daw, M.S. and M.I. Baskes (1984), "Embedded-atom method - derivation and application to impurities, surfaces, and other defects in metals",*Phys. Rev. B*, Vol. 29, pp. 6 443–6 453.

Devanathan, R. (2009), "Radiation damage evolution in ceramics", *Nucl. Inst. Meth. B*, Vol. 267, pp. 3 017–3 021.

Diaz de la Rubia, T. (1996), "Irradiation-induced defect production in elemental metals and semiconductors: a review of recent molecular dynamics studies", *Annu. Rev.Mater. Sci*., Vol. 26, pp. 613–49.

Duffy, D.M., S.L. Daraszewicz and J. Mulroue (2012), "Modelling the Effects of Electronic Excitations in Ionic-Covalent Materials", *Nucl. Instr. Meth. B*, Vol. 277, pp. 21-27.

Duffy D.M. and A.M. Rutherford (2007), "Including the effects of electronic stopping and electron-ion interactions in radiation damage simulations", *J. Phys.: Condens. Matter*, Vol. 19, 016207.

Finnis M.W., P. Agnew and A.J.E. Foreman (1991), "Thermal excitation of electrons in energetic displacement cascades", *Phys. Rev. B*, Vol. 44, pp. 567–74.

Finnis, M.W. and J.E. Sinclair (1984), "A simple empirical n-body potential for transition-metals",*Phil. Mag. A*, Vol. 50, pp. 45-55.

Flynn, C.P. and R.S. Averback (1998), "Electron–phonon interactions in energetic displacement cascades", *Phys. Rev. B,* Vol. 38, pp. 7118–20.

Germann, T.C. and K. Kadau (2008), "Trillion-atom molecular dynamics becomes a reality", *Inter. J. of Mod.Phys. C*, Vol. 19, pp. 1315-1319.

Gibson J.B. et al. (1960), "Dynamics of radiation damage", *Phys. Rev., Vol.* 120, pp. 1229–53.

Justo, J.F. et al. (1998), "Interatomic potential for silicon defects and disordered phases", *Phys. Rev. B*, Vol. 58, p. 2 539.

Kinchin, G.H. and R.S. Pease (1955), "The displacement of atoms in solids by radiation", *Rep. Prog. Phys., Vol.* 18, pp. 1–51.

Klaumünzer, S. (2006), "Thermal-spike models for ion track physics: A critical examination",*Mat. Fys. Medd. Dan. Vid. Selsk*., Vol. 52, p. 293.

Malerba L. (2006), "Molecular dynamics simulation of displacement cascades in a-Fe: A critical review", *J. Nucl. Mater*., Vol. 351, pp. 28–38.

Malerba, L. et al. (2010), "Comparison of empirical interatomic potentials for iron applied to radiation damage studies", *J. Nucl. Mater.*, Vol. 406, pp. 19-38.

Mason, D.R. et al. (2009), "Electronic damping of atomic dynamics in irradiation damage of metals", *J. Phys: Conden. Matt*., Vol. 19, 436209.

Mulroue, J. and D.M. Duffy (2011), "An ab initio study of the effect of charge localisation on oxygen defect formation and migration energies in magnesium oxide", *Phil. Trans. Roy. Soc. A*, Vol. 467, pp. 2 054.

Nordlund, K. et al. (1998), "Defect production in collision cascades in elemental semiconductors and fcc metals", *Phys. Rev. B,* Vol. 57, pp. 7556–70.

Norgett, M.J., M.T. Robinson and I.M. Torrens (1975), "A proposed method of calculating displacement dose rates, *Nucl. Eng. Des.*, Vol. 33, pp. 50–4.

Olsson, P. et al. (2005), "Two-band modeling of alpha-prime phase formation in Fe-Cr", *Phys. Rev.* Vol. B 72, 214119.

Race, C.P. et al. (2010), "The treatment of electronic excitations in atomistic models of radiation damage in metals", *Rep. Prog. Phys*., Vol. 73, 116501.

Rappel, A.K. and W.A. Goddard III (1991), "Charge Equilibration for Molecular Dynamics Simulations", *J. Phys. Chem.*, Vol. 95, pp. 3358-3363.

Robinson M. T. and I.M. Torrens (1974), "Computer simulation of atomic-displacement cascades in solids in the binary-collision approximation", *Phys. Rev. B, Vol.* 9, pp. 5008–24.

Rutherford, A. M. and D.M. Duffy (2007), "The effect of electron-ion interactions on radiation damage simulations", *J. Phys.: Condens. Matter*, Vol. 19, 496201.

Samaras, M. et al. (2002), "Computer simulation of displacement cascades in nanocrystalline Ni", *Phys. Rev. Lett*. Vol. 88, 125505.

Serruys, Y. et al. (2009), "JANNUS: A multi-irradiation platform for experimental validation at the scale of the atomistic modelling", *J. Nucl. Mater.,* 386-388, 967-970.

Shluger, A.L. and A.M. (1933), "Small polarons in real crystals - concepts and problems", J. Phys.: Condens. Matter, Vol. 5, pp. 3049-30.

Stoller, R.E. (1996), "Point defect survival and clustering fractions obtained from molecular dynamics simulations of high energy cascades", *J. Nucl. Mater*., 233-237, 999-1003.

Stoller, R.E., G.R. Odette, B.D. Wirth (1997), "Primary damage formation in bcc iron", *J. Nucl. Mater*., Vol. 251, pp. 49-60.

Stoller, R.E. (2012), "1.11 – Primary Radiation Damage Formation", in *Comprehensive Nuclear Materials*, Vol. 1, pp. 293-332.

Tadmor, E.B. et al. (2011), "The potential of atomistic simulations and the knowledgebase of interatomic models", *JOM,* Vol. 63, p. 17.

Terentyev, D. et al. (2006), "Effect of the interatomic potential on the features of displacement cascades in a-Fe: A molecular dynamics study", *J. Nucl. Mater.,* Vol. 351, pp. 65–77.

Trachenko, K. et al. (2006), "Atomistic simulations of resistance to amorphization by radiation damage", *Phys. Rev. B*, Vol. 73, 174207.

Van Duin A.C.T. et al. (2001) "ReaxFF: A Reactive Force Field for Hydrocarbons", *J. Phys. Chem. A*, Vol. 105, pp. 9396-9409.

Zarkadoule, E. et al. (2013), "The nature of high energy radiation damage in iron", *J. Phys.: Conden. Matt*., Vol. 25, 125402.

Ziegler J. F. (2004), "SRIM—The stopping and range of ions in matter", *Nucl. Instr. and Meth. B,* 219-220, 1027.

Zinkle, S.J., V.A. Skurato and D.T. Hoelzer (2002), "On the conflicting roles of ionizing radiation in ceramics", *Nucl. Inst. Meth. B,* Vol. 191, p. 758.

5. Atomistic simulation methods for long-time dynamics in materials for nuclear energy systems

B. P. Uberuaga[1], D. Perez[2] and A. F. Voter[2]

[1]Materials Science and Technology Division, Los Alamos National Laboratory, United States

[2]Theoretical Division, Los Alamos National Laboratory, United States

5.1. Introduction

Many important processes in materials systems are intrinsically atomistic in nature but involve time scales that span many orders of magnitude, thus exceeding what can be directly simulated using molecular dynamics (MD). This is especially true for materials in nuclear energy applications, in which defects created by collision cascades on the femtosecond-picosecond time scale cause microstructural changes that continue to evolve for years, in many cases leading to failure of the material. In this chapter, we review atomistic methods for reaching long-time scales in systems like these. These accelerated MD and adaptive kinetic Monte Carlo (aKMC) methods exploit the infrequent-event nature of the diffusive events that comprise this long-time evolution. In favourable cases, these methods can predict state-to-state evolution that approximates what would result from an extremely long molecular dynamics simulation, and the most accurate of the methods can do this to arbitrary accuracy. We present some examples of applications of these methods to problems relevant to nuclear energy materials, the subject of this volume. We then discuss situations that are difficult for the methods, causing them to be less efficient, and we conclude with a short list of the most pressing issues in the further development of these approaches to make them as powerful and predictive as possible for realistic problems.

The evolution of radiation damage in materials spans many time and length scales. While the initial damage production occurs on the atomic scale over picoseconds via collision cascades, the damage ultimately manifests itself macroscopically, often in the form of swelling or cracking, which can take years to develop. There is a wide range of phenomena that bridge these two extremes, including defect diffusion, annihilation and aggregation, the formation of interstitial loops and voids, and the development of a more complex microstructure. As a result, no one simulation method can be employed to study the problem of radiation damage on all relevant time and length scales. Rather, a combination of many techniques must be used to address this problem. MD simulation, in which atom positions are evolved by integrating the classical equations of motion in time, can probe timescales of ps to ns, making it ideal for studying collision cascades. The initial, post-thermal-spike damage produced in the collision cascade can be directly simulated using MD. However, once that damage has been formed, diffusion and subsequent annihilation or aggregation of those defects can occur on much longer time scales, perhaps even seconds or beyond, depending on the conditions (temperature, pressure, etc.) and the material. Such phenomena must be accounted for to accurately predict the formation and evolution of larger scale features such as interstitial dislocation loops and vacancy voids. These are important in the evolving microstructure, leading to the macroscopic response to radiation damage such as swelling and cracking. While at some larger time and length scale, the material evolution can be described by higher-level models, for much longer times than the sub-microsecond time scale that direct MD can reach, full atomistic simulation is necessary

to understand the complex and often competing phenomena that must be accounted for and quantified to parameterise these higher-level models.

In this chapter we discuss some methods developed over the past 20 years for reaching these longer time scales without sacrificing atomistic detail or dynamical accuracy. These long-time methods work for *infrequent-event systems*, in which the system spends a long period of time in a given state (typically corresponding to a particular basin of the potential energy surface), occasionally making a transition to a neighbouring state (e.g. through the motion of a defect). These transitions are usually activated processes, whose rate is limited by a potential energy barrier. The characteristic that the state residence time is much longer than the time duration of the transition event (typically 1 ps) is what defines these as infrequent-event systems.

These long-time methods are most valuable for treating what we often term “complex infrequent event systems," and many if not most radiation damage systems fall into this class. These are characterised by an interplay of reaction pathways with different activation energies (and hence different temperature dependences), so that one cannot simply raise the temperature of the system to accelerate the dynamics in a meaningful way, and transition pathways that are complicated enough that many important pathways might be omitted if one resorts to a simple model (e.g. lattice-based kinetic Monte Carlo (KMC) with a pre-calculated rate catalogue). The point of these long-time-dynamics methods is that they do not make assumptions about transition pathways or the nature of the states, assumptions that would limit their accuracy. Consequently, these methods give dynamical evolution from state to state that is roughly the same as that of a very long direct-MD simulation. For the most accurate of these methods, parallel replica dynamics, the state-to-state evolution can be made arbitrarily accurate.

These methods generally fall into two classes: accelerated molecular dynamics (AMD) methods and aKMC approaches. In the AMD approach, the key concept is to let the trajectory find its own way out of the current state of the system – as a trajectory in a standard MD simulation would – but to coax it into to finding this escape pathway more quickly. The methods in this class are parallel replica dynamics (Section 5.3) hyperdynamics (Section 5.4), and temperature accelerated dynamics (Section 5.5). The aKMC approach, described in Section 5.6, is a generalisation of standard KMC. The goal in aKMC is to find all possible, or all relevant, escape pathways from the current state of the system, typically by performing a set of saddle-point searches. One of these pathways is then selected to move the system to the next state. In Section 5.7. we briefly describe a very recently developed method in this long-time simulation class, κ-dynamics.

After presentation of the methods and a review of some recent advances in Sections 5.3. through 5.7., in Section 5.8. we give some examples of applications of these methods to problems relevant for nuclear materials. We note that some of the method presentation draws from another review we wrote recently (Perez et al., 2009). We then conclude in Sections 5.9. and 5.10. with a discussion of the current limitations of these methods and the areas in which we believe further development is needed to make the methods maximally useful for simulating atomistic evolution in nuclear energy materials.

Due to space constraints, the descriptions of the methods that follow omit many foundational details. The interested reader is encouraged to consult the original references (e.g. Voter, 1998, 1997a, 1997b; Sørensen, Voter, 2000; Henkelman, Jónsson, 2001, Lu, Makarov, Henkelman, 2010) or more in-depth review articles (e.g. Uberuaga et al., 2005a; Voter, 2006) for a more thorough treatment. The reader can also find some more discussion

of the relative merits of the methods and how to choose the best method for a particular problem in Uberuaga and Voter (2006).

Finally, we note that the manuscript for this chapter was prepared in 2013, but not published until 2020. We have left the main body of the chapter largely unchanged, as we believe it still represents a valuable introduction to the methods and review of how they can be applied to nuclear-energy materials. We have added a short section near the end of the chapter that describes the important developments in the past five years.

5.2. Parallel replica dynamics

Parallel replica dynamics (ParRep) is the most general and accurate of the AMD methods. It assumes only that the system will escape from the current state in a way that obeys first-order kinetics; i.e. for any trajectory that has been in a state long enough to have lost its memory of how it entered the state (longer than the correlation time τ_{corr}), the probability distribution function for the time of the next escape from that state is given by

$$p(t) = ke^{-kt}, \tag{1}$$

where k is the rate constant for escape from the state. ParRep allows for the temporal parallelisation of the state-to-state dynamics of such a system on M processors. This is to be contrasted with standard parallelisations of MD simulations in which spatial decomposition schemes are used.

We briefly sketch the derivation here. For a state in which the total rate of escape is k, simultaneously explored on M processors, the effective escape rate for the first escape of *any* replica is Mk. If the simulation time accumulated on one processor is t, the total time on the M processors will then be $t_{sum} = Mt$. Thus, using a simple change of variable, $p(t)$ can be written as

$$p(t)dt = Mke^{-Mkt}dt \tag{2a}$$

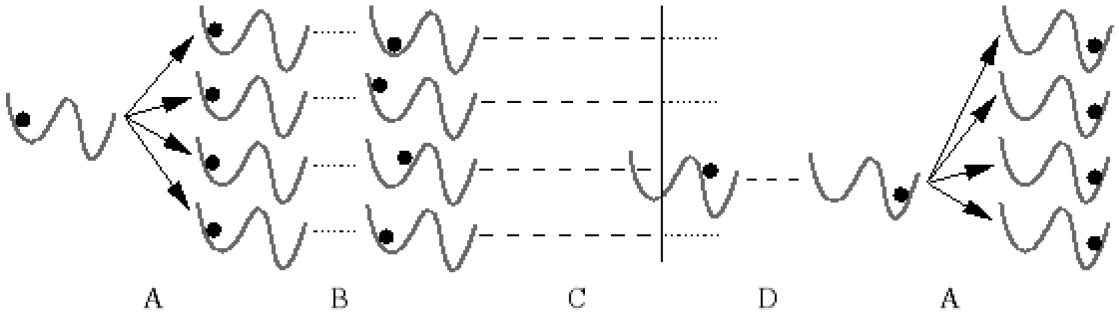
$$= ke^{-kt_{sum}}dt_{sum} \tag{2b}$$

$$= p(t_{sum})dt_{sum}\,, \tag{2c}$$

implying that the probability to leave the state per unit MD time is the same whether that time is accumulated on one or M processors. While this sketched derivation applies for processors of equal speed, we emphasise that the same conclusion can be shown to be valid even if the processors run at different speeds and/or with time-varying speeds (Voter, 1998).

Figure 5.1. shows a schematic of the algorithm. Starting with a system in a particular state, it is replicated on each of the M processors. Each replica, after an independent random momentum assignment, is evolved forward for a time $\Delta t_{deph} \geq \tau_{corr}$ to eliminate correlations between replicas, a stage referred to as dephasing. After dephasing, each processor carries out an independent MD trajectory, typically performed at a constant temperature. Together, the trajectories explore phase space within the particular basin M times faster than a single trajectory would. Once a transition is detected on any processor, all processors are stopped. The simulation clock is then advanced by t_{sum}, the accumulated trajectory time summed over all replicas until the first transition occurred.

Figure 5.1. Schematic illustration of the parallel replica dynamics method



Note: The four steps, described in the text, are (A) replication of the system into M copies, (B) dephasing of the replicas, (C) propagation of independent trajectories until a transition is detected in any of the replicas, and (D) brief continuation of the transitioning trajectory to allow for correlated events such as recrossings or follow-on transitions to other states. The resulting configuration is then replicated, beginning the process again.
Source: Motalenti and German, 2002

ParRep also correctly accounts for correlated dynamical events (Chandler, 1978; Voter Doll, 1985), i.e. there is no requirement that the system obeys transition state theory (Marcelin, 1915; Wigner, 1932; Eyring, 1935; Pechukas, 1981), unlike the other AMD methods (or aKMC). This is accomplished by allowing the trajectory that made the transition to continue for a further amount of time $\Delta t_{deph} \geq \tau_{corr}$, during which recrossings or follow-on events may occur. The simulation clock is then advanced by Δt_{corr}, the new state is replicated on all processors, and the whole process is repeated.

The computational efficiency of the method is limited by both the dephasing stage, which does not advance the system clock, and the correlated event stage, during which only one processor accumulates time. (This is illustrated schematically in Figure 5.1., where dashed line trajectories advance the simulation clock but dotted line trajectories do not.) Thus, the overall efficiency will be high when

$$\frac{\tau_{rxn}}{M} \gg \Delta t_{deph} + \Delta t_{corr} , \tag{5}$$

where $\tau_{rxn} = 1/k$.

An extension to ParRep allows the method to be applied to driven systems (Uberuaga, Stuart, Voter, 2007). To result in valid dynamics, the processors must run at the same speed, and the drive rate must be slow enough that at any given time the rates for the different escape pathways in the system depend only on the instantaneous configuration of the system.

While the derivation of the ParRep method does not impose a particular definition of a "state" of the system, the operational definition often used in practice corresponds to a single basin of the potential energy surface, i.e. a state is taken to be the ensemble of points of configuration space that converge to the same fixed point under a local minimisation of the energy of the system (e.g. using a steepest-descent algorithm). An exponential distribution of escape times is then obtained if the typical timescale for a transition out of the state is long compared to the characteristic vibrational period of the system around that fixed point. While this definition has the virtue of being conceptually and computationally simple, it limits the range of possible applications to systems where the basins are deep

enough (relative to k_BT) and well separated from each other and leaves many other, more complex, systems out of reach. This limitation can sometimes be overcome by a more general definition of a state. For example, in the case of pyrolysis of hexadecane, it was shown that states could be defined as the ensemble of all configuration space points that share the same network of covalent bonds (Kum et al., 2004). In that case, these "superstates" contain a large number of simple energy basins of the potential energy surface, each corresponding to a different conformation of the molecular backbone. There, the method exploited the separation of timescale between the rapid changes of dihedral angles of the backbone (intrasuperstate transitions) and the slow covalent bond breaking process (intersuperstate transitions) rather than between the vibrational timescale and that of sampling of the different dihedral angles. This enables one to ignore the "irrelevant" fast transitions that would demand incessant dephasing and decorrelation and concentrate directly on the real kinetic bottlenecks. This approach has been applied to other cases as well, such as surface diffusion kinetics while the surface is in contact with a liquid phase (with its huge number of shallow states) (Perez et al., 2009) and diffusion of heteroepitaxial surface clusters in metallic systems (Uche et al., 2009). This approach was recently mathematically formalised (Le Bris et al., 2012), showing that the ParRep procedure gives correct results for arbitrary state definitions, provided the dephasing and decorrelation times are made long enough.

We will give specific examples below of the application of ParRep in situations relevant to nuclear energy materials, but we mention here that ParRep dynamics has been successfully applied to a number of varied problems, including the diffusion of H_2 in crystalline C_{60} (Uberuanga et al., 2003a), the pyrolysis of hexadecane (Kum et al., 2004), the diffusion of defects in plutonium (Uberuanga et al., 2003b), the transformation of voids into stacking fault tetrahedra in face centred cubic (fcc) metals (Uberuaga et al., 2007a), the stretching of carbon nanotubes (Uberuaga, Stuart, Voter, 2007), grain boundary sliding in Cu (Mishin et al., 2007), friction-force microscopy (Martini et al., 2009; Dong et al., 2011), the diffusion of Li through a polymer matrix (Duan et al., 2005), the fracture process in metals (Warner, Curtin, Qu, 2007), and the folding dynamics of small proteins (Zagrovic, Sorin, Pande, 2001). As parallel-computing environments become even more common, ParRep will become an increasingly important tool for reaching long-time scales in complex infrequent-event systems.

5.3. Hyperdynamics

Another way to accelerate the state-to-state evolution of an infrequent-event system is to construct an auxiliary system in such a way that the escape dynamics of the latter are faster than those of the former, while enforcing that one maps onto the other by a suitable renormalisation of time. Hyperdynamics (Voter, 1997a) realises this objective by building on the concept of importance sampling (Valleau, Whittington, 1977; Berne, Ciccotti, Coker, 1998) and extending it into the time domain. In this approach, the auxiliary system is obtained by adding a non-negative *bias* potential $\Delta V_b(\boldsymbol{r})$ to the potential of the original system $V(\boldsymbol{r})$ so that, effectively, the height of the barriers between different states is reduced, as schematically shown in Figure 5.2. The relationship between the dynamical evolution of the original and biased systems is recovered using transition state theory (TST). According to TST, the rate of escape of the original system out of a given state A is given by

$$k_{A\rightarrow}^{TST} = < |v_A| \delta_A(\boldsymbol{r}) >_A \qquad (6)$$

where $\delta_A(\boldsymbol{r})$ is a Dirac delta function centred on the dividing surface between state A and the neighbouring state(s) (here the hypersurface is at $\boldsymbol{r} = 0$), v_A is the velocity normal to it and $\langle P \rangle_A$ is the canonical ensemble average of a quantity P for a system confined to state A. By standard importance sampling manipulations, the last equation can be recast in a form where the averages are obtained on the *biased* potential instead. We find:

$$k_{A\rightarrow}^{TST} = \frac{< |v_A| \delta_A(\boldsymbol{r}) e^{\beta \Delta V_b(r)} >_{A_b}}{< e^{\beta \Delta V_b(r)} >_{A_b}}, \tag{7}$$

where $\beta = 1/k_B T$ and k_B is the Boltzmann constant. If we impose the condition that the bias potential must vanish at the dividing surface, the last equation can be rewritten as

$$k_{A\rightarrow}^{TST} = \frac{< |v_A| \delta_A(\boldsymbol{r}) >_{A_b}}{< e^{\beta \Delta V_b(r)} >_{A_b}}. \tag{8}$$

This result is very appealing since the relative rates of escapes from A to other states is invariant under the addition of the bias potential, i.e.

$$\frac{k_{A_b \rightarrow B}^{TST}}{k_{A_b \rightarrow C}^{TST}} = \frac{k_{A \rightarrow B}^{TST}}{k_{A \rightarrow C}^{TST}}. \tag{9}$$

Thus, the state-to-state dynamics on the biased potential is equivalent to that on the original potential as long as the time in each state is renormalised to account for the uniform relative increase of all the rates introduced by the biased potential. This rescaling of time is obtained by multiplying the MD timestep Δt_{MD} by the inverse Boltzmann factor for the bias potential, so that n MD timesteps on the biased potential are equivalent to an elapsed time of

$$t_{\text{hyper}} = \sum_{j=1}^{n} \Delta t_{MD} e^{\beta \Delta V_b [\boldsymbol{r}(t_j)]} \tag{10}$$

on the original potential. Even though the speedup is typically different for each state, this definition of the hypertime gives an ongoing estimate of the accelerated time with statistical error bars that are unbiased, and in the long-time limit (e.g. after many transitions) it converges on the exact value with vanishing relative error. The speedup for escape from state A is given by the average boost factor B,

$$B_{\text{hyper}} = \frac{t_{\text{hyper}}}{t_{MD}} = < e^{\beta \Delta V(\boldsymbol{r})} >_{A_b}, \tag{11}$$

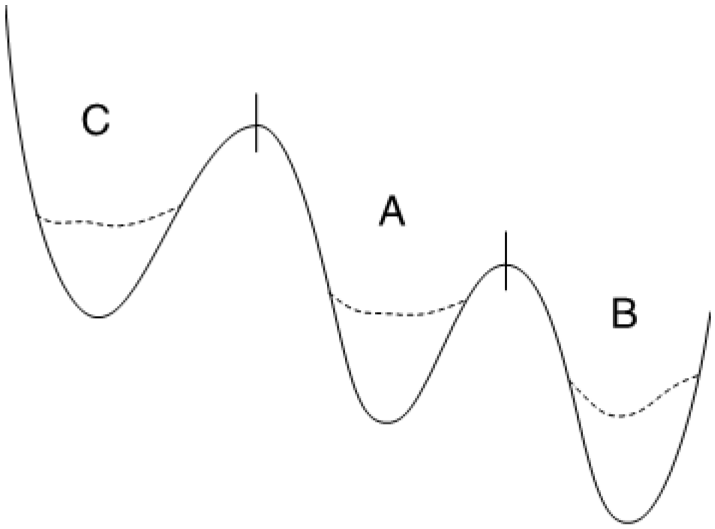
and the overall computational boost is this boost divided by the relative extra cost of calculating the bias potential and associated forces.

If both the original and biased systems obey TST and the bias potential is zero on all the ridgetops, so that the above derivation holds, hyperdynamics can provide considerable, accurate time acceleration compared to direct-MD simulations. It has been successfully applied to a variety of systems, including desorption of organic molecules from graphitic substrates (Beker, Mignogna, Fichthorn, 2009), surface diffusion of metallic clusters (Voter, 1997b), heteroepitaxial growth (Miron, Fichthorn, 2005), microscopic studies of sliding friction (Kim, Falk, 2010, 2011), plastic response of compressed nanopillars (Hara,

Li, 2010) and the dynamics of biomolecules (Hamelberg, Mongag, McCammon, 2004; Xin, Doshi, Hamelberg, 2010).

Figure 5.2. Schematic illustration of the hyperdynamics method

A bias potential (ΔV(r)), is added to the original potential (V(r), solid line). Provided that ΔV(r) meets certain conditions, primarily that it be zero at the dividing surfaces between states, a trajectory on the biased potential surface (V(r)+ΔV(r), dashed line) escapes more rapidly from each state without corrupting the relative escape probabilities. The accelerated time is estimated as the simulation proceeds.



Source: Uberuaga et al., 2019

In practice, the applicability of hyperdynamics is limited by the availability of bias potentials that meet the necessary requirements and have reasonable computational overhead. The construction of bias potentials that are valid, efficient and transferable remains a challenge and a subject of active research. A review of bias potential forms as of 2002, such as the Hessian-based bias (Vote, 1997b) and the flat bias (Steiner, Genilloud, Wilkins, 1998), can be found in Voter, Montalenti and Germann (2002).

An important advance on this front was taken by Miron and Fichthorn, with the introduction of their "bond-boost" bias potential (2003). As the name suggests, the bond-boost bias potential is composed of pairwise terms that tend to soften the bonds between atoms. The key assumption here is that transitions between states will involve the formation or breaking of some bond so that the proximity to a transition state will be signalled by an unusually large distortion of a bond. If the overall bias potential is then designed to vanish when any bond in the system distorts by more than some critical amount (say by more than 20% of its equilibrium length), then it should be possible to safely turn off the bias before a dividing surface is reached. While this approach relies on the assumption that bond lengths alone are reliable indicators of the distance to a potential energy ridge, it provides a very flexible and efficient way to carry out hyperdynamics.

The flexibility of this bond-boost approach can be further exploited to make hyperdynamics more tolerant to low barriers. Indeed, the presence of low barriers is detrimental to the performance of hyperdynamics because it will frequently cause the bias potential to vanish. In cases where appropriate "superstates" can be defined, it is possible to leave the bias potential turned on when the system crosses intrasuperstate barriers without affecting the accuracy of the simulation. This power of this promising type of "bridging" approach,

proposed by Miron and Fichthorn (2004), has been demonstrated for deposition of Co on Cu(100) (Miron, Fitchorn, 2005).

Hamelberg et al (2004) have developed a variation on Steiner's flat bias (Steiner, Genilloud, Wilkins, 1998) that is effective for accelerating the dynamics in biomolecules. The rugged landscape makes these very challenging systems for accelerated dynamics, and this bias is typically applied aggressively enough that the bias potential is not zero at all the key barriers. Hamelberg and co-workers (Xin, Doshi, Hamelberg, 2010) have recently developed a way to extract the correct rate constant for a known pathway by appealing to Kramers rate theory after performing multiple hyperdynamics simulations with bias potentials that are non-zero at the dividing surface.

Hara and Li (2010) developed a bias potential to accelerate the nucleation of dislocations in a system under stress, by constructing a form that is sensitive to the local shear distortion, turning off when it reaches a critical value. Although this mechanism-specific approach runs the risk of slowing the rates for processes that the tailored bias is insensitive to (because the bias may be non-zero at the dividing surfaces for these other processes), this can be a powerful approach. They achieved boost factors greater than 10^{10} in simulations of compressed nanopillars. Kim and Falk (2010, 2011) have developed bias potentials suitable for the study of microscopic sliding friction, and have shown how in a driven system, hyperdynamics can be effectively combined with parallel computation by speculatively simulating future system configurations. Fichthorn and co-workers (2009) have demonstrated that hyperdynamics can be carried out with electronic-structure-based forces, although the much greater expense of electronic structure calculations compared to empirical potentials means that with present-day computers and electronic structure methods, this is barely viable for most systems.

Chen and Horing (2007) have proposed a variation on the hyperdynamics approach, path integral hyperdynamics (PIHD), which releases the requirements on the bias potential. The bias is allowed to be non-zero at the dividing surfaces, and TST-violating correlated events are allowed as well. These effects are properly accounted by appealing to a Langevin path integral formalism to define a weight for each trajectory. As such, the calculation of the rate constant for a given rare event requires multiple simulations, in contrast to the standard hyperdynamics where a single trajectory is representative of the long-time evolution. Recently, this PIHD approach has been shown to provide an effective strategy in a study of diffusion in a periodic potential, including the case of a time-dependent forcing (Ikonen et al., 2011), and in a study of barrier crossing by self-avoiding polymer chains (Shin et al., 2010).

5.4. Temperature accelerated dynamics

One natural way of speeding up the dynamics of a system is to simply raise the temperature. However, while the rates of processes will increase with higher temperatures, the relative probabilities of different events occurring will be different than at the original temperature of interest. Correcting for this reordering is the basic idea behind temperature accelerated dynamics (TAD) (Sørensen, Voter, 2000). In TAD, transitions are sped up by increasing the temperature to some T_{high}, but transitions that should not have occurred at the original temperature T_{low} are filtered out. The TAD method assumes that the system obeys harmonic TST (Vineyard, 1957), in which the transition pathway is associated with a saddle point on the energy landscape, and the rate constant has a simple Arrhenius temperature dependence with a fixed prefactor. As a result, TAD is more approximate than the other AMD methods.

However, for many applications, especially in solids, this additional approximation is acceptable.

In each basin, the system is evolved at T_{high}. When a transition is detected, the saddle point for that transition is found. The trajectory is then reflected back into the basin and continued. This procedure is referred to as "basin constrained molecular dynamics" (BCMD). During the BCMD, a list of escape paths and escape times at T_{high} for each pathway is generated. Assuming that harmonic TST holds, and knowing the saddle point energy for the transition, we can then extrapolate each escape time observed at T_{high} to obtain a corresponding escape time at T_{low}. This extrapolation, which does not require knowing the preexponential factor, can be illustrated graphically on an Arrhenius-style plot ($ln(1/t)$ vs. $1/T$), as shown in Figure 5.3. The time for each event seen at T_{high} extrapolated to T_{low} is

$$t_{low} = t_{high} e^{E_a(\beta_{low} - \beta_{high})}, \tag{12}$$

where $\beta = 1/k_B T$ and E_a is the activation energy.

As the BCMD is continued, a new shorter-time event may be discovered. With the additional assumption that there is a minimum preexponential factor, ν_{min}, which bounds from below all the preexponential factors in the system, we can define a time at which the BCMD trajectory can be stopped. This time has the property that the probability any transition observed later would replace the first transition at T_{low} is less than δ, an uncertainty set by the user. This "stop" time is given by

$$t_{high,stop} \equiv \frac{\ln(1/\delta)}{\nu_{min}} \left(\frac{\nu_{min} t_{low,short}}{\ln(1/\delta)} \right), \tag{13}$$

where $t_{low,short}$ is the shortest transition time at T_{low}. When this stop time is reached, the system clock is advanced by $t_{low,short}$, and the corresponding transition is accepted. The TAD procedure is then started again in the new basin. Thus, in TAD, two parameters govern the accuracy of the simulation: δ and ν_{min}.

The average boost in TAD can be dramatic when barriers are high and T_{high}/T_{low} is large. However, as TAD relies upon harmonic TST for validity, any anharmonicity error at T_{high} will lead to inaccuracy in the predicted dynamics. This anharmonicity error can be controlled by choosing a T_{high} that is not too high.

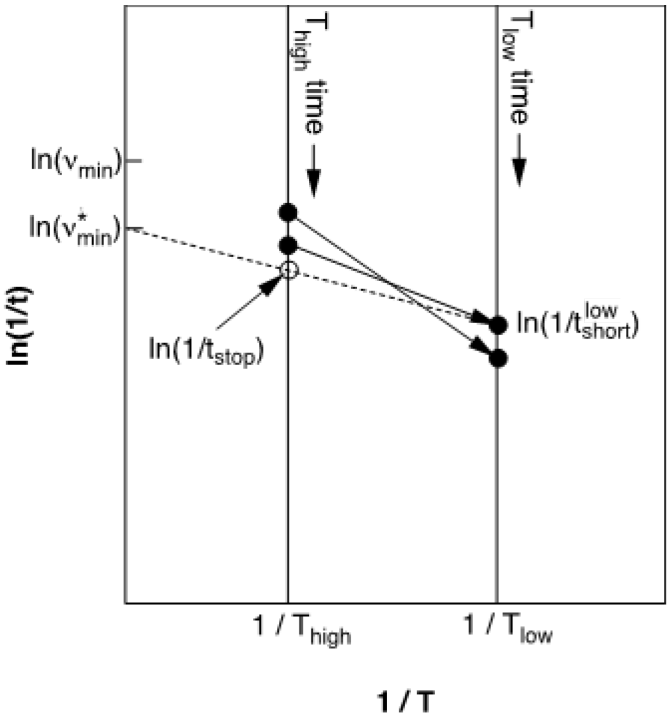
A number of advances have led to increased efficiency in particular systems. "Synthetic" mode (Sørensen, Voter, 2000), a KMC treatment of low-barrier transitions, can significantly improve the efficiency in cases where low-barrier events are repeated often. Furthermore, if we know something about the minimum barrier to leave a given state, either because we have visited the state before and have a lower bound on this minimum barrier or because the minimum barrier is supplied *a priori*, we can accept a transition and leave the state earlier than the time given by equation (13). See Montalenti and Voter (2002) for details.

The AMD methods in general do not scale particularly well with system size, but Shim et al. (2007, 2008) have recently shown how the TAD method can be spatially parallelised, to make it more efficient for larger systems.

Temperature accelerated dynamics has been demonstrated to be very effective for studying the long-time behaviour of defects produced by collision cascades in some ceramics (Uberuaga et al., 2004, 2007b), and more nuclear material applications will be

discussed below. Coupled with MD simulation of surface deposition events, TAD has been applied to simulate the long-time-scale growth of Ag (Mishin, Sørensen, Voter, 2001), Cu (Shim et al., 2008), Cu on Ag(100) (Sprague et al., 2002), and magnetron sputter-deposition growth of Mg-Al-O thin films (Georgieva, Voter, Bogaerts, 2011). Other applications for which TAD has proven effective include defect diffusion on oxide surfaces (Harris et al., 2004), the dissolution of boron-interstitial clusters in silicon (Cogoni et al., 2005), and topological changes in carbon nanotube fragments at high temperature (Uberuaga et al, 2012). Tsalikis et al., (2010a) have recently combined TAD with advanced sampling techniques to study long-time dynamical properties of glass-forming systems.

Figure 5.3. Schematic illustration of the temperature accelerated dynamics method



Note: Progress of the high-temperature trajectory can be thought of as moving down the vertical time line at $1/T_{high}$. For each transition detected during the run, the trajectory is reflected back into the basin, the saddle point is found, and the time of the transition (solid dot on left time line) is transformed (arrow) into a time on the low-temperature time line. Plotted in this Arrhenius-like form, this transformation is a simple extrapolation along a line whose slope is the negative of the barrier height for the event. The dashed termination line connects the shortest-time transition recorded so far on the low temperature time line with the confidence-modified minimum preexponential ($\nu^{\star}_{min}=\nu_{min} / \ln(1/\delta)$) on the y axis. The intersection of this line with the high-T time line gives the time (t_{stop}, open circle) at which the trajectory can be terminated. With confidence 1-δ, we can say that any transition observed after t_{stop} could only extrapolate to a shorter time on the low-T time line if it had a preexponential factor lower than ν_{min}.
Source: Uberuaga et al., 2019

5.5. Adaptive kinetic Monte Carlo

In the aKMC method, the ultimate goal of accelerating the state-to-state dynamics is the same as in the AMD methods, but the conceptual approach is different. aKMC is a generalisation of the KMCo method (Voter, 2006) which we briefly describe first.

In KMC, there is no classical trajectory at all. Instead, from a list of possible pathways and rates for escape from the current state, one escape path is chosen randomly with a probability proportional to its rate constant, and the system is advanced along that path to a new state. The clock is then incremented in an appropriate way, by drawing a time from the exponential distribution (equation (1)) with exponent equal to the sum of the escape rates (Young, Elcock, 1996; Bortz, Kalos, Lebowitz, 1975; Gillespie, 1976). In a typical implementation of KMC, sometimes called lattice KMC, it is assumed each atom can be associated with a fixed lattice position, and the state-to-state transitions all occur along well-understood unit-process pathways. In the bulk, the key unit process is the jump of a vacancy, and on a surface, it is the jump of an adatom to an adjacent site. In principle, somewhat more complicated unit processes (e.g. surface exchange events (Feilbelman, 1990)) can be included, but typically they are not. With the additional assumption that the rate constant for a unit-process pathway is influenced only by the nearby atomic environment, one can compile a list of all possible pathways and their associated rate constants, forming a rate catalogue (Voter, 1986) that is used to drive the lattice KMC simulation. The rates are usually based on harmonic transition state theory (Vineyard, 1957), so that a required step is finding the saddle point connecting the initial and final state for every entry in the rate catalogue.

An attractive formal feature of KMC is that if the rate catalogue is complete (all possible escape paths from each state are included) and the rate constant for each path is exact (which might require going beyond the TST approximation), then the state-to-state dynamical evolution for this system of states is exact (Gillespqie, 1976, 1977). In standard KMC, however, this exact limit is essentially never achieved, and in most cases probably not even well approximated, due to the simplifying on-lattice and unit-process assumptions.

In the aKMC approach (Henkelman, Jónsson, 2001; Middleton, Wales, 2004; El-Mellouhi, Mousseua, Lewis, 2008) the aim is to generate a much more complete rate catalogue. Rather than assuming a set of unit-process escape mechanisms, for each state that is visited during the simulation, one carries out a systematic search for saddle points corresponding to escape paths out of the basin. This can be accomplished efficiently using eigenvector-following approaches such as the dimer method (Henkelman, Jónsson, 1999; Munro, Wales, 1999) or the activation-relaxation technique (Barkeman, Mousseau, 1996; Beland et al., 2011). A key advance, relative to the state of the art of the mid-1990s, is that these methods only require first derivatives of the potential to find the saddle point. When this type of saddle search is carried out, it is not unusual to find a complex pathway involving multiple atoms (e.g. see Henekelman and Jónsson (1999)), and the process may lead to an off-lattice geometry as well. This behaviour is consistent with the complex pathways found during AMD simulations.

This aKMC approach comes much closer to the ideal of exact KMC than the lattice KMC described above. The more searches that are performed, the more complete the aKMC catalogue for the current state is, and the more likely that a kinetically appropriate pathway is selected for the next event. Although it is difficult, if not impossible, to prove that all saddle points surrounding a basin can be found in a finite number of searches, Henkelman has recently discussed how to approximate the confidence that the relevant saddles have

been found after a certain number of randomly initialised saddle searches, by monitoring the rate, per search, at which redundant saddles are found (Xu, Henkelman, 2008).

The computational efficiency of aKMC can be increased in several ways. It can be readily parallelised by performing saddle searches independently on different processors. Also, for large systems, each search can be spatially localised to a subset of the system, although this should be done with care, to avoid missing pathways that are more delocalised than expected. Storing information about each state allows substantial savings when a state is revisited (as is true for all the methods discussed in this chapter), and combining this with the localisation concept and pattern recognition techniques can lead to significant computational gain when locally equivalent, or locally similar, defects are identified (El-Mellouhi, Mousseau, 2008; Xu, Henkelman, 2008; Konwar, Bhute, Chatterjee, 2011).

The aKMC approach has been applied to a variety of processes, including post-cascade defect evolution (Xy, Stoller, Osetsky, 2013), surface diffusion and growth of metal surfaces (Xu, Henkelman, 2008; Konwar, Bhute, Chatterjee, 2011; Henkelman, Jónsson, 2003; Scott et al., 2011), surface erosion by sputtering (Henkelman, Jónsson, 2003), self-diffusion in copper grain boundaries (Pedersen et al., 2009), decomposition of methanol on Cu(100) (Xu, Mei, Henkelman, 2009), elastic effects on the diffusion of vacancies in silicon (El-Mellouhi, Mousseau, 2008), and kinetics in a model glass-forming system (Middleton, Wales, 2004). Using a related approach, an off-lattice self-learning KMC developed specifically for surfaces, Rahman and co-workers have studied surface diffusion of clusters on fcc(111) (Kara et al., 2009). Henkelman's group has discussed the use of electronic structure forces with aKMC (Xu, Henkelman, 2008), and shown that they can be implemented effectively, although this remains very computationally expensive. Jónsson and co-workers (Henkelman, Jónsson, 2003; Pedersen, Jónsson, 2010) have discussed the effective use of aKMC in a distributed computing environment.

5.6. κ-Dynamics

Here we briefly describe a promising approach to long-time dynamics developed by Lu, Makarov and Henkelman (2001), which they term κ-dynamics. Building on the concept that trajectories initiated at an arbitrary TST dividing surface offer a way to recover the exact rate constants for an infrequent-event system (Voter, Doll, 1985), they have designed a method that combines importance sampling with parallelised dividing-surface trajectories to give accelerated state-to-state transitions. They start by specifying a dividing surface that totally encloses the initial state of the system, and which would be infrequently reached by a standard trajectory. Although the efficiency would be higher if this dividing surface was close to the ridgetop surrounding the state, it is not a formal requirement of the method. Hence, simple and tractable definitions can be used to approximate this exceedingly complex ideal surface. Using importance sampling, the TST rate constant for escape through this dividing surface is determined, and a set of representative points on this surface (i.e. Boltzmann distributed within this dividing surface subspace) are generated. These representative points are used as the starting points for trajectories that are run forward and backward in time as far as necessary to determine whether an escape trajectory has been found. An escape trajectory is one that in the outgoing direction proceeds directly to some new state of the system without recrossing the surface, and in the reverse direction settles into the initial state before deeply entering any other states, perhaps after some number of recrossings. Although many trajectories may be required before an escape trajectory is found, this stage of the simulation is parallelisable, as is the importance sampling. In the limit where the importance sampling produces a perfectly unbiased set of starting points on

the dividing surface, κ-dynamics gives correct state-to-state dynamics. Using a bond-boost bias potential form for their importance sampling, Lu et al (2010) applied this approach to surface diffusion events on an Al(100) surface, demonstrating its accuracy, as well as its ability to properly describe important events in which the final state is not connected to the initial state, which requires that correlated dynamical events are treated correctly. Of the AMD and aKMC methods discussed in this chapter, only ParRep and κ-dynamics can treat these correlated events correctly (although in hyperdynamics these correlated events can occur, and may be approximately correct).

5.7. Example applications

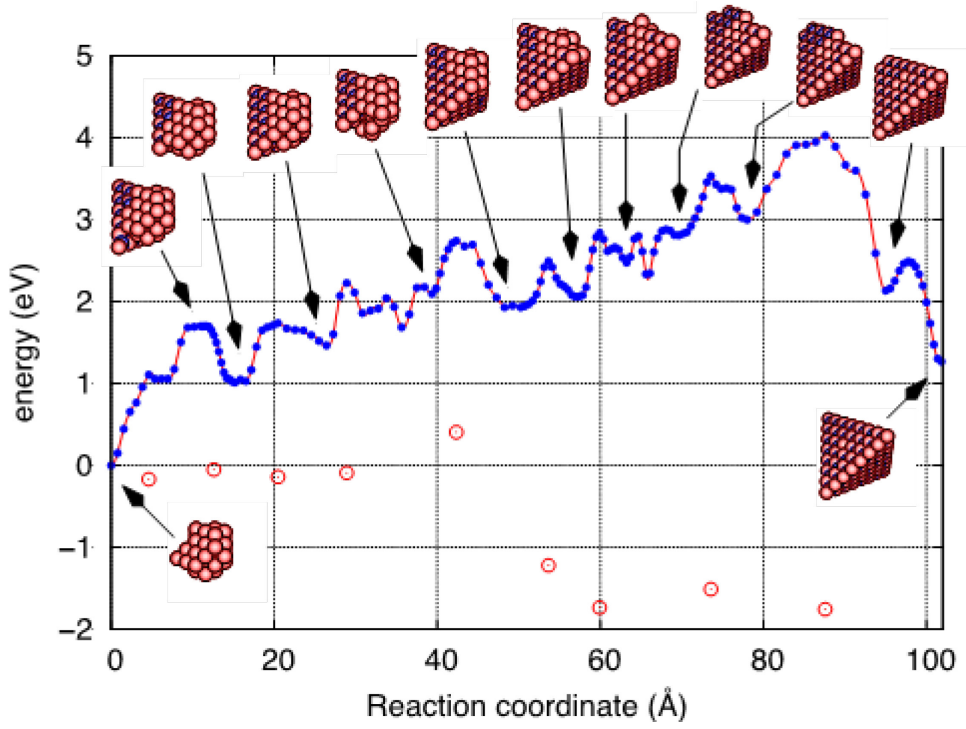
Long-time dynamics methods have been used to examine a number of phenomena in nuclear materials or model materials meant to give basic insight that is applicable to nuclear materials. Here we describe three examples of such simulations and how they provided new insight into the properties and behaviour of these materials. In each of these example cases, an AMD method was employed, but except for the first case (void transformation), in which correlated dynamical events were important, similar results could have been obtained using an aKMC approach.

5.7.1. Void evolution in fcc metals

At high temperatures, the defects produced under irradiation can aggregate, leading to detrimental effects such as swelling. This occurs when a vacancy bias is produced as a consequence of interstitials effectively being removed from the system via the formation of interstitial loops. While it is in fact the loops themselves that cause the swelling, the characteristic signature of this effect is the formation of voids from the aggregation of the excess vacancies. On very long-time scales, these voids can diffuse and aggregate to form even larger clusters. Using AMD methods, primarily parallel replica, we examined the long-time behaviour of voids in both Cu, as a simple model fcc material, and Pu.

These simulations revealed surprising behaviour. It had been assumed in the literature that once voids form, they could not transform to other types of vacancy clusters, such as stacking fault tetrahedra (SFTs), as the barrier for that transformation was considered too high (Zinkel, Seitzman, Wolfer, 1987). However, as described in Uberuaga, et al. (2007a) and shown in Figure 5.4., we found that there is such a high entropic preference for the SFT that the free energy barrier at temperatures of 475 K was very small, meaning that the transformation can occur on the timescale of microseconds at that temperature. As shown in Figure 5.4., from an energy landscape perspective, the transformation is very complicated, with many local minima connected by a variety of barrier heights. However, once the process begins, the entropic driving force is so strong that it happens very quickly on an MD time scale. That is, the system does not thermalise within each local minimum, but rather progresses very quickly along the collection of states connecting the initial void and the final SFT. This illustrates the complexity of the types of processes that can be found via the AMD methods, but also the potential dangers of simulating these types of systems: a method that effectively thermalised the system within each and every state (such as a KMC simulation based upon this landscape) would most likely result in the wrong dynamics. This transformation happens in both Cu and Pu (Uberuaga, Valone, 2008). It is interesting to note that repeating this simulation with even a single He atom placed in the centre of the void significantly slows down the transformation.

Figure 5.4. Transformation pathway of a 45-vacancy void to a stacking fault tetrahedron at *T*=475 K



Note: The figures show the structure of the vacancy cluster as the transformation proceeds. Red spheres are vacancies and blue spheres are interstitials, as determined by comparison to the perfect fcc lattice. The blue points and solid line represent the energy profile for the process. The red points are an estimate of the free energy profile along the path. The large difference between the two is a consequence of the large change in entropy that results from the collapse of the void.
Source: Uberuaga et al. 2007a

5.7.2. Non-equilibrium transport processes

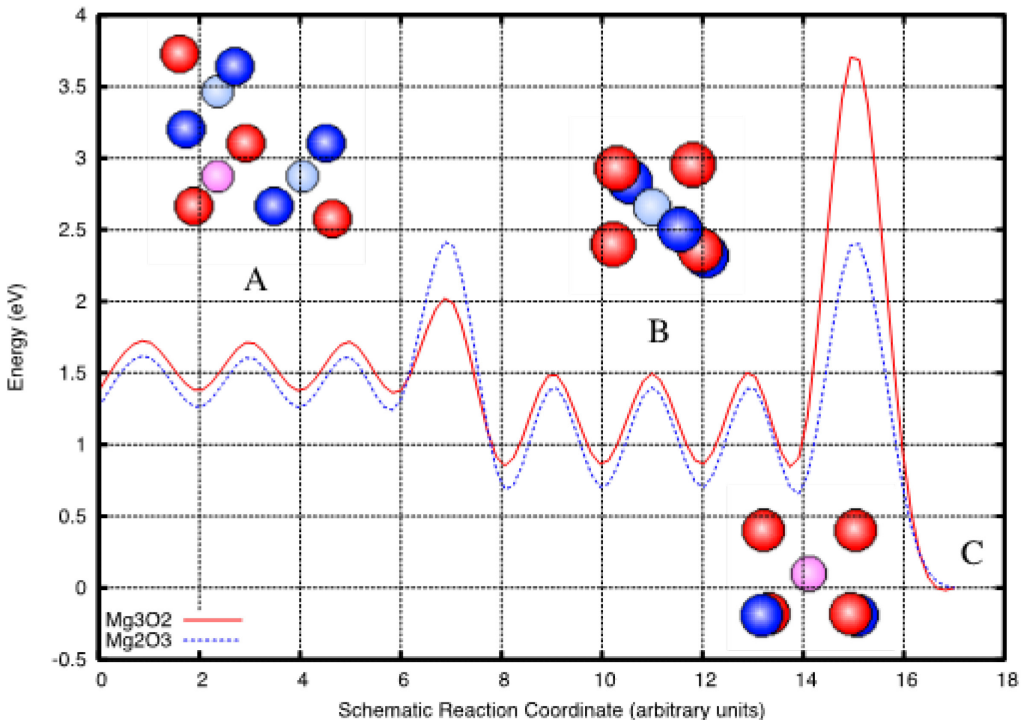
Under irradiation, extremely large amounts of energy are deposited into the material, leading to mass transport mechanisms that would essentially be irrelevant under equilibrium conditions. This effect can take two forms. First, the defect configuration produced directly by the cascade might be metastable, having kinetic properties significantly different than the equilibrium structure. Second, as the defects produced during collision cascades aggregate, they may form structures that are high in energy, again metastable, but that also have unique kinetics.

Two examples illustrate these effects. In the case of the fissioning of U in UO_2, fission products are generated with large amounts of energy. When the fission product is Xe, the equilibrium configuration is a substitution on the U sublattice. However, after a fission event, the final state of the Xe atom depends on the available defects and MD simulations have shown that the final position of the Xe atom can be either interstitial or substitutional (Parfitt, Grimes, 2009). Most studies of Xe diffusion in bulk UO_2 have focused on the substitutional state as it is the most thermodynamically favoured. However, based on these MD simulations, we examined the mobility of interstitial Xe using basin constrained MD, the foundation of TAD (Liu et al., 2011). We found that Xe can diffuse much faster interstitially than it can substitutionally via a complex concerted mechanism on the oxygen sublattice. In this mechanism, Xe pushes out one oxygen, replacing it, and

then is pushed back out into a new interstitial site. This process has a barrier significantly lower than vacancy-assisted diffusion and may be very important for Xe transport in-reactor, where Xe diffusion is known to be much faster than expected based on equilibrium transport measurements.

A second example, illustrated in Figure 5.5., involves the aggregation of defects in MgO, chosen here as a model oxide to understand radiation effects, though MgO also has been proposed as a material in some advanced fuel concepts. Using a combination of MD and TAD, we examined the clustering of defects produced during collision cascades (Uberuaga et al., 2004; 2005b; 2006). We found that large interstitial clusters could adopt a number of structural variants and that higher energy structures, structures that would have essentially zero concentration at equilibrium, can form as a consequence of aggregation reactions and that these high-energy clusters are very mobile. In fact, the fastest diffusing species we have seen in MgO is a metastable structure of the six-interstitial cluster, which diffuses on the ns time scale at room temperature. In contrast, the ground state structure of this same six-interstitial cluster is extremely immobile, with an average hop time measured in years.

Figure 5.5. Schematic of the energy landscape for five-interstitial clusters in MgO



Note: Red indicates oxy- gen, blue magnesium, dark colours interstitials and light colours vacancies. There are at least three structural variants for each of the Mg_2O_3 and Mg_3O_2 clusters. Each variant exhibits unique diffusive characteristics.
Source: Adapted from Uberuaga et al, 2006

The case shown in Figure 5.5. demonstrates the complexity associated with even this simple oxide. Here, the landscape of two versions of a five-interstitial cluster, Mg_2O_3 and Mg_3O_2, are shown. Two striking aspects are illustrated. First, for each of these clusters, there are at least three structural variants with different relative stabilities. Second, each of these variants exhibits its own diffusive behaviour. Given that, during irradiation, these clusters form from even higher energy states (dispersed interstitials), they can easily form

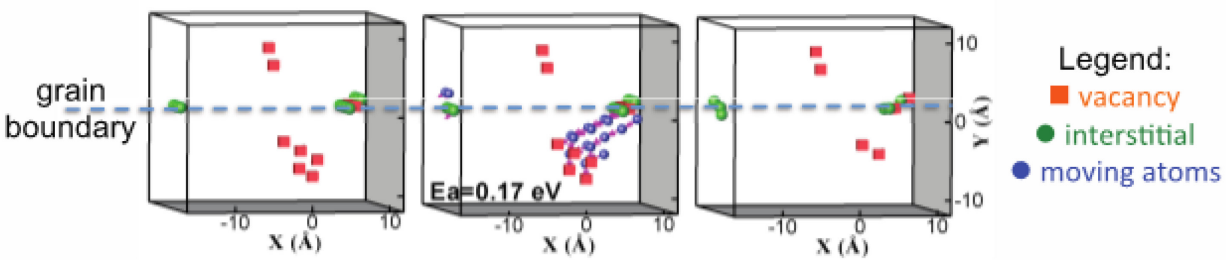
from the defects produced during irradiation. A simple reaction-diffusion model allowed us to determine that these types of clusters would have measurable effects on the nucleation and growth rate of interstitial loops within the material. Thus, it would be critical to include these in mesoscale models of radiation damage evolution in materials such as MgO.

5.7.3. Defect interactions with grain boundaries

The key challenge in advanced nuclear energy concepts – both fission and fusion – is developing materials that can withstand the harsh environments, particularly irradiation. Proposed design concepts require materials that can sustain damage doses of 200 dpa or higher (dpa, or displacements per atom, is a typical measure of damage, and indicates the average number of times an atom is displaced from its lattice site during irradiation). Thus, it is imperative to design materials that can withstand these doses. One promising avenue is to use internal interfaces – phase boundaries and grain boundaries – to promote defect recombination and, effectively, self-healing.

We used MD and TAD to examine the interaction between irradiation-induced defects and grain boundaries in Cu (Bai et al., 2010), again chosen as a model system that can be effectively modelled on the atomistic scale. We simulated collision cascades near the boundaries with MD and the long-time evolution of those defects with TAD. The simulations revealed three important phenomena. First, as seen before (Sugio, Shimomura, de la Rubia, 1998) during the cascades, the boundaries tend to absorb many more interstitials than vacancies, effectively reducing bulk defect recombination. This implies that if no subsequent thermally activated processes were active, boundaries would worsen radiation tolerance by increasing the defect content compared to a single crystal. Second, at moderate temperatures, the interstitials at the boundary could interact with the residual vacancies in the bulk over relatively large distances (~1 nm) with barriers that are very small compared to bulk vacancy diffusion. This so-called "interstitial emission" mechanism, illustrated in Figure 5.6., leads to enhanced interstitial-vacancy recombination compared to what would be expected if the vacancies directly diffused to the boundary. Finally, these results indicate that the sink strength of a boundary is not a fixed quantity, but depends on the absorbed defect concentration and thus the irradiation conditions and the microstructural state of the material (e.g. the density of boundaries). This has clear consequences for mesoscale models of nanomaterials under irradiation.

Figure 5.6. An example of the type of complex processes that can occur near damaged grain boundaries.



Note: In this case, the grain boundary had absorbed interstitials during a collision cascade. Those absorbed interstitials interact much more strongly with nearby vacancies than did the pristine grain boundary. The interstitials effectively "re-emit" from the boundary, annihilating those vacancies, with barriers much smaller than the competing process of conventional vacancy diffusion.
Source: Adapted from Bai et al., 2010

5.8. Limitations and ongoing challenges

While the methods described above have enjoyed considerable success so far, significant challenges remain. The foremost issue is that in many systems, these methods fail to provide significant acceleration due to the presence of large numbers of states connected by low barriers. In essence, the methods described in this chapter achieve time acceleration by collapsing the time to the first event down to a time (or a time-equivalent in computational work) that is on the order of the vibrational time scale. If low barriers cause the time to the first event to be only marginally longer than this vibrational time scale, the boost factor will not be much greater than unity, and can even be much less than unity when the overhead is taken into account (e.g. in TAD). Some strategies have been put forward to mitigate this issue, such as super-state parallel replica dynamics (Perez et al., 2009), state-bridging hyperdynamics (Miron, Fichthorn, 2004), TAD with synthetic mode (Søren, Voter, 2000), TAD with superbasin sampling (Tsalikis et al., 2010a), on-the-fly superbasin detection (Tsalikis et al., 2010b), and incorporation of Tabu sampling (El-Mellouhi, Mousseau, Lewis, 2008). For the aKMC methods, absorbing Markov-chain (Novotny, 1995), and related kinetic system acceleration methods can be employed (e.g. see Chatterjee and Voter (2010) and references therein). However, we believe more work is required before victory can be claimed, as many systems of interest remain out of reach with the current methods. For example, an on-the-fly state definition algorithm that automatically identifies an exploitably large separation of timescales would tremendously extend the reach of parallel replica dynamics. Statistical analysis tools could also be used to identify dynamically “irrelevant” states that could be ignored or lumped with others without affecting the long-time dynamics. Many of these ideas are now being explored and will hopefully lead to more general AMD methods in the next few years.

The second issue is the overall accuracy of the predictions that come out of these long-time methods. We have limited our discussion to methods that give a high-quality description of the long-time dynamics of a system, given particular interatomic potential that provides the forces guiding the dynamics. The burden for providing results accurate enough that these methods can be reliably useful in guiding and interpreting experiment thus falls on the quality of this interatomic potential, which unfortunately is often not up to the task. As discussed above, we are beginning to see implementations of AMD and aKMC methods with electronic structure methods, which can provide much more accurate forces. Further development in this area will be valuable. However, this is still extremely expensive with present-day computers and methods, and a substantial number of systems we wish to study will probably remain out of reach of this kind of direct electronic-structure treatment for the foreseeable future. Thus, it remains important to develop advanced interatomic potential forms, perhaps coupled tightly with electronic structure methodology to gain greater accuracy at reasonable cost. Further, by extending this concept, it should be possible to develop methods that also take specific advantage of the characteristics of infrequent-event systems and the AMD or aKMC method advancing the dynamics. An example of an approach in this general class, designed for regular molecular dynamics, is the “learn on the fly" method of Csanyi et al. (2004), in which spatially local potentials are repeatedly refit, after each set of a few MD steps, to density functional theory results. In the case of infrequent-event systems, it should be possible, for example, to take advantage of the long residence time in a particular energy basin to fit a highly accurate “basin-specific" interatomic potential for each state the system visits.

The final issue we will note is the trend towards increasingly parallel, and increasingly specialised, computer architectures, which offer both new opportunities and new

challenges. Machines at the exascale, anticipated to be less than five years away, will probably have ~10^8 processing cores, each running slower than present-day processors to keep the total power requirements under control. In addition to the challenge of harnessing this many processors effectively without becoming hopelessly communication bound, the nonstandard architectures expected for these systems will make programming a non-trivial task. We have begun to see this added difficulty in programming cell processors (Germann, Kadau, Swaminarayan, 2009) and GPUs (Anderson, Lorenz, Travesset, 2008), for example. On the other hand, the payoff can be substantial. Moreover, the increasing ease with which specialised processors can be designed and manufactured for specific tasks offers the opportunity for co-designed computers and algorithms that achieve an efficiency and speed that is not possible on general purpose machines. The Anton (Shaw et al., 2008) computer, designed specifically to perform ultra-fast MD simulations for biological systems, is an excellent example of this. It is capable of reaching more than a millisecond on realistic biomolecules. One could envision a similar project with the aim of ultra-fast simulation of materials systems, and additional speed might be gained if the characteristics of infrequent-event methods are directly incorporated.

5.9. Advances from 2013 to 2018

Since the time this manuscript was written in 2013, there have been a few noteworthy advances in the methods, which we very briefly describe in this section. For a review of some recent applications, see references: (Zamora et al., 2016a; Perez, Uberuaga, 2015; Ervin, Xu, 2018).

An approximate version of the hyperdynamics method has been developed, designed to maintain constant boost for large systems. In standard hyperdynamics, because making the system larger means there are more escape pathways, the system is more frequently brought near a dividing surface, a condition requiring that the bias potential go to zero, as discussed above. Consequently, for any valid form of bias potential, the boost in hyperdynamics decays towards unity as the system is made very large. In *local hyperdynamics* (Kim, Perez, Voter, 2015), a bias force is defined locally, rather than globally, to circumvent this problem. Although this approach is approximate, it gives strikingly accurate results, and opens the door to massively parallel hyperdynamics simulations on systems with millions or billions of atoms.

A version of the TAD method exploiting speculative parallelisation (SpecTAD) has been developed (Zamora et al., 2016a). The key concept here is that any time an attempted transition to a state *j* is observed during the high-temperature basin constrained MD in state *i*, a TAD simulation on a separate processor can be initiated in state *j*, on the chance that this transition to state *j* will be the one that is ultimately accepted. This spawned process is continued until/unless it becomes clear that state *j* is no longer a candidate for the accepted transition. SpecTAD can give substantially higher boost factors than TAD alone under the right conditions, as can be seen in a study of the radiation-damage-relevant defects in $MgAl_2O_4$ (Zamora et al., 2016b).

A spatially parallelised version of ParRep has been developed (Martinez, Uberuaga, Voter, 2014), using the same sort of sublattice approach mentioned above for spatially parallelised TAD (Shim et al., 2007). For a certain range of system sizes and processor counts (e.g. 10^6 atoms, 10^7 processors, total defect event rate 10^5 s^{-1}), this approach gives a substantially greater parallel speedup than is available from either direct ParRep or spatially parallel MD.

Additionally, a generalisation of parallel replica dynamics with greater parallel efficiency, parallel trajectory splicing (ParSplice), has been developed (Perez, Huang, Voter, 2018). In ParSplice, trajectory segments are generated in many states at the same time – any known state becomes a candidate for execution of one or more segment-generating processes. Building on the quasi-stationary distribution theory discussed in Le Bris et al., (2012), a trajectory segment is "spliceable" if it begins by spending at least a correlation time (or dephasing time) in one state, and similarly ends with a full correlation time in one (same or different) state. Segments are then spliced together end to end – i.e. a segment that ends in a particular state can be spliced to another segment that begins in that same state – to obtain an MD-accurate overall trajectory. Using information from the observed behaviour of the system, one can build a Markovian statistical model for predicting where new trajectory segments will be most likely needed as the system evolves. This model improves the efficiency of ParSplice, but has no impact on the accuracy. For systems whose kinetics are characterised by superbasin trapping, ParSplice can give a substantial speedup compared to ParRep (Perez, Huang, Voter, 2018). ParSplice has been effectively applied to systems with radiation-induced disorder (Perriot et al., 2017) to fusion materials (Perez et al., 2017), and to metallic nanoparticles (Huang et al., 2017).

Finally, AMD techniques have started to evolve beyond the original goal of generating individual state-to-state trajectories towards the efficient generation of higher-scale models of long-time evolution. A recent example of such an approach is TAMMBER (temperature-accelerated Markov models with Bayesian estimation of rates) (Swinburne, Perez, 2018). In TAMMBER, the objective is to use parallel-computing resources to most efficiently create a Markovian (KMC-like) model of the state-to-state dynamics while precisely accounting for the effect of the incompleteness of the model due to the finite amount of (A)MD that was used to parameterise it. The computational engine that underpins TAMMBER is a combination of ParSplice and TAD: short burst of high-temperature MD are carried out in a distribution of states chosen so as to optimally increase the validity time of the model at some lower temperature (the amount of time a trajectory is expected to evolve through known transitions before, statistically, a transition that is not yet part of the model would occur). While the validity time is a global quantity, it can be inferred using estimates of the contribution of the yet-unobserved pathways to total state-wise transition rates; in TAMMBER, this is carried out using a Bayesian formalism that exploits the Poisson statistics of first-escape times. This formalism also allows for the propagation of the uncertainties to observables calculated from the model.

5.10. Conclusion

Since their introduction 20 years ago, the AMD and aKMC methods have proven useful in a variety of situations, including some relevant to nuclear energy materials, where the timescales of interest are out of reach of direct molecular dynamics and where the kinetics are too rich to be adequately described with a limited list of pre-determined pathways. When the activation barriers between the different states are high relative to the thermal energy, any of the AMD methods can yield colossal accelerations, providing a view of atomistic dynamics over unprecedented timescales. Further, by leveraging the particular strength of each of the methods, or by generalising and combining them with other techniques, an even wider variety of situations can be efficiently simulated. Nonetheless, there also remain a large number of systems where persistent low barriers severely limit the available computational boost. In the previous section, we identified this as the greatest limitation of the methods and the most important frontier for future development. Two additional important areas for development are designing methods to provide fast and

accurate forces to guide the dynamics, perhaps by coupling with electronic structure methods and exploiting the properties of the AMD and aKMC methods; and harnessing the full power of specialised computer architectures and the upcoming exascale parallelism.

Acknowledgements

Work at Los Alamos National Laboratory (LANL) was supported by the US Department of Energy (DOE), Office of Basic Energy Sciences, Division of Materials Science, by the LANL Laboratory Directed Research and Development programme, and by the Center for Materials at Irradiation and Mechanical Extremes, an Energy Frontier Research Center funded by the US Department of Energy, Office of Basic Energy Sciences. LANL is operated by Los Alamos National Security, LLC, for the National Nuclear Security Administration of the US DOE under Contract No. DE-AC52-06NA25396.

References

Anderson, J.A., C.D. Lorenz and A. Travesset (2008), "General purpose molecular dynamics simulations fully implemented on graphics processing units", *J. Comp. Phys.*, Vol. 227, 5342.

Bai, X.M. et al. (2010), "Efficient annealing of radiation damage near grain boundaries via interstitial emission", *Science*, Vol. 327, p. 1 631.

Barkema, G.T. and N. Mousseau (1996), "Event-based relaxation of continuous disorderd systems", *Phys. Rev. Lett.*, Vol. 77, 4358.

Becker, K.E., M.H. Mignogna and K.A. Fichthorn (2009), "Accelerated molecular dynamis of temperature-programmed desorption", *Phys. Rev. Lett.*, Vol. 102, 046101.

Beland, L.K. et al. (2011), "Kinetic activation-relxation technique", *Phys. Rev. E*, Vol. 84, 046704.

Berne, B.J., G Ciccotti and D.F. Cocker (eds.) (1998), *Classical and Quantum Dynamics in Condensed Phase Simulations*, World Scientific, Singapore.

Bortz, A.B., M.H. Kalos and J.L. Lebowitz (1975), "A new algorithm for Monte Carlo simulation of Ising spin systems", *J. Comp. Phys.*, Vol. 17, p. 10.

Chandler, D. (1978), "Statistical mechanics of isomerization dynamics in liquids and the transition state approximation", *J. Chem. Phys.*, Vol. 68, 2959.

Chatterjee A. and A.F. Voter (2010), "Accurate acceleration of kinetic Monte-Carlo simulations through the modification of rate constants", *J. Chem. Phys.*, Vol. 132, 194101.

Chen, L.Y. and N.J.M. Horing (2007), "An exact formulation of hyperdynamics simulations", *J. Chem. Phys.*, Vol. 126, 224103.

Cogoni, M. et al. (2005), "Atomistic study of the dissolution of small boron interstitial clusters in *c*-si", *Appl. Phys. Lett.*, Vol. 87, 191912.

Csanyi, G. et al. (2004), "Learn on the Fly: A hybrid classical and quantum-mechanical molecular dynamics simulation", *Phys. Rev. Lett.*, Vol. 93, 175503.

Dong, Y. et al. (2011), "The roles of statics and dynamics in determining transitions between atomic friction remines", *Tribol. Lett.*, Vol. 42, p. 99.

Duan, et al. (2005), "Mechanisms of lithium transport in amorphous polyethylene oxide", *J. Chem. Phys.*, Vol. 122, 054702.

El-Mellouhi, F., N. Mousseau and L.J. Lewis (2008), "Kinetic activation-relaxation technique: An off-lattice self-learning kinetic Monte Carlo algorithm", *Phys. Rev. B*, Vol. 78, 153202.

Ervin, A and H. Xu (2018), "Mesocale simulations of radiation damage effects in material: a SEAKMC perspective", *Computational Materials Science*, Vol. 150, p. 180.

Eyring, H. (1935), "The activated complex in chemical reactions", J. Chem. Phys., Vol. 3, p. 107.

Feibelman, P.J. (1990), "Diffusion path for an Al adamtom on Al(001)", *Phys. Rev. Lett.*, Vol. 65, p. 729.

Fichthorn, K.A. et al. (2009), "Accelerated molecular dynamics simulation of thin-film growth with the bond-boost method", *J. Phys. Condens. Matter*, Vol. 21, 084212.

Georgieva, V., A.F. Voter and A. Bogaerts (2011), "Understanding the surface diffusion processes during magnetron sputter-deposition of complex oxide Mg-Al-O thin films", *Cryst. Growth and Des.*, Vol. 11, 2553.

Germann, T.C., K. Kadau and S. Swaminarayan (2009), "369 Tflop/s molecular dynamics simulations on the petaflop hybrid supercomputer "Roadrunner", *Concurrency and Computation: Practice and Experience*, Vol. 21, 2143.

Gillespie, D.T. (1976), "A general method for numerically simulating the stochastic time evolution of coupled chemical reactions", *J. Comp. Phys.*, Vol. 22, p. 403.

Gillespie, D.T. (1977), "Exact stochastic simulation of coupled chemical reactions", *J. Phys. Chem.*, Vol. 81, 2340.

Hamelberg, D., J. Mongan and J.A. McCammon (2004), "Accelerated molecular dynamics: a promising and efficient simulation method for biomolecules", *J. Chem. Phys.*, Vol. 120, 11919.

Hara, S. and J. Li (2010), "Adaptive strain-boost hyperdynamics simulations of stress-driven atomic processes", *Phys. Rev. B*, Vol. 82, 184114.

Harris, D.J. et al. (2004), "Novel exchange mechanisms in the surface diffusion of oxides", *J. Phys. Cond. Matt.*, Vol. 16, L187.

Henkelman, G. and H. Jónsson (1999), "A dimer method for finding saddle points on high dimensional potential surfaces using only first derivatives", *J. Chem. Phys.*, Vol. 111, 7010.

Henkelman, G. and H. Jónsson (2001), "Long time scale kinetic Monte Carlo simulations without lattice approximation and predefine event table", *J. Chem. Phys.*, Vol. 115, 9657.

Henkelman, G. and H. Jónsson (2003), "Multiple time scale simulations of metal chrystal growth reveal the importance of multiatom surface processes", *Phys. Rev. Lett.*, Vol. 90, 116101.

Huang, R. et al. (2017), "Cluster analysis of accelerated molecular dynamics simulations: A case study of the decahedron of icosahedron transition in Pt nanoparticles", *J. Chem. Phys.*, Vol. 147, 152717.

Ikonen, T. et al. (2011), "Difussion in periodic potentials with path integrals hyperdynamics", *Phys. Rev. E*, Vol. 84, 026703.

Kara, A. et al. (2009), "Off-lattice self-learning kinetic Monte Carlo: application to 2D cluster diffusion on the fcc(111) surface", *J. Phys. Condens. Matter*, Vol. 21, 084213).

Kim, S.Y., D. Perez and A.F. Voter (2013), "Local hyperdynamics", *J. Chem. Phys.*, Vol. 139, 144110.

Kim, W.K. and M.L. Falk (2010), "Accelerated molecular dynamics simulations of low-velocity frictional sliding", *Modelling Simul. Mater. Sci. Eng.*", Vol. 18, 034003.

Kim, W.K. and M.L. Falk (2011), "Role of intermediate states in low-velocity friction between amorphous surfaces", *Phys. Rev. B*, Vol. 84, 165422.

Konwar, D., V.J. Bhute and A. Chatterjee (2011), "An off-lattice, self-learing kinetic Monte Carlo method using local environments", *J. Chem. Phys*., Vol. 135, 174103.

Kum, O. et al. (2004), "Parallel replica dynamics with a heterogenous distribution of barriers: Application to *n*-hexadecane pyrolysis", *J. Chem. Phys.*, Vol. 121, 9808.

Le Bris, C. et al. (2012), "A mathematical formalization of the parallel replica dynamics", *Monte Carlo Methods and Applications*, Vol. 18, p. 119.

Liu, X.-Y. et al. (2001), "Mechanism for transient migration of xenon in UO_2", *Appl. Phys. Lett.*, Vol. 98, 151902.

Lu, C.-Y., D.E. Makarov and G. Henkelman (2010), "Communication: κ-dynamics — An exact method for accelerating rare event classical molecular dynamics", *J. Chem. Phys.*, Vol. 133, 201101.

Marcelin, R. (1915), "Contribution à l'étude de la cinétique physico-chimique", *Ann. Physique*, Vol. 3, p. 120

Martinez, E., B.P. Uberuaga and A.F. Voter (2014), "Sublattice parallel replica dynamics", *Phys. Rev. E*, Vol. 89, 063308.

Martini, A et al. (2009), "Low-speed atomistic simulation of stick-slip friction using parallel replica dynamics", *Tribol. Lett.*, Vol. 36, p. 63.

Middleton, T. and D.J. Wales (2004), "Comparison of kinetic Monte Carlo and molecular dynamics simulations of diffusion in a model glass former", *J. Chem. Phys.*, Vol. 120, 8134.

Miron, R.A. and K.A. Fichthorn (2003), "Accelerated molecular dynamics with the bond-boost method", *J. Chem. Phys.*, Vol. 119, 6210.

Miron, R.A. and K.A. Fichthorn (2004), "Multiple-time scale accelerated molecular dynamics: addressing the small-barrier problem", *Phys. Rev. Lett.*, Vol. 93, 128301.

Miron, R.A. and K.A. Fichthorn (2005), "Heteroepitaxial growth of Co⁄Cu(001): An accelerated molecular dynamics simulation study", *Phys. Rev. B*, Vol. 72, 035415.

Mishin, Y. et al. (2007), "Stick-slip behaviour of grain boundaries studies by accelerated molecular dynamics", *Phys. Rev. B*, Vol. 75, 224101.

Mishin, Y., M.R. Sørensen and A.F. Voter (2001), "Calculation of point defect entropy in metals", *Phil. Mag. A*, Vol. 81, 2591.

Montalenti F. and A.F. Voter (2002), "Exploiting past visits or minimum-barrier knowledge to gain further boost in the temperature-accelerated dynamics method", *J. Chem. Phys*. Vol. 116, 4819.

Munro, L.J. and D.J. Wales (1999), "Defect migration in crystalline silicon", *Phys. Rev. B*, Vol. 59, 3969.

Novotny, M.A. (1995), "Monte Carlo algorithms with absorbing Markov Chains: Fast local algorithms for slow dynamics", *Phys. Rev. Lett.*, Vol. 74, p. 1, Erratum Vol. 75, p. 1 424.

Parfitt, D.C. and R.W. Grimes (2009), "Predicting the probability for fission gas resolution into uranium dioxide", *J. Nucl. Mater.*, Vol. 392, p. 28.

Pechukas, P. (1981), "Transition State Theory", *Ann. Rev. Phys. Chem.*, Vol. 32, p. 159.

Pedersen, A. et al. (2009), "Long time scale simulation of a grain boundary in copper", *New J. Phys.*, Vol. 11, 73034.

Pedersen, A. and H. Jónsson (2010), "Distributed implementation of the adaptive kinetic Monte Carlo method", *Mathematics and Computers in Simulation*, Vol. 80, p. 1 487.

Perez, D. et al (2009), "Accelerated molecular dynamics methods: introduction and recent developments", *Ann. Rep. Comp. Chem.*, Vol. 5, p. 79.

Perez, D. et al. (2016), "Long-time dynamics through parallel trajectory splicing", *Journal of Chemical Theory and Computation*, Vol. 12, p. 18.

Perez, D. et al. (2017), "The mobility of small vacancy/helium complexes in tungsten and its impact on retention in fusion-relevant conditions", *Scientific Reports*, Vol. 7, 2522.

Perez, D. B.P. Uberuaga and A.F. Voter (2015), "The parallel replica dynamics model: Introduction and recent developments", *Comp. Mat. Sci.*, Vol. 100, p. 90.

Perez, D., R. Huang and A.F. Voter (2018), "Long-time molecular dynamics simulations on massively parallel platforms: A comparison of parallel replica dynamics and parallel trajectory splicing", *J. Mater. Res.*, Vol. 33, p. 813.

Perriot, R. et al. (2017), "Evidence for percolation diffusion of cations and reordering in disordered prochlore from accelerated molecular dynamics", *Nature Comm.*, Vol. 8, p. 618.

Scott, C. et al. (2011), "Atomistic surface erosion and thin film growth modelled over realistic time scales", *J. Chem. Phys.*, Vol. 135, 174706.

Shaw, D.E. et al. (2008), "Anton, a special-purpose machine for molecular dynamics simulation", *Communications of the ACM*, Vol. 51, p. 91.

Shim, Y. et al. (2007), "Reaching extended length scales and time scales in atomistic simulations via spatially parallel temperature-accelerate dynamics", *Phys. Rev. B*, Vol. 76, 205439.

Shim, Y. et al. (2008), "Vacancy formation and train in low-temperature Cu/Cu(100) growth", *Phys. Rev. Lett.*, Vol. 101, 116101.

Shin, J. et al. (2010), "Polymer escape from a metastable Kramers potential: Path integral hyperdynamics study", *J. Chem. Phys.*, Vol. 133, 184902.

Sørensen, M.R. and A.F. Voter (2000), "Temperature-accelerated dynamics for simulation of infrequent events," *J. Chem. Phys.*, Vol., 112, 9599.

Sprague, J.A. et al. (2002), "Simulation of growth of Cu on Ag(001) at experimental deposition rates", *Phys. Rev. B.,* Vol. 66, 205415.

Steiner, M.M., P. A. Genilloud and J. W. Wilkins (1998), "Simple bias potential for boosting molecular dynamics with the hyperdynamics scheme", *Phys. Rev. B*, Vol. 57, 10236.

Sugio, K., Y. Shimomura and T.D. de la Rubia (1998), "Computer simulations of displacement damage cascade formation near sigma 5 twist boundary in silver", *J. Phys. Soc. Japan*, Vol. 67, p. 882.

Swinburne, T. and D. Perez (2018), "Self-optimized constructure of transition rate matrices from accelerated atomicstic simulations with Bayesian uncertainty quantification", *Phys. Rev. Mat.*, Vol. 2, 053802.

Tsalikis, D.G., N. Lempesis, G.C. Boulougouris, D.N. Theodorou (2010a), "Temperature accelerated dynamics in glass-forming materials", *J. Phys. Chem. B*, Vol. 114, pp. 7 844–7 853.

Tsalikis, D.G., N. Lempesis, G.C. Boulougouris, D.N. Theodorou 2010b), "Efficient parallel decomposition of dynamical sampling in glass-forming materials based on an "on the fly"definition of metabasins", *J. Chem. Theory Comput.*, Vol. 6, p. 1 307.

Uberuaga, B.P. et al. (2003), "Mechanisms and rates of interstitial H2 diffusion in crystalline C60", *Phys. Rev. Lett.*, Vol. 91, 105901.

Uberuaga, B.P. et al. (2003), "Accelerated molecular dynamics study of vacancies in Pu", *AIP Conference Proceeding*, Vol. 673, p. 213.

Uberuaga, B.P. et al. (2004), "Structure and mobility of defects formed from collision cascades in MgO", *Phys. Rev. Lett.*, Vol. 92, 115505.

Uberuaga, B.P. et al (2005a), "Accelerated molecular dynamics methods", *Handbook of Materials Modeling, Part A - Methods*, Springer, p. 629.

Uberuaga, B.P. et al. (2005b), "Dynamical simulations of radiation damage and defect mobility in MgO", *Phys. Rev. B*, Vol. 71, 104102.

Uberuaga, B.P. et al. (2006), "Accelerated molecular dynamics simulations of interstitial clusters in pure and Al-doped MgO", *Nucl. Instr. and Meth.*, Vol. 250, p. 12.

Uberuaga, B.P. et al. (2007a), "Direct transformation of vacancy voids to stacking fault tetrahedral", *Phys. Rev. Lett*, Vol. 99, 135501.

Uberuaga, B.P et al. (2007b), "Defect kinetics in spinels: Long-time simulations of MgAl2O4, MgGa2O4, and MgIn2O4", *Phys. Rev. B*, Vol. 75, 104116.

Uberuaga, B.P. and S.M. Valone (2008), "Simulations of vacancy cluster behaviour in δ-Pu", *J. Nucl. Mater.*, Vol. 375, p.144.

Uberuaga, B.P. et al. (2012), "Fullerene and graphene formation from carbon nanotube fragments", *Computational and Theoretical Chemistry*.

Uberuaga, B.P. and A.F. Voter (2006), "Accelerated molecular dynamics methods", in *Radiation Effects in Solids*, Springer, NATO Publishing Unit, Dordrecht, pp. 25-43.

Uberuaga, B.P., S.J. Stuart, and A.F. Voter (2007), "Parallel replica dynamics for driven systems: derivation and application strained nanotubes", *J. Chem. Phys.*, Vol. 75, 014301.

Uche, O.U. et al. (2009), "Rapid diffusion of magic-sized islands by combined glide and vacancy mechanism", *Phys. Rev. Lett.*, Vol. 103, 046101.

Valleau, J.P. and S.G. Whittington (1977), "A guide to Monte Carlo for statistical mechanics: 1. Highways", in *Modern Theoretical Chemistry*, Plenum, New York, Vol. 5, p 137.

Vineyard, G.H. (1957), "Frequency factors and isotop effects in solid state rate processes", *J. Phys. Chem. Solids*, Vol. 3, p. 121.

Voter, A.F. (1986), "Classically exact overlayer dynamics: Diffusion of rhodium cluster on Rh(100)", *Phys. Rev. B*, Vol. 34, 6819.

Voter, A.F. (1997a), "A method for accelerating the molecular dynamics simulations of infrequent events", *J. Chem. Phys.*, Vol. 106, 4665.

Voter, A.F. (1997b), "Hyperdynamics: Accelerated molecular dynamics of infrequent events", *Phys. Rev. Lett.*, Vol. 78, 3908.

Voter, A.F. (1998), "Parallel replica method for dynamics of infrequent events", *Phys. Rev. B*, Vol. 57, 13985.

Voter, A.F., F. Montalenti and T.C. Germann (2002), "Extending the time scale in atomistic simulation of materials", *Annu. Rev. Mater. Res.*, Vol. 32, p. 321.

Voter, A.F. (2006), "Introduction to the kinetic Monte Carlo method", in *Radiation Effects in Solids*, Springer, NATO Publishing Unit, Dordrecht, pp. 1-24.

Voter, A.F. and J.D. Doll (1985), "Dynamical corrections to transition state theory for multistate systems: Surface self-diffusion in the rare-event regime", *J. Chem. Phys.*, Vol. 82, p. 80.

Warner, D.H., W.A. Curtin and S. Qu, (2007), "Rate dependence of crack-tip processes predicts twinning trends in f.c.c. metals", *Nature Materials*, Vol. 6, p. 877

Wigner, E. (1932), "Über das Überschreiten von Potentialschwellen bei chemischen Reaktionen", *Z. Phys. Chem.*, Vol. B 19, p. 203.

Xin, Y., U. Doshi, and D. Hamelberg (2010), "Examining the limits of time reweighting and Kramers'rate theory to obtain correct kinetics from accelerated molecular dynamics", *J. Chem. Phys.*, Vol. 132, 224101.

Xu, H., R.E. Stoller and Y.N. Osetsky (2013), "Casade defect evolution processes: Comparison of atomostic methods", *Journal of Nuclear Materials*, Vol. 443, pp. 66–70.

Xu, L. and G. Henkelman, "Adaptive kinetic Monte Carlo for first-principles accelerated dynamics", *J. Chem. Phys.*, Vol. 129, 114104.

Xu, L., D. Mei and G. Henkelman (2009), "Adaptive kinetic Monte Carlo simulation of methanol decomposition on Cu(100)", *J. Chem. Phys.*, Vol. 131, 244520.

Young, W.M. and E.W. Elcock (1966), "Monte Carlo studies of vacancy migration in binary ordered alloys: I", *Proc. Phys. Soc.*, Vol. 89, p. 735.

Zagrovic, B., E.J. Sorin and V.S. Pande (2011), "Deta-hairpin folding simulations in atomistic detail using an implicit solvent model", *J. Mol. Biol.*, Vol. 313, p. 151.

Zamora, R.J. et al. (2016a), "The modern temperature-accelerated dynamics approach", *Annu. Rev. Chem. Biomol. Eng.*, Vol. 7, 3.1.

Zamora, R.J. et al. (2016b), "The effects of cation-anion clustering of defect migration in $MgAl_2O_4$", *Phys. Chem. Chem. Phys*., Vol. 18, 19647.

Zinkle, S.J., L.E. Seitzman and W.G. Wolfer (1987), "I. Energy calculations for pure metals", *Philos. Mag. A*, Vol. 55, p. 111.

6. Kinetic Monte-Carlo methods

María J. Caturla[1] and Roger E. Stoller[2]

[1]Dept. Física Aplicada, Facultad de Ciencias, Fase II, Universidad de Alicante, Spain

[2]Materials Science and Technology Division, Oak Ridge National Laboratory, United States

6.1. Introduction

The different kinetic Monte-Carlo methods that are most widely used in the field of nuclear materials are described in this chapter. These methods allow for calculations of defect evolution in time for time and length scales that can not be reached by atomistic models such as molecular dynamics simulations. Therefore, the results of these calculations take us closer to the experimental time and length scales. However, they require information from other more fundamental models or from experiments. The KMC methods included in this paper are classified as: atomistic kinetic Monte Carlo, object kinetic Monte Carlo and event kinetic Monte Carlo. In the last section improved models such as parallel algorithms are described.

One of the challenges of multi-scale modelling in nuclear reactor materials is to be able to predict microstructure evolution of irradiated materials for times that expand the lifetime of the reactor. The intrinsic difficulty in developing these models lies in the fact that time scales from picoseconds to years must be studied. The initial damage produced by the energetic particles, the so-called cascade damage, lasts only a few tens of picoseconds and extends only a few nanometers. However, it must be accurately described since the initial distribution of damage is crucial for the later damage evolution. Molecular dynamics (MD) simulations with empirical potentials, have been very successful in modelling this initial stage of damage production. However, these models can only reach time scales of nanoseconds and length scales of nanometers. For a description of these models see Chapters 2 to 4 in this issue.

Kinetic Monte Carlo (KMC) models on the other hand, are able to simulate the evolution in time of a set of objects that have associated certain events with known probabilities of occurrence. Unlike in MD simulations, the vibrations of each atom around its equilibrium position is not modelled, and much longer time scales can be achieved. However, KMC requires information of the types of events that can occur as well as the probability for each one of these events. This information is obtained in different ways depending on the specific KMC approach used. We will describe three different KMC methods: atomistic kinetic Monte Carlo (AKMC), object kinetic Monte Carlo (OKMC) and event kinetic Monte Carlo (EKMC). Firstly, a general background of these models will be given.

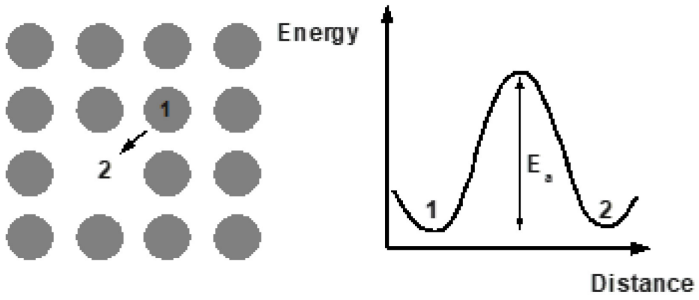
6.2. Background

KMC methods follow the kinetics of slow processes or rare events. Unlike MD simulations where time scales of only a few nanoseconds can be achieved since the minimum time step is on the order of femtoseconds, in KMC methods time scales of seconds or hours can be reached. This allows for direct comparison with experimental measurements, since it is at the right time and also length scale. One of the first KMC simulations were performed by Beeler (1966), followed by those of Besco (1967) and Doran (1970, 1972) all in the field

of radiation effects, since they studied short-term annealing of defects in face centred cubic (fcc) and body centred cubic (bcc) materials. Since then, different algorithms based on the KMC approach have been developed and applied to many different fields.

The KMC method is based on knowing the transition rates between events. And these transition rates are obtained from energy barriers between two different states of the system. For example, the migration of an atom from one lattice position, 1, to a neighbouring vacant site, 2, will be driven by an activation energy as shown schematically in Figure 6.1. This implies that all the possible barriers for the different transitions of the system to be studied must be known. Generally, transition state theory (TST) is used to obtain the rates between two different states (Marcelin, 1915). For a detailed and in-depth description of TST and KMC see the review by A. Voter (2007).

Figure 6.1. Schematic diagram of the activation barrier for an atom in position 1 to move to a vacant site (position 2)



Source: Caturla et al., 2019

Once the activation energy is known, the rate for the event, Γ_e, is then obtained considering:

$$\Gamma_e = \Gamma_{e0} \exp\left(-E_a / KT\right) \tag{1}$$

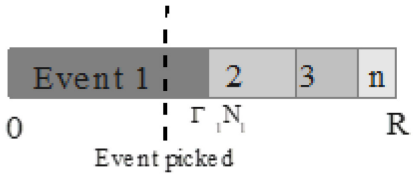
where Γ_{e0} is the jump or attempt frequency, E_a is the activation energy for that particular event, (in the example above, a migration energy), K is Boltzmann's constant and T is the temperature.

Once all the different type of events and event rates are known, the KMC evolves the system in time. The residence-time algorithm or Bortz-Kalos-Liebowitz (BKL) algorithm (Kalos, 1986) is used both in OKMC and in atomistic or lattice KMC. The algorithm proceeds by selecting one event from all possible, with the appropriate weight for each event. First, the total rate for all events is calculated as:

$$R = \sum \Gamma_e N_e \tag{2}$$

that is the sum over all events of the probability of each event, Γ_e, times the number of objects that can undergo that event, N_e. An event is chosen randomly between 0 and R, therefore ensuring that each event is weighted by the appropriate probability of occurrence. Figure 6.2. shows a diagram of the total rate and the event picked for a particular simulation step.

Figure 6.2. Total rate as the sum of individual rates of event up to n events



Source: Caturla et al., 2019

The time of the simulation is then increased by Δt:

$$\Delta t = \frac{-\log \xi}{R} \tag{3}$$

where R is the total rate given by equation (2) and ξ is a random number between 0 and 1, that is used to give a Poisson distribution of the time. Once the event has been selected, a random particle is chosen from all those that can undergo that event. The particle is then moved and the total rate has to be computed again for the next simulation step.

A different KMC algorithm for advancing the time is used in the so-called EKMC and first passage kinetic Monte Carlo (FPKMC) methods, as explained later in this chapter. However, both methods use, as in the case of AKMC and OKMC, a set of tabulated rates for all possible events in the simulation.

6.3. Atomistic kinetic Monte Carlo

In AKMC atoms are fixed in a rigid lattice (Young, Elcock, 1966). Events are usually the migration of an atom from a lattice position to a neighbouring position, considering the appropriate activation energies for migration. For further details on this method see references (Bellon, 2003; Soisson 2006, 2007, 2010).

This method is used very successfully to study processes such as precipitation or segregation in alloys driven by point defects (Soisson, 2006). It is also widely used to study thin film deposition and growth (Gilmer et al., 2000; Huang, Gilmer, Diaz de la Rubia, 1998). AKMC is often also known as lattice kinetic Monte Carlo (LKMC) as in reference (Khrushcheva et al., 2003).

AKMC has a high computational cost, since the jump of every single atom is followed in the simulation. That limits the total time that can be reached as well as the system size that can be modelled. However, it has other advantages with respect to the OKMC that will be described below.

6.4. Object kinetic Monte Carlo

OKMC started to be widely used in the field of radiation damage after the work of H. Heinisch in the early 90s (1990). Unlike in the case of AKMC, not all atoms are described, but only the objects of interest. In the case of radiation, these objects are the defects produced during the irradiation, that is vacancies, self-interstitials, impurities and their clusters. The events these objects can perform are diffusion events, dissolution from a cluster, interaction between different defects or defects with other objects such as grain

boundaries or dislocations. The probabilities of these events are given by the migration energies and binding energies of the defects. For example, the probability of a defect of type i undergoing a migration event is given by:

$$\Gamma_m^i = \Gamma_0^i \exp\left(-E_m^i / KT\right) \tag{4}$$

where, as mentioned above, Γ^i_0 is the jump frequency, E^i_m is the migration energy for that particular defect, K is Boltzmann's constant and T is the temperature. When a migration event is selected the object is moved a distance δ , the jump distance, which is often selected between first and second nearest neighbours. When the object can migrate in any direction (three-dimensional migration) the jump is performed by randomly placing the object within a sphere of radius δ. When the migration of the object is restricted to one particular direction (one-dimensional migration), such as the case of self-interstitial clusters in metals, a particular direction of motion (<111> for Fe and <110> for Cu, for example) with respect to the simulation box is given to the object when it is created, and the jumps are performed only along that direction and with a distance δ.

The probability of a defect of type i undergoing a dissolution event from a cluster is given by:

$$\Gamma_d^i = \Gamma_0^i \exp\left(-\left(E_m^i + E_b^i\right) / KT\right) \tag{5}$$

where E^i_b is the binding energy of the defect to the cluster. This energy depends on the number of defects in the cluster.

The initial conditions of the simulation are the (x,y,z) co-ordinates of those defects produced by the irradiation as well as their type. In case of a continuous irradiation, new defects are introduced in the simulation box with a rate according to the dose rate of the experiment that is being simulated. The positions and types of defects are obtained from molecular dynamics, from binary collision approximation (BCA) calculations, such as those obtained from SRIM (SRIM-ref) or Marlow (MARLOW-ref), or as a random distribution of Frenkel-pairs, depending on the type of calculation. For example, when damage is produced by electrons the last approximation can be used (Fu et al., 2005). In the case of damage produced by light ions such as He, calculations using the binary collision approximation are appropriate. However, for self-irradiation and heavy-ions it is necessary to use those results obtained from MD simulations. Often times a combination of BCA and MD calculations is used to obtain the distribution of defects during irradiation for energies that can not be reached by MD alone. In this case the BCA is used to obtain the energies of those recoils produced by the energetic particle along its path, but the final defect distribution produced by those recoils is the one obtained from MD simulations. This approximation is based on the existence of a threshold for sub-cascade formation.

Most of the OKMC calculations for radiation effects in metals describe the objects as points in the simulation box with a capture radius that depends on the number of defects of that object. This capture radius is normally defined spherical as:

(6)

where n is the number of defects in the cluster and Ω is the atomic volume. This capture radius is used to define when two defects interact. Also when a defect dissolves from a

cluster it is positioned outside this capture radius. When using this approach information regarding the lattice structure is lost. It is however possible to keep the location of every defect during the calculation with the consequent increase in memory for the calculation. This is in fact done in models of dopant diffusion in silicon (Johnson, Carturla, Diaz de la Rubia, 1998). Strain effects such as the bias interaction between interstitials and dislocations can be included in this capture radius, increasing the capture radius for interstitials. However, it is also possible to include strain effects in OKMC using elasticity theory (Sun et al., 2000; Hudson et al., 2005). Note that in the AKMC method described above there is no need to define a capture radius since all atoms, although in a rigid lattice, are included in the calculation.

As mentioned above, the system evolves in this KMC following the residence-time algorithm. Examples of applications of this method can be found in (Domain, Becquart, Malerba, 2004; Becquart et al., 2010; or Caturla et al., 2000 among many others).

6.5. Event and first-passage kinetic Monte Carlo

Another type of algorithm to evolve the time in a kinetic Monte Carlo model is the one used by the so-called EKMC and the FPKMC methods. Like in AKMC or OKMC these methods require a list of event rates. However, the evolution in time of those events is performed differently. In these KMC models the time delay for all the events in the simulation is calculated first. And the event with the shortest-time delay is picked first. That event is then performed, and all the new time delays have to be calculated again for the next simulation step. The time advances by the delay time of the event picked. This algorithm is used in codes such as JERK (Lanore, 1974; DallaTorre et al., 2005) and has been used very successfully to compute the electrical resistivity of irradiated Fe (Fu 2005) and Fe in the presence of impurities such as carbon (Jourdan, Crocombette, 2012). In the EKMC method some approximations are made in order to calculate the delay time for events such as the interaction between to neighbouring objects. For more details on EKMC see references (DallaTorre et al., 2005; Jourdan, Crocombette, 2012; Becquart et al., 2010).

A general method, without any approximations, was recently developed by Opplestrup et al (2006, 2009) named FPKMC. In this method, as in the EKMC algorithm, the event selected is the one that would occur in the shortest time from all possible events. However, FPKMC lacks the approximations included in EKMC. In FPKMC, each one of the walkers is surrounded by a "protective domain" and when a walker is selected it is moved to the edge of that domain. That results in a tremendous computational gain with respect to OKMC since many small jumps, that would have to be done in OKMC where the jump distance is fixed, are automatically avoided in FPKMC. However, the efficiency of FPKMC decreases significantly with respect to OKMC when the number of particles is high, since for each step, all the times associated to the jump of each walker to the edge of their "protective domains" have to be calculated. Some applications and comparisons between FPKMC and OKMC can be found in Donev et al. (2010).

6.6. Other advanced models

In recent years there has been considerable efforts to improve the KMC. In particular, there is an increased interest and need to produce more accurate values for the energy barriers of the different states, especially in the case of OKMC. One of the great advancements in this area is what is known as on-the-fly KMC. The basic idea of this method is to compute the rates of the specific processes at the same time as the kinetic algorithm is evolving. That is,

the event rates are not tabulated before the KMC calculation starts. This is particularly important for those systems where the type of events possible are very large and it is not possible, *a priori*, to know or define every possible scenario. This is, for example, the case of defect diffusion in alloys, in particular, in concentrated alloys, where the rate of a particular reaction (a vacancy jump) will depend on the local environment.

The methodologies used to implement an on-the-fly KMC algorithm differs between different groups and is adapted to the type of problem that needs to be solved. Probably the first on-the-fly KMC model is the one by Henkelman and Jónsson (2001). The authors use the dimer method (Helkelman, Jónsson, 1999) to obtain the saddle points between different states and construct an event catalogue. Stress-assisted diffusion of hydrogen in iron has been studied (Ramasubramanian et al., 2008) combining on-the-fly calculations of barriers using empirical potentials with pre-calculated barriers with more accurate density functional theory. For other examples of on-the-fly KMC see reference (Blackwell et al., 2013; Robinson et al., 2012; Vernon et al., 2011; El-Mellouhi, Mousseau, 2008).

A new method named self-evolving atomistic kinetic Monte Carlo (SEAKMC) also uses an on-the-fly determination of reaction saddle points and energy barriers in a KMC model (Xu, Osetsky, Stoller, 2011). Longer simulation times can be reached in SEAKMC because it focuses the defect analysis on defined "active volumes" which are regions that encompass the defects of interest. The atomistic fidelity of SEAKMC to the underlying physics has been demonstrated in direct comparisons with MD simulations and the ability of SEAKMC to reach long times has identified new phenomena (Xu, Osetsky, Stoller, 2012; Xu, Osetsky, Stoller, Terentyev, 2013). In addition, when the systems to be studied are very complex, such as in defect evolution in concentrated alloys, special methods have been developed to produce the catalogue of transition rates in a more efficient way. One such method consists of using artificial neural networks to predict the values of the energy barriers (Castin, Malerba, 2010).

The application of KMC simulations in more and more complex systems also requires efficient methods of solving the KMC algorithm. Nowadays, with the existence of supercomputers, parallelisation would seem an obvious way to boost the time in these calculations. Parallelisation is relatively simple in an EKMC or FPKMC algorithm. However, this is not the case in OKMC. In the OKMC algorithm described above the total rate for all events, R, must be computed at each time step. And in order to compute this rate all the events must be known. This would mean that all nodes should know of all events at every time step, therefore making the parallelisation very inefficient. Recently, Martinez et al. (Martinez, 2011) have developed a synchronous parallel algorithm that, unlike previous attempts, solves the same master equation as the serial algorithm. In this case, the total rate of all events on each processor is kept fixed for all processors by including null events. SEAKMC has been parallelised using a hybrid OpenMP and MPI scheme.

6.7. Conclusions

KMC algorithms are useful simulation tools to follow the evolution in time of a set of events of known probability of occurrence. They can reach time and length scales that allow for direct comparison to experiments, and therefore, for validation of the models. In fact, KMC can be used for validation of other, more fundamental results, such as DFT calculations of migration energies and binding energies of defects (see for example Fu et al., 2005).

One of the major disadvantages of KMC is the need to know all possible events for a particular system, as well as all rates associated to those events. This can only be achieved in simple systems. However, the systems of interest for nuclear applications are never simple, they are often multicomponent alloys with almost unavoidable impurities. Fortunately, an increasing number of groups are developing new methodologies to tackle these problems such as on-the-fly KMC algorithms, where the rates are not catalogued but computed on-the-fly by different methods, including predictive models such as those based on artificial neural networks.

Besides the effort to increase the accuracy of the KMC method for a particular problem, there are also significant advances in terms of speeding up the KMC algorithm itself without using approximations. Such is the case of FPKMC, which can be extremely efficient for low radiation doses (low defect concentrations) or parallel OKMC.

In the view of this author, the development of KMC methods for radiation damage applications will come from different fronts. On-the-fly methods are very recent and promising developments, provided they can be implemented in an efficient manner to reach reasonable time and length scales. On the other hand, one important effect that is lacking in KMC calculations for radiation damage applications is an efficient way of computing defect diffusion including elastic interactions. These interactions are extremely important to understand, for example, loop interactions in metals or dislocation-defect interactions, particularly under strain. This requires a combination of KMC and dislocation dynamics calculations. Some examples already exist in the literature including one single dislocation (Wen, Takahashi, Ghoniem, 2009) but further developments are needed.

Besides the development of methods and algorithms, having more groups in the radiation field using KMC models for real applications would also help improve the field. Even though the KMC algorithm is fairly simple to implement, developing an efficient code, capable of modelling many different conditions and many different species, is not straight forward. However, unlike for other simulation methods such as MD, there are few KMC programmes available to researchers, at least in the case of radiation damage. An effort in this respect is being done at the IMDEA Materials Institute in Madrid (Martín-Bragado et al., 2013) for AKMC and OKMC. Further developments along this line for other KMC algorithms such as FPKMC or parallel OKMC would also be desirable. On the other hand, as mentioned above, standard KMC simulations rely on a set of parameters as input data. Researchers could also benefit from a well-structured data base of parameters properly documented. Such a project is also currently taking place at CETA-Ciemat[3]. Finally, one would expect that, in the spirit of the multi-scale modelling methodology often used in simulations of radiation damage, the combination of all these algorithms (AKMC, on-the-fly KMC, FPKMC and OKMC), together with cluster dynamics, could reach the time and length scales needed to accurately predict the effects of radiation.

Acknowledgements

We thank Dr Ignacio Martín-Bragado for reviewing and helpful comments on this chapter.

3. www.ceta-ciemat.es

References

Becquart, C.S. et al. (2010), "Modeling the long-term evolution of the primary damage in ferritic alloys using coarse-grained methods", *J. Nucl. Mater.*, Vol. 406, p. 39.

Beeler, Jr., J.R. (1996), "Displacement Spikes in Cubic Metals. I. α-Iron, Copper, and Tungsten", *Phys. Rev.*, Vol. 150, p. 470.

Bellon, P. (2003), "Kinetic Monte-Carlo simulations in the crystalline alloys: principles and selected application", in *Thermodynamics, Microstructures and Plasticity*, Kluwer Academic, Dordrecht, The Netherlands, Vol. 108.

Besco, D.G. (1967), Computer Simulation of Point Defect Annealing in Metals, *USA-AEC Report GEMP-644*, October 1967.

Blackwell, S. et al. (2013), "Modelling the growth of ZnO thin films by PVD methods and the effects of post-annealing", *J. of Physics-Condensed Matter*, Vol. 25, 135002.

Castin, N. et al. (2009), "Modelling radiation-induced phase changes in binary FeCu and ternary FeCuNi alloys using an artificial intelligence-based atomistic kinetic Monte Carlo approach", *Nucl. Inst. Meth. Phys. Res. B*, Vol. 267, p. 3 002.

Castin N. and L. Malerba (2010), "Calculation of proper energy barriers for atomistic kinetic Monte Carlo simulations on rigid lattice with chemical and strain field long-range effects using artificial neural networks", *J. Chem. Phys.*, Vol. 132, 074507.

Caturla, M.J. et al. (2000), "Comparative study of radiation damage accumulation Cu and Fe", *J. Nucl. Mater.*, Vol. 276, 1.

Dalla Torre, J. et al. (2005), "JERK, an event-based Kinetic Monte Carlo model to predict microstructure evolution of materials under irradiation", *Phil. Mag.*, Vol. 85**,** 549.

Dalla Torre, J. et al. (2006), "Resistivity recovery simulations of electron-irradiated iron: kinetic Monte Carlo versus cluster dynamics", *J. Nucl. Mater.*, Vol. 352, p. 42.

Domain, C., C.S. Becquart and L. Malerba (2004), "Simulation of radiation damage in Fe alloys: an object kinetic Monte Carlo approach", *J. Nucl. Mater.*, Vol. 335, p. 121.

Donev, A. et al. (2010), "A First-Passage Kinetic Monte Carlo algorithm for complex diffusion-reaction systems", *J. of Computational Physics*, Vol. 229, p. 3 214.

Doran, D.G. (1970), "Computer simulation of displacement spike annealing", *Radiat. Eff.*, Vol. 2, p. 249.

Doran, D.G. and R.A. Burnett (1972), "Computer Simulation of the Short-Term Annealing of Displacement Cascade", in Gehlen, P.C., J.K.R. Beeler and R. I. Jafee (Eds.), *Interatomic Potentials and Simulations of Lattice Defects*, Plenum, New York, NY, p. 40.

El-Mellouhi, F., N. Mousseau and L.J. Lewis (2008), "Kinetic activation-relxation technique: An off-lattice self-learning kinetic Monte Carlo algorithm", *Phys. Rev. B*, Vol. 78, 153202

Fu, C-C. et al. (2005), "Multiscale modelling of defect kinetics in irradiated iron", *Nature Materials*, Vol. 4, p. 68

Gilmer, G.H. et al. (2000), "Lattice Monte Carlo models of thin film deposition", *Thin Solid Films*, Vol. 365, pp. 189-200.

Heinisch, H.L. (1990), "Computer simulation of high energy displacement", *Radiat. Eff. Def. Solids*, Vol. 113, p. 53.

Henkelman, G. and H. Jónsson (1999), "A dimer method for finding saddle points on high dimensional potential surfaces using only first derivatives", *J. Chem. Phys.*, Vol. 111, 7010.

Henkelman, G. and H. Jónsson (2001), "Long time scale kinetic Monte Carlo simulations without lattice approximation and predefined event table", *J. of Chem. Phys.*, Vol. 115, 9657.

Huang, H., G.H. Gilmer and T. Diaz de la Rubia (1998), "An atomistic simulator for thin firlm deposition in three dimensions", *J. Applied Physics*, Vol. 84, 3636.

Hudson, T. et al. (2005), "Effects of elastic interactions on post-cascade radiation damage evolution in kinetic Monte Carlo simulations", *Phil. Mag.*, Vol. 85, p. 661.

Johnson, M., M. J. Caturla and T. Diaz de la Rubia (1998), "A kinetic study of the effectice diffusivity of the silicon self-interstitial in the presense of Carbon and Boron", *J. of Applied Physics*, Vol. 84, p. 1963.

Jourdan, T. and J.-P. Crocombette (2012), "Rate theory cluster dynamics simulations including spatial correlations within displacement cascades", *Phys. Rev. B*, Vol. 86, 54113.

Lanore, J.M. (1974) "Simulation de l'evolution des defauts dans un reseau par le methode de monte-carlo", *Radiat. Eff.*, Vol. 22, pp. 153-162.

Kalos, M.H. and P.A. Whitlock (1986), *Monte Carlo Methods, Vol. I, Basics*, John Wiley & Sons, New York.

Khrushcheva, O et al. (2003), "Copper precipitation in iron: a comparison between metropolis Monte Carlo and lattice kinetic Monte Carlo methods", *Nucl. Instrum. and Meth. B*, Vol. 202, p. 68.

Marcelin, R. (1915) "Contribution à l'étude de la cinétique physico-chimique", *Ann. Physique*, Vol. 3, p. 120.

Martín-Bragado, I. et al. (2013), "MMonCa: An object kinetic Monte Carlo simulator for damage irradiation evolution and defect diffusion", *Computer Physics Communications*, under review.

Martínez, E., P.R. Monasterio and J. Marian (2011), "Billion-atom synchronous parallel kinetic Monte Carlo simulations of critical 3D Ising systems", *J. Comp. Phys.*, Vol. 230, p. 1 359.Opplestrup, T. et al. (2006), "First-Passage Monte Carlo Algorithm: Diffusion without All the Hops", *Phys. Rev. Lett.*, Vol. 97, 230602.

Opplestrup, T. et al. (2009), "First-passage kinetic Monte Carlo method", *Phys. Rev. E*, Vol. 80, 066701.

Ramasubramaniam, A. et al. (2008), "Effect of atomic scale plasticity on hydrogen diffusion in iron: Quantum mechanically informed and on-the-fly kinetic Monte Carlo simulations", *J. of Materials Research*, Vol. 23, p. 2 757.

Robinson, M.T. (1989), "Slowing-down time of energetic atoms in solids", *Phys. Rev. B*, Vol. 40, 10717.

Robinson, M. et al. (2012), "Point defect formation and migration in Ga stabilised δ-Pu", *J. of Nucl. Mat.*, Vol. 423, 16.

Soisson, F. (2006), "Kinetic Monte Carlo simulations of radiation induced segregation and precipitation", *J. Nucl. Mater.*, Vol. 349, p. 235.

Soisson F and C.-C. Fu (2007), "Cu-precipitation kinetics in α−Fe from atomistic simulations: Vacancy-trapping effects and Cu-cluster mobility", Phys. Rev. B, Vol. 76, 214102.

Soisson, F. et al. (2010), "Atomistic kinetic Monte Carlo studies of microchemical evolutions driven by diffusion processes under irraditation", *J. of Nucl. Mat.*, Vol. 406, p. 55.

Sun, L.Z. et al. (2000) "3D dislocation dynamics study of plastic instability in irradiated copper", *J. Nucl. Mater.*, Vol. 283-287, 741.

Vernon, L. et al. (2011), "Growth mechanisms for TiO_2 at its rutile (11) surface", *Phys. Rev. B*, Vol. 83, 075412.

Voter A.F (2007), "Introduction to the kinetic Monte Carlo Method", in *Radiation effects in Solids*, Nato Science Series, Vol. 235, pp. 1-23.

Wen, M, A. Takahashi and N. M. Ghoniem (2000), "Kinetics of self-interstitial cluster aggregation near dislocations and their influence on hardening", *J. of Nucl. Mat.*, Vol. 392, p. 386.

Xu, H., Y.N. Osetsky and R.E. Stoller (2011), "Simulating complex atomistic processes: On-the-fly kinetic Monte Carlo scheme with selective active volumes", *Phys. Rev. B*, Vol. 84, 132103.

Xu, H., Y.N. Osetsky and R.E. Stoller (2012), "Self-evolving atomistic kinetics Monte Carlo: fundamentals and applications", *J. Phys.: Condens. Matter*, Vol. 24, 375402.

Xu, H., R.E. Stoller, Y.N. Osetsky and D. Terentyev (2013), "Solving the puzzle of <100> interstitial loop formation in bcc iron", *Physical Review Letters,* Vol. 110/26, 265503.

Young, W.M. and E.W. Elcock (1966), "Monte Carlo studies of vacancy migration in binary ordered allots: I", *Proc. Phys. Soc.*, Vol. 89, 735.

7. Dislocation dynamics

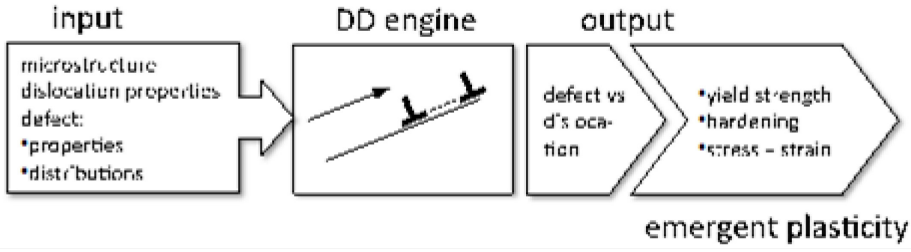
A. I. Topuz[1,3], N. V. Luzginova[2] and E. van der Giessen[3]
[1]Materials Innovation Institute (M2i), Netherlands;
[2]Nuclear Research and consultancy Group (NRG), Netherlands;
[3]University of Groningen, Netherlands

7.1. Introduction

Dislocation dynamics (DD) is a methodology that, in terms of length and time scales, sits in between atomistics and continuum plasticity. It is powerful for problems in which there are too many dislocations to be treated by molecular dynamics (MD), but too few dislocations for continuum plasticity to be applicable. DD applies to both face centred cubic (fcc) and body centred cubic (bcc) materials, by the application of appropriate slip systems and constitutive rules.

As illustrated in Figure 7.1., DD requires input from MD or kinetic Monte Carlo (KMC) etc. in terms of material parameters to describe dislocation evolution and microstructural information, such as grainsize and defect density. Using this, it predicts quantities such as critical resolved shear stress or hardening rate, which may serve as input for a continuum plasticity description, or entire stress-strain curves. Also, it potentially may serve as a routine to describe plastic deformation during other phenomena, such as crack growth or fatigue.

Figure 7.1. Illustration of the position of DD in multi-scale modelling



Source: Topuz et al., 2019

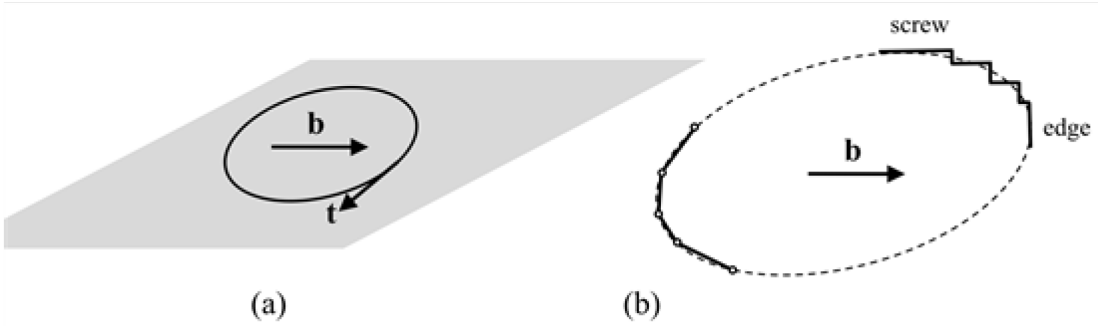
This review will start with a brief review of the underlying assumptions of DD, the theoretical models used in setting up the framework and its numerical most popular implementation. Subsequently, we will describe how the method has been applied in the last two decades to study the influence of irradiation-induced defects on plasticity. For this, we will first consider unit events, with individual dislocations interacting with individual defects, and then present the predictions of plastic deformation as the result of the collective behaviour of multiple dislocations.

7.2. Dislocation dynamics: Methodology

7.2.1. Continuum description of dislocations

While dislocations are line defects in the atomic lattice, at the lengthscale of DD the atoms are averaged out to an elastic continuum and the dislocations are retained as discrete entities. A dislocation loop on a crystallographic plane is characterised by the tangent vector $\mathbf{t}(s)$ as a function of the local co-ordinate s along the loop, and the Burgers vector of the displacement discontinuity across the plane, $\mathbf{b}$, inside the loop, see Figure 7.2.a.

Figure 7.2. Geometry of a single dislocation (a) and top view on a slip plane (b) showing two discretisation schemes of a dislocation: edge-screw (top right) and nodal (bottom left)



Source: Topuz et al., 2019

In DD, the state of stress and strain associated with each dislocation is a continuum elasticity problem. Most DD models adopt linear elasticity, in which case the dislocation line itself becomes a singularity. Closed-form analytical solutions currently exist only for special cases. For straight dislocations (or dislocation segments) the ones in an infinitely large, isotropic elastic medium are best known (see, e.g. Hirth, and Lothe (1973)). Comparison with atomistic simulations has shown that the continuum solutions are accurate at distances of 5~6 times the magnitude of the Burgers vector, b, away from the core of the dislocation. These elastic fields incorporate the long-range interaction between dislocations but generally do not account for finite dimensions of the crystal. The incorporation of traction-free boundary conditions is classically done by means of image dislocations, but since 1995 a versatile method for general boundary-value problems is available that exploits superposition in linear elasticity (Van der Giessen, Needleman, 1995; Weygand et al., 2002).

7.2.2. Constitutive rules

Effects that occur on atomic length scales are not incorporated in the elastic stress fields and need to be formulated in the form of constitutive rules. Such short-range effects include the friction stress; dislocation mobility; annihilation of dislocation segments of opposite character; the interaction of a dislocation with an obstacle or with a grain boundary; the nucleation of a dislocation (or its multiplication in two dimensions (2D)), etc. All of these are essentially governed by the Peack-Koehler force. Line tension can be incorporated either directly through the lengthening of a dislocation or through its local curvature. Cross-slip of screw dislocations is generally treated within the framework of the Friedel-Escaig model. Constitutive rules tend to be relatively simple and involve material parameters whose value is obtained from smaller-scale computations, typically MD simulations, or

from experiments. The difference between implementations for fcc versus bcc crystals is the number of slip systems and the appropriate constitutive rules.

7.2.3. Numerical issues

A DD simulation involves an incremental time-marching algorithm comprising two elementary steps: (1) update the dislocation structure, (2) determine the interactive forces in the updated situation. Changes in the dislocation structure are governed by the Peach-Koehler force along the dislocation line, which is a configurational force. It is computed from the stress field of all dislocations, the image (or applied) stress and line tension. Neglecting the mass attributed to a dislocation line, the dislocation structure is generally evolved in a quasi-static way.

Three classes of analyses can be identified: (a) real three-dimensional (3D); (b) projected 2D; (c) cross-sectional 2D. In (c), dislocations are assumed to be straight, infinitely long edge dislocations (with $\mathbf{b} \perp \mathbf{t}$) or screw dislocations (with $\mathbf{b} \parallel \mathbf{t}$), so that the problem is truly 2D. In both (a) and (b), dislocations are considered to be loops on a plane, but in (b) only a single slip plane is being considered. In either case, dislocations are discretised (see Figure 7.2.b.): either as a string of curved (parametric representation) or straight segments (nodal representation), or in terms of segments that are either edge or screw. In practice, adaptive re-meshing of the loop is used in order to capture large changes in curvature.

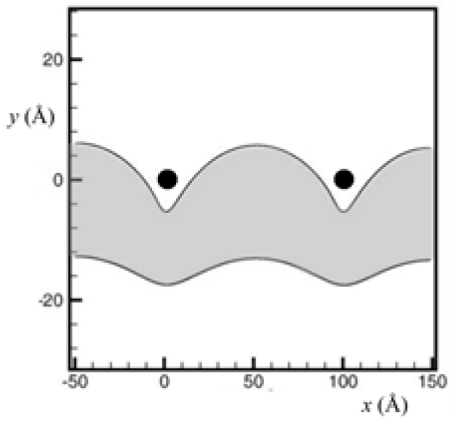
7.2.4. Input for dislocation dynamics for irradiation-induced defects

Theoretical or computational approach begins with *ab-initio* electronic structure calculations which determine basic defect properties such as formation, binding and migration energies of point defects or small defect clusters. In the next stage, MD can describe especially primary damage production. Also, the properties of larger defects like voids and nanoscopic particles are acquired from MD simulations. However, defect and solute diffusion as well as clustering require larger periods (seconds) and distances (micrometrers) with respect to MD. Hence, the information obtained from *ab-initio* and MD is introduced to KMC as input data. KMC plays an important role in informing DD about the 3D spatial distribution, the number and the size of dislocation loops and defects (Wirth et al., 2004). Besides these computational strategies, experimental approaches like transmission electron microscopy (TEM) can be used in order to get spatial distributions, volume fraction, average diameter, and density of defects (Bako et al., 2007).

7.3. Interaction between dislocations and defects

Metals inevitably contain defects as a consequence of production. The defects induced by irradiation as they emerge at the sizescale of individual dislocations, depend on many processing conditions but also on crystallography. In fcc materials, stacking fault tetrahedra (SFT) are a predominant form of damage in low stacking fault energy (SFE) materials like Cu, while in high SFE fcc materials interstitials give rise to Frank sessile loops (Osetsky et al., 2006). Frank loops serve as hardening particles in bcc crystals as well, but in these materials irradiation also gives rise to voids. In this section, we summarise the state of the art in the DD modelling of the interaction between a single dislocation and these defects.

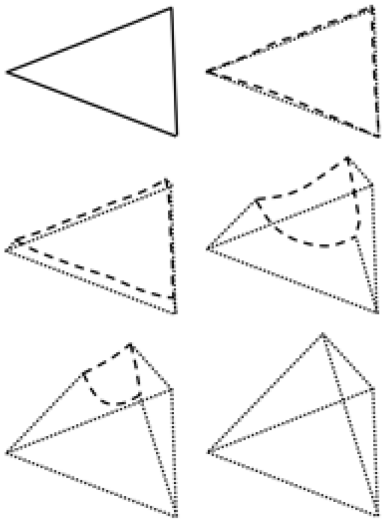
Figure 7.3. Dissociated dislocation with stacking fault (grey area) under an applied shear stress in the vertical direction interacting with small precipipates in Cu, according to projected 2D DD simulations



Source: Adapted from Hardikar et al., 2001

The modelling of SFT's requires a representation of partial dislocations, characterised by the stacking fault energy γ. This energy level scales inversely with the separation distance between adjoining Shockley partials. Hardikar et al. (2001) proposed a simple scheme by representing two partial dislocations as individual dislocations that apart from long-range elastic energy interact with each other via the energy cost corresponding to the stacking fault they bound. They determined the value of γ by fitting the separation distance to the value predicted by molecular statics (through the quasicontinuum method). An example of the profile of the dislocation, showing leading and trailing edges, is shown in Figure 7.3.

Figure 7.4. Pathway of formation of an SFT from a triangular Frank loop (solid line) through dissociation into a Shockley partial (dashed) and a stair-rod dislocation (dotted)

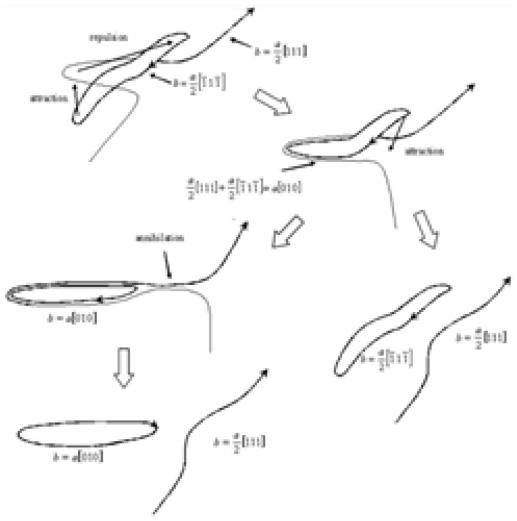


Source: Martinez et al., 2008

A more generic approach was introduced recently by Martinez et al. (2008) in which segments of a dissociated dislocation are characteried explicitly by the quantity $\gamma\mathbf{n}$, in addition to $\mathbf{b}$ and $\mathbf{t}$. Here, $\mathbf{n}$ is the unit normal to the slip plane with opposite directions for the two partials. This allows for a seamless unified treatment of partials and full dislocations (for which $\gamma \to \infty$) as well as Frank loops (a partial edge dislocation with $\mathbf{b}$ not contained in the slip plane). Traditionally, Frank's energy criterion is used to govern the dissociation of dislocations, but within the framework of DD these authors have shown that it is more convenient to adopt a power dissipation criterion. Using their methodology, Martinez et al. (2008) elucidated how the stacking fault contained in the Frank loop is unfaulted in a way that naturally leads to a SFT, as shown Figure 7.4. The method is likely to be amenable to study the interaction between a dislocation and an SFT, as done by means of MD by Wirth et al. (2001), but has not been performed to the best of our knowledge.

Interstitials or vacancies are the results of fast neutron irradiation; as individual atomic defects they are in the realm of MD or KMC, but once they cluster and collapse they can be regarded as sessile prismatic dislocation loops. The stress field of a planar prismatic loop is given in closed-form in (Khraishi et al., 2001) or in more general 3D configurations can be computed in much the same way as for glide dislocations, as demonstrated, for example, by Novokshanov and Roberts (2009). Figure 7.5. shows snapshots of a DD simulation of the passage of a mobile dislocation over an 80nm prismatic loop, which can cross-slip as well as glide. Depending on the height at which the mobile dislocation enters, two interaction mechanisms were obtained leading to different strengths and a different character of the retained prismatic loop. These simulations have also shown that randomly distributed loops are weaker 'obstacles' than loops that are decorating dislocations and thereby locking them.

Figure 7.5. Interaction between a mobile edge dislocation and an inclined interstitial prismatic loop

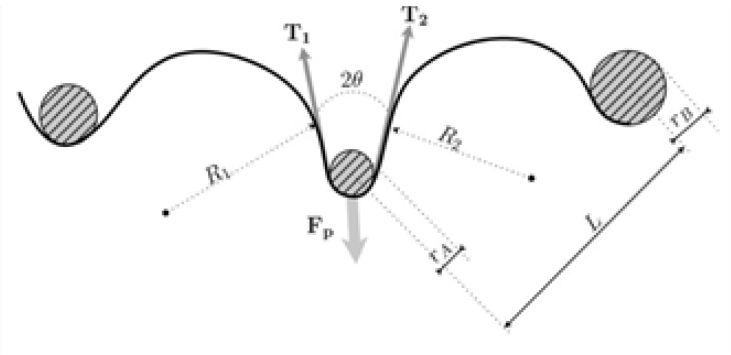


Source: Novokshanov and Roberts, 2009

The interaction between a moving dislocation and a dispersoid or another reinforcing particle can be described by DD provided it is sufficiently large. Various ways to describe the direct interaction between the dislocation and the particle can be found in the literature. A rather sophisticated method is to add an explicit repulsive force field between particle

and dislocation, as proposed by Bako et al. (2007), which will pin the dislocation against the particle. This applies to incoherent and unshearable particles. For such particles, a simpler approach, which is particularly popular in 2D projected simulations, is to represent the particle geometrically as an impenetrable boundary. In this approach, as a dislocation folds around the particle, it is easily seen that the line tension is the most dominant contribution to the force on the particle, see Figure 7.6. Introduction of a maximum value of this force is, in effect, an extra rule which allows to model shearable particles.

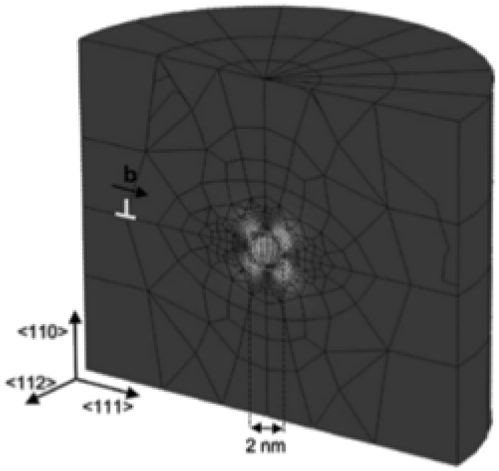
Figure 7.6. Simple geometrical interaction between a dislocation and a particle



Source: Adjanor et al., 2010

Voids are shearable particles in ultimate form. While it can be justified to neglect the mismatch in elastic properties between the crystal and reinforcing particles, the treatment of voids requires proper account of the traction-free conditions on the void surface. Hafez Haghigat et al. (2009) have done so by making use of superposition in conjunction with a finite element method. Confrontation of the stress field around a 2nm void as induced by an edge dislocation at some distance away from it with MD simulations confirmed that this effect can indeed be captured by elasticity, see Figure 7.7. The internal free surface of the void attracts the dislocation, both as it moves towards the void and when it needs to break away from it. It is this interaction that gives rise to the void effectively acting as a particle, with a strength that depends on the lattitude of the slip plane (Hafez Haghighat et al., 2009).

Figure 7.7. Stress field around a void induced by an edge dislocation according to DD



Source: Hafez Haghiaghat et al., 2009

7.4. Collective behaviour: What can dislocation dynamics predict?

The emergent phenomenon caused by the collective behaviour of moving dislocations is plasticity. DD is geared to predicting plasticity on the basis of knowledge regarding the interaction with individual defects and other microstructural features such as precipitates, grain boundaries, twins and laths. There is a large body of literature on DD plasticity, mostly dedicated to stress-strain behaviour, including characteristics like yield strength and hardening rate, but with limited yet increasing attention for the role of plasticity in embrittlement. For space reasons, in the following we concentrate entirely on ways in which irradiation-induced defects have been represented in such studies and some salient findings.

7.4.1. 3D simulations

Fully 3D computations require a large amount of input data, including the position of initial dislocations and defects as well as their size and shape. For this reason (probably), most studies so far have focused on the effect of a single type of defect at a time.

A very detailed study has been performed by Bako et al. (2007) in which the distribution of dispersoids in an Fe alloy were obtained from TEM images. The critical resolved shear stress (CRSS) was computed as the applied stress at which a screw dislocation propagated through the obstacle field (see Figure 7.8. just before the critical stage). The dislocation was discretised using the nodal method, the dispersoids were assumed to be spherical and the interaction with the dislocation treated as mentioned in the previous section. Hundred different realisations of dispersoid distributions were analysed in order to get statistics on the predicted CRSS, which translates into the initial yield strength via the Schmid factor.

Figure 7.8. Snapshot of a screw dislocation moving through a 3D distribution of dispersoids, occasionaly leaving behind small Orowan loops

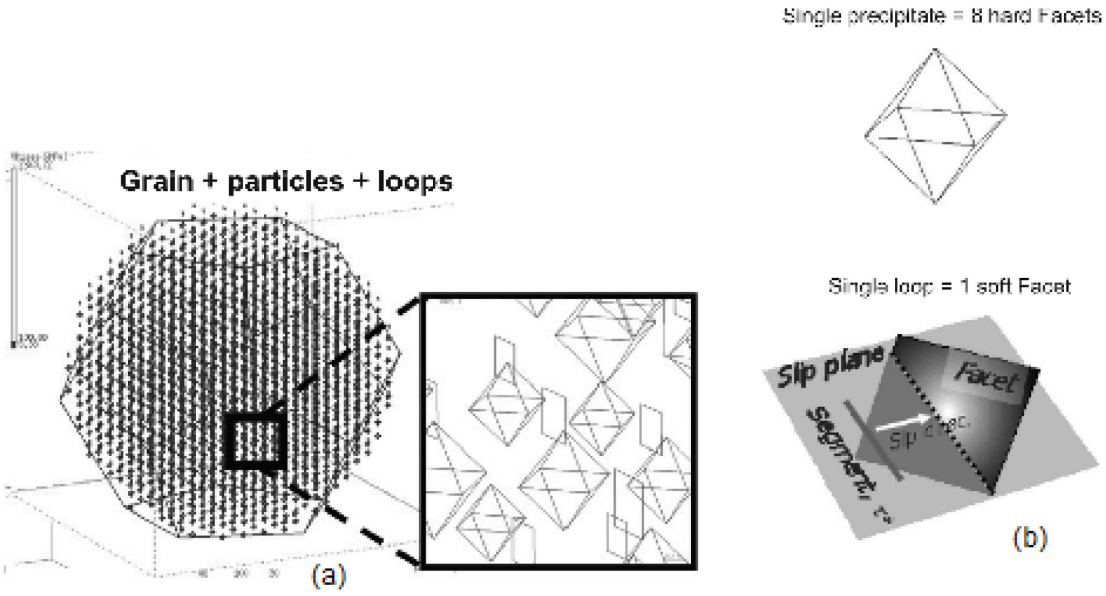
Source: Bako et al., 2007

In a similar spirit, Shin et al. (2009) performed simulations to study the increase of the yield strength of an fcc crystal due to spatial distribution of Frank loops. However, a simplified interaction rule was used: each loop was treated as an obstacle with a barrier strength proportional to the loop radius and independent of the impact location (that is, in contrast with Figure 7.6.). As the final step of a multi-scale analysis, Diaz de la Rubia et al. (2000) reported a similar DD simulation but now including both Frank loops and SFT's. These computations were the first to explain the development of dislocation-free channels. In the

studies mentioned above the various defects were incorporated each by their own representation. Robertson and Gururaj (2011) recently proposed a method in which both hard precipitates and sessile loops were treated by approximating their shapes with facets, see Figure 7.9.

In so far as DD studies go beyond the initial yield stress and predict stress-strain curves it is important to note that in most cases (and certainly the ones cited here) the computations are load controlled. This means that a uniform macroscopic stress is increased incrementally and that strain at each instant is computed. Sometimes only plastic strain is computed from the area swept by the mobile dislocations (e.g. Robertson and Gururaj (2011)), sometimes the elastic part of the strain is added by application of Hooke's law (e.g. Kraishi et al. (2001)). This kind of computation is notably different, though, from a displacement or strain driven tensile test which gives rise to stress drops rather than strain jumps. Simulation of a displacement driven test requires the solution of a boundary-value problem (e.g. Weygand et al. (2001)).

Figure 7.9. (a) Simulation model for the effect of particles and sessile dislocation loops on plastic deformation in ferritic grains, by means of (b) a unified facet representation of hard particles or sessile loops



Source: Robertson and Gururaj, 2011

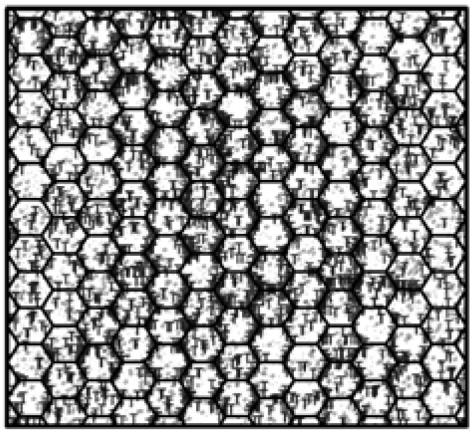
7.4.2. 2D simulations

The 3D computations mentioned above are able to capture many of the details of the interactions between dislocations and (irradiation-induced) defects, but they are limited by the number of dislocations that can be handled. Many more dislocations and larger strains can be achieved with 2D simulations where all dislocations are either pure edge or pure screw, since dislocation lines reduce to points, but, obviously, the detailed interactions are lost. In fact, 2D obstacles are imaginary entities that, somehow, are to represent the (spatial) distribution of real defects and their variation in strength. Likewise, a 2D nucleation site is the projection of a pinned segment that serves as a Frank-Read source. Nevertheless, with appropriately tuned rules, it has been possible, for example, to accurately capture the size

dependence of the stress-strain response of thin polycrystalline films (Shishvan and Van der Giessen (2010)). In this cited paper, the source parameters were fit from experiments but, in principle, it should be possible to extract them from 3D (or projected 3D) computations (in the spirit of the suggestion made by Hardikar et al. (2001)). To the best of the authors' knowledge this latter approach has not yet been applied to study the influence of irradiation-induced defects.

Durgaprasad et al. (2011) have presented an experimentally-based implementation of 2D DD simulations for irradiated materials (see Figure 7.10.). All defects were represented by point obstacles whose ("locking") strength is assumed to obey a Gaussian distribution. The value of the average strength was related by these authors to the radiation fluence via the confrontation of single-crystal simulations with experimental yield strength values. For simplicity, only single slip was considered and shear loading was strain controlled. The grain boundaries in this work are assumed to be impenetrable and because of this interaction give rise to a significant Hall-Petch effect.

Figure 7.10. Shear-induced distribution of edge dislocations in a 2D polycrystalline material containing randomly distributed obstacles with a normal distribution of strengths to represent irradiation-induced defects



Source: Durgaprasad et al., 2011

7.5. Limitations, needs and challenges

Just like any computational model, DD has restrictions; some of these are inherent, and define the range of applicability of the model, others are limitations that are of more practical nature and may be overcome in future implementations. We close the exposition with improvements and extensions that are needed, especially for nuclear materials.

7.5.1. Inherent limitations

- DD only applies to phenomena occurring on length scales $\geq 6b$. Below this, the dislocation core becomes important, and this cannot be captured by linear elasticity.
- The quality of the predictions is controlled by the extent to which physical mechanisms are represtented by constitutive rules and by their accuracy. In particular, the modelling of cross-slip and climb is not as robust as the desciption of dislocation glide.

- The glide velocity of a dislocation needs to be smaller than the shear wave speed, since the stress fields adopted for individual dislocations are for quasi-static dislocations.

7.5.2. Practical limitations (at this moment)

- The fields for individual dislocations and dislocation segments are well known in terms of isotropic elasticity. In truly 2D computations the fields based on cubic isotropy are also known, yet somewhat more cumbersome. Elastically anisotropic fields in 3D, however, are only known formally in integral form, which makes them useless.
- The computing time increases rapidly with number of dislocations because their long-range interactions involves an $O(N^2)$ problem. Fast multipole methods have been developed to remedy this, but the break-even point lies at rather large numbers of dislocation segments.
- The necessary information content about the initial microstructure, including defects and the distribution of their properties, is very large and, in most cases, unknown. It is common to generate random microstructures based on known global measures such as densities, but these microstructures may not be realistic or even attainable in practice. An alternative approach therefore is to perform DD simulations of a preceding deformation process in order to obtain, for example, more realistic dislocation structures that serves as initial microstructure for the desired computation. At any rate, multiple realisations need to be analysed in order to obtain statistically meaningful results.
- The accuracy with which boundary conditions can be incorporated in 3D is quite limited. Superposition methods with a finite element solution are versatile but rely quite sensitively on the mesh size, especially where dislocations intersect boundaries (cf. Weygand et al. (2001)). The treatment of free surfaces recently proposed by Tang et al. (2006) is accurate, but restricted to traction-free boundary conditions.

7.5.3. Needs

- Interfaces are known to have a significant effect on dislocation plasticity. Grain boundaries, for instance, can contribute significantly to hardening (leading to the well-known Hall-Petch grain size effect). Till this date, DD studies of the effect of radiation on yield and hardening have been carried out (almost) exclusively on single crystals, which, at least, raises questions about the relative contribution of irradiation-induced damage. It has been demonstrated in purely 2D studies that grain boundary effects can be incorporated to a first approximation in DD by assuming the boundaries to be impenetrable (e.g. Balint et al. (2008), Shishvan, Van der Giessen (2010), Durgaprasad et al. (2011)) and refined grain boundary models are under construction. This work should be extended to 3D.
- The interaction between dislocations and defects has been modelled at varying levels of sophistication, but have all ignored the elastic mismatch between the metal crystal and, for instance, particles. The influence of this mismatch may be small but this can only be established once a method is available to incorporate it.

- As a further improvement of the previous item, there is a need for methods to study the propagation of a dislocation through a shearable particle, as a function of its strength, size and shape.
- All DD studies referenced in the above have assumed that new dislocations are generated by multiplication through the Frank-Read process. Although relevant, especially for single crystals, this ignores the possibility of nucleation of dislocations from (free) surfaces and interfaces. The need for models of dislocation nucleation is quite generic, and not less so in order to capture the possible enhancement of nucleation from defects such as voids, grain boundaries and other interfaces as well as He bubbles.
- By the very nature of DD, thermal effects cannot be represented directly but have to be captured in constitutive rules. For some processes, like cross-slip, this has been done by means of a probabilistic approach. However, while it is known that thermal activation is key in the evolution of microstructures, to the best of the authors' knowledge this has not been incorporated yet in any DD model. As dislocation crossing of obstacles is a dominant phenomenon in irradiated materials, the incorporation of thermal acitivation seems to be of high priority.
- To this date, DD studies have focused almost exclusively on the effect of irradiation damage on plastic deformation, viz. yield and hardening, while the effects on fracture properties are likely to be of equal importance from a nuclear engineering point of view. Hence, irradiation embrittlement is still a relatively unexplored area. Plastic properties, of course, play an important role in fracture, but there is no simple one-to-one mapping. Especially in many popular nuclear materials which possess rather limited plasticity, fracture is phenomenon that necessitates a multi-scale approach (cf., e.g. Odette et al. (2003)). While the interaction between a dislocation and a stationary crack has been investigated to quite some detail, there have been only a handful of DD-based simulation studies of crack growth. The recent work by Noronha and Ghoniem (2007) is the first attempt to study the effect of irradiation on the ductile-to-brittle transition temperature, but by using a very simplified model. There is a lot of room for improvement here.

References

Adjanor, G. et al. (2010), "Overview of the RPV-2 and INTERN-1 packages: From primary damage to microplasticity", *J. Nucl. Mater.* Vol. 406, pp. 175–186.

Bako, B. et al. (2007), "Discrete dislocation dynamics simulations of dislocation interactions with Y2O3 particles in PM2000 single crystals", *Phil. Mag*, Vol. 81, pp. 3 645–3 656.

Balint, D. et al. (2008), "Discrete dislocation plasticity analysis of the Hall-Petch effect", *Int. J. Plasticity*, Vol. 24, pp. 2 149–2 172.

Diaz de la Rubia, T. et al. (2000), "Multiscale modelling of plastic flow localization in irradiated materials", *Nature*, Vol. 406, pp. 871–874.

Durgaprasad, P.V. et al. (2011), "Application of dislocation dynamics to assess irradiation effect on materials with different grain sizes", *Nucl. Engrg. Design*, Vol. 241, pp. 603–607.

Hafez Haghighat, S.M. et al. (2009), "Dislocation–void interaction in Fe: A comparison between molecular dynamics and dislocation dynamics. *J. Nucl. Mater.,* Vol. 386–388, pp. 102–105.

Hardikar, K., V. B. Shenoy and R. Phillips (2001), "Reconciliation of atomic-level and continuum notions concerning the interaction of dislocations and obstacles", *J. Mech. Phys. Solids*, Vol. 49, pp. 1951–1967.

Hirth, J.P. and J. Lothe (1968) *Theory of Dislocations*, McGraw-Hill, New York.

Khraishi, T.A. et al. (2001), "Modelling of irradiation-induced hardening in metals using dislocation dynamics", *Phil. Mag. Lett.,* Vol. 81, pp. 583–593.

Martinez, E. et al. (2008), "Atomistically informed dislocation dynamics in fcc crystals", *J. Mech. Phys. Solids*, Vol. 56, pp. 869–895.

Noronha, S.J. and N. M. Ghoniem (2007), "Brittle–ductile transition in F82H and effects of irradiation", *J. Nucl. Mater.*, Vol. 367–370, pp. 610–615.

Novokshanov, R. and S. Roberts (2009), "3D Dislocation dynamics modelling of interactions between prismatic loops and mobile dislocations in pure iron", *J. Nucl. Mater.* Vol. 386–388, pp. 64–66.

Odette, G.R. et al. (2003) "Cleavage fracture and irradiation embrittlement of fusion reactor alloys: mechanisms, multiscale models, toughness measurements and implications to structural integrity assessment" *J. Nucl. Mater.*, Vol. 323, pp. 313–340.

Osetsky, Y.N., D. Rodney, and D.J. Bacon (2006), "Atomic-scale study of dislocation–stacking fault tetrahedron interactions. Part I: mechanisms", *Phil. Mag.*, Vol. 86, pp. 2 295–2 313.

Robertson, C. and K. Gururaj (2011), "Plastic deformation of ferritic grains in presence of ODS particles and irradiation-induced defect clusters: A 3D dislocation dynamics simulation study", *J. Nucl. Mater.*, Vol. 415, pp. 167–178.

Shin, C., J. Kwon, and W. Kim (2009), "Prediction of radiation-induced yield stress increment in austenitic stainless steels by using a computational approach", *J. Nucl. Mater.*, Vol. 386–388, pp. 610–612.

Shishvan, S.S. and E. Van der Giessen (2010), "Distribution of dislocation source length and the size dependent yield strength in freestanding thin films", *J. Mech. Phys. Solids*, Vol. 58 pp. 678–695.

Tang, M. et al. (2006), "A hybrid method for computing forces on curved dislocations intersecting free surfaces in three-dimensional dislocation dynamics", *Model. Simul. Mat. Sci. Engrg.*, Vol. 14, pp. 1 139–1 151.

Van der Giessen, E. and A. Needleman (1995), "Discrete Dislocation Plasticity: A Simple Planar Model", *Model. Simul. Mat. Sci. Engrg.*, Vol. 3, pp. 689–735.

Weygand, D. et al. (2002), "A. Aspects of boundary-value problem solutions with three-dimensional dislocation dynamics", *Model. Simul. Mat. Sci. Engrg.*, Vol. 10, pp. 437–468.

Wirth, B.D. et al. (2001), "Mechanical property degredation in irradiated materials: A multiscale modelling approach", *Nuclear Instruments and Methods in Physics Research B*, Vol. 180, pp. 23–31.

Wirth, B.D. et al. (2004), "Multiscale modeling of radiation damage in Fe-based alloys in the fusion environment", *J. Nucl. Mater.*, Vol. 329-333, pp. 103–111.

Zinkle, S.J. (2005), "Fusion Material Science: Overview of challenges and recent progress", *Physics of Plasmas*, Vol. 12, 058101.

8. Introduction to the phase-field modelling technique: A primer on the Allen-Cahn and Cahn-Hilliard models

M. J. Welland
Computational Techniques, Canadian Nuclear Laboratory, Canada

8.1. Introduction

Phase-field models have emerged as a robust modelling approach to understand and predict microstructure evolution with complex interface morphology including topology changes (Allen, Cahn, 1972; Cahn, Hilliard, 1958; Provatas, Elder, 2010; Steinbach, 2013; Bellon, 2012; Chen, 2002, Boettinger et al., 2002). Models of this type can quantitatively simulate phenomena that involve phase changes and/or interfacial energy such as fracture (Borden et al., 2012; Bourdin et al., 2008; O'Connor et al., 2016, Chakraborty et al., 2014), grain boundary movement (Kennouche et al., 2016; Zhang et al., 2016; Tonks et al., 2014 ; Ahmed et al., 2017), bubble evolution (Chakraborty et al., 2014; Millet et al., 2012; Li et al, 2010), melting/solidification (Welland et al., 2009; Böhler et al., 2014), and nano-scale phenomena (Welland et al., 2015; Ulvestad et al., 2015), some of which are shown in Figure 8.1. Models may be derived in a thermodynamically self-consistent fashion, integrated with thermodynamic databases, and used to demonstrate kinetic phenomena such as dendritic solidification and solute trapping. They are readily coupled with other continuum-scale transport equations such as heat and mass transport, elasto-mechanics, surface effects, etc.

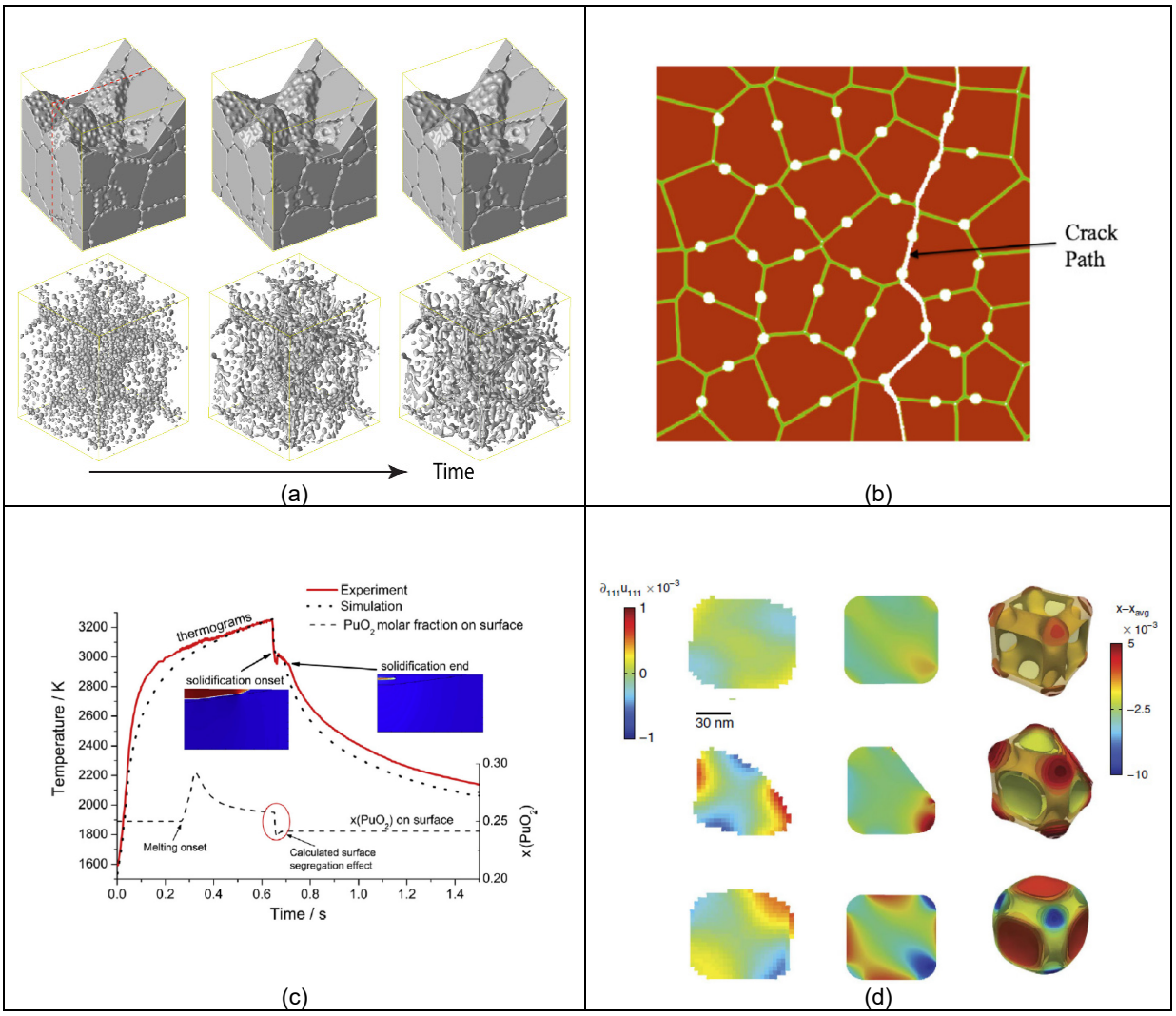
This document is an introduction to the phase-field modelling concept and discussion of the canonical Allen-Cahn and Cahn-Hilliard models. The theory refers to both equilibrium thermodynamics and transport processes and a basic familiarity of both is assumed. Several comprehensive reviews have been written on the current state of phase-field models (Provatas, Elder, 2010; Steinbach, 2013; Bellon, 2012; Chen, 2002, Boettinger et al., 2002) to which the reader is referred for a more in-depth discussion. Similarly, there are several review articles on phase-field modelling of irradiation damage (Devanthan et al., 2010; Li et al., 2017; Tonks et al., 2017).

Phase-field models represent the local phase at a point in space by a field variable which varies smoothly and continuously in space between phases, representing a diffuse interface. Contrasted to sharp interface models, this approach captures complex interface morphology in a robust fashion while naturally including interfacial energy which dominates many nano- and microscopic processes (Welland, Lewis, Thompson, 2008). However, this robustness has substantial computational expense, requiring small mesh sizes in typical numerical solver techniques, as the interface, which is typically very thin, must be resolvable in the system.

Phase-field models are typically derived at the mesoscale which is characterised as the intermediate scale between the description of material as discrete components and a continuum. This scale is large compared to atomic distances and includes enough discrete components to be treated statistically such that thermodynamic properties such as temperature, concentration, pressure, and *phase* are robustly defined (Steinbach, 2013). However, the scale is small compared to other characteristic, macroscale properties such as the size of the system of interest, and the gradients of such thermodynamic properties.

Figure 8.1. Example applications of phase-field modelling

(a): Evolution of intergranular fission gas bubbles in U-Mo for calculation of the effective thermal conductivity. (b): Intergranular fracture with the presence of bubbles along the grain boundary. (c): Non-congruent melting of $(U,Pu)O_2$ fuel in laser flash experiments. (d): Hydrogen accomodation in PdH nanocubes compared to experimental results.



Source: (a) Hu et al., 2015, (b) Chakraborty, Tonks, Pastore, 2014, (c) Böhler et al., 2014, (d) Ulvestad et al., 2015

The governing equations for the phase-field models, along with any other coupled physics, is derivable through the Theory of Irreversible Processes, also known as Non-equilibrium Thermodynamics. Through the governing equations, the system is allowed to evolve so as to generate entropy based on the energy of the system as a whole and is represented by a system of coupled partial differential equations (PDE). This derivation provides a robust mathematical link between equilibrium thermodynamics and the time-dependant evolution of the system with multiple coupled phenomena such as heat and mass transport in the presence of non-congruent phase change (Welland et al., 2009; de Groot, Mazur, 1984; Welland, 2012).

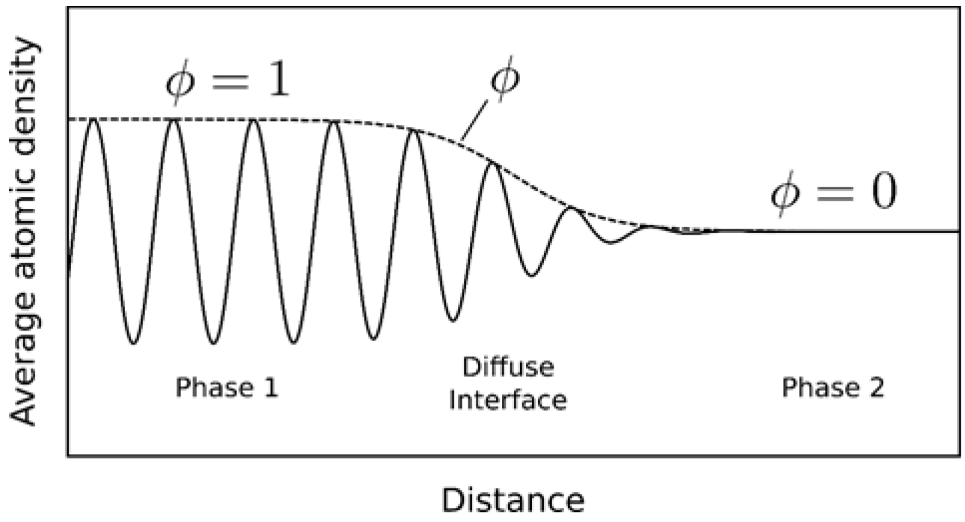
Both the Allen-Cahn and Cahn-Hilliard models have non-uniform, diffuse interface solutions controlled by two factors which together determine the interfacial energy and width. The first factor deals with the local state of the material within the diffuse interface which is not in either of the equilibrium phases and therefore has a non-optimal energy. The second factor is related to the relative state of the surrounding material and is mathematically captured by the inclusion of a *gradient energy* term in the free energy functional of the system, which is sometimes referred to as the Ginzburg-Landau functional (Cahn, Hilliard, 1958; Provatas, Elder, 2010). Minimisation of the excess and gradient energies leads to shrinking and widening of the diffuse interface respectively resulting in a minimal, spatially-dependant profile with a characteristic interface width, and a non-zero interfacial energy.

8.2. The Allen-Cahn model

The Allen-Cahn equation, also referred to as the non-conserved phase-field model or Model A, is useful for representing a first-order phase transformation such as melting/solidification and evaporation/condensation (Allan, Cahn, 1972; Provatas, Elder, 2010; Hohenberg, Halperin, 1977). These transformations feature a discontinuous transition in the first derivative of the free energy, e.g. a latent heat of transformation.

A set of state variables are introduced akin to reaction co-ordinates for the phase transition, and are not conserved quantities. Historically, these variables were envisioned as representing the ordering / symmetry of the phase. This ordering is schematically depicted in Figure 8.2. in which the average atomic density is periodic to the left and becomes smooth to the right corresponding to the transition of, e.g. a solid to a liquid. In modern applications it is common to treat the phase-field variables as abstract 'phase indicators' to spatially represent the local phase and disregard potential physical meaning (Steinbach, 2013).

Figure 8.2. Schematic representation of a one-dimensional (1D) average atomic density field



Note: The figure illustrates a crystal-like ordering on the left and fluid-like disordering on the right with a diffuse interface between them. The phase-field variable ϕ can be envisioned as the amplitude of the oscillation.
Source: Welland, 2019

A fundamental example of the Allen-Cahn model is a two-phase material, which requires only a single phase-field variable ϕ. The free energy of the system is represented as the integral over the system volume, V, of the local free energy density and additional terms present in the interface,

$$F = \int \underbrace{f(\phi, \ldots)}_{free\ energy\ density} + \underbrace{W\phi^2[1-\phi]^2 + \frac{\epsilon^2}{2}|\nabla\phi|^2}_{interfacial\ terms}\ dV\,, \tag{1}$$

where $f(\phi, \ldots)$ is the phase-dependent free energy density, and the ellipses indicate other state variables such as concentration, temperature, etc. The first interfacial term is a *double-well potential* which has minima at $\phi = 0{,}1$. The form of this term is not unique and other models are in use such as the double-obstacle function (Steinbach, 2013). The interfacial terms are controlled by parameters W and ϵ whose product and ratio are related to the interfacial energy and width respectively as described below.

An interpolation function, $p(\phi)$, which interpolates between single-phase properties, can be envisioned as the local volume fraction of a phase. This function also ensures each phase is stable with respect to infinitesimal variations in ϕ. A simple and commonly used form for $p(\phi)$ is,

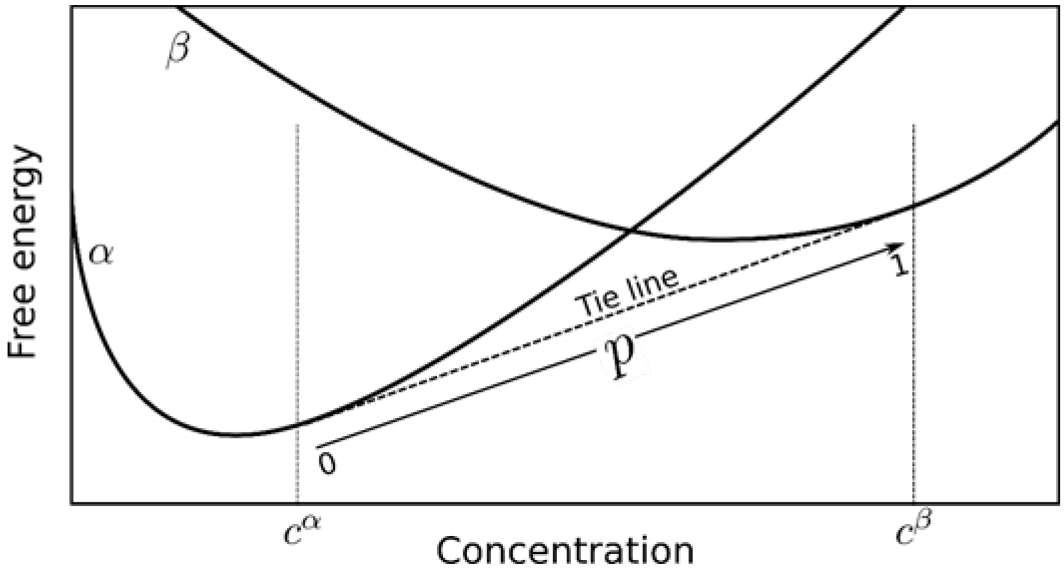
$$p(\phi) = \phi^3[6\phi^2 - 15\phi + 10], \tag{2}$$

which satisfies $p(\phi = 0) = 0$, $p(\phi = 1) = 1$, and $\frac{dp}{d\phi}(\phi = 0) = \frac{dp}{d\phi}(\phi = 1) = 0$. The free energy density of the multiphase system can then be interpolated between those of the two pure phases, f^α and f^β as,

$$f(\phi, \ldots) = [1-p]f^\alpha(\ldots) + pf^\beta(\ldots). \tag{3}$$

The system may be visualised as in Figure 8.3. where the free energy curves for the α and β phases are shown as functions of the concentration, c. The optimal, equilibrium state is the lowest common tangent of the curves which imply the partitioning of species between phases to c^α and c^β. The phase function p is then the distance along the lowest common tangent (tie line).

Figure 8.3. Free energy curves for coexisting phases showing the tie line between equilibrium concentrations



Source: Welland, 2019

The change in the energy of the system with respect to a change in the phase-field profile is given by the functional derivative,

$$\frac{\delta F}{\delta \phi} = \frac{\partial f}{\partial \phi} + W\frac{\partial(\phi^2[1-\phi]^2)}{\partial \phi} - \epsilon^2\nabla^2\phi. \tag{4}$$

In equilibrium, the variation of F with ϕ must be zero, $\frac{\delta F}{\delta \phi} = 0$. Furthermore, for a planar system in equilibrium with two coexisting phases, the driving force for local phase transformation is also zero, $\frac{\partial f}{\partial \phi} = 0$. The shape of ϕ is then controlled by the interfacial terms,

$$W\frac{\partial(\phi^2[1-\phi]^2)}{\partial \phi} - \epsilon^2\nabla^2\phi = 0, \tag{5}$$

which has a solution with the normal co-ordinate x,

$$\phi(x) = \frac{1}{2} + \frac{1}{2}\tanh\frac{x}{2d}, \tag{6}$$

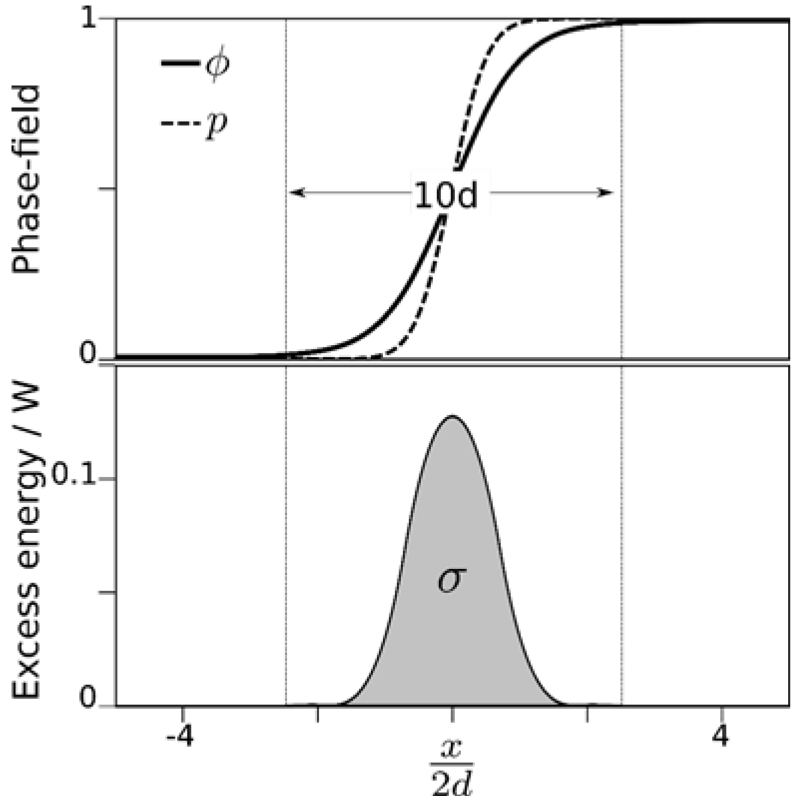
Where

$$d = \sqrt{\frac{\epsilon^2}{2W}}, \tag{7}$$

and is referred to as the interfacial width. Equation (6) is plotted in Figure 8.4. with the size of a factor of d indicated.

Figure 8.4. Demonstrative phase-field profiles and excess energy

Above: equilibrium solution with the interface width shown. Below: Corresponding excess energy which integrates to the interfacial energy.



Source: Welland, 2019

Inserting the equilibrium profile in equation (6) into the energy functional in equation (1), we can determine the total free energy of the system including the interface. The difference between this energy and that of a system without the interfacial terms is identified as the interfacial energy, σ,

$$\sigma = \frac{1}{3}\sqrt{\frac{\epsilon^2 W}{2}}, \tag{8}$$

and is shown as the shaded area in Figure 8.4.

The phase-field parameters in equation (1) are related to the material properties of interface width and energy by inverting equations (7) and (8),

$$\epsilon^2 = 6\sigma d, \tag{9}$$

$$W = 3\sigma/d. \tag{10}$$

Finally, there is also a phase-field parameter related to the kinetics of phase change. Since the phase-field variable is non-conserved, it is driven directly from the reduction of free energy. The governing equation for this parameter is simply,

$$\frac{\partial \phi}{\partial t} = -M_\phi \frac{\delta F}{\delta \phi}, \tag{11}$$

where M_ϕ is a mobility term which can be related to interface attachment kinetics (Karma, 1996), and the variational derivative is given in equation (4).

The Allen-Cahn formula can be generalised to multiphase systems through the introduction of additional phase-field variables, ϕ_1, ϕ_2 ... and the modification of the interfacial terms in equation (1) to include multi-well potentials and other gradient energy terms (Nestler, Garcke, Stinner, 2005; Zhang, Steinbach, 2012; Welland, Wolf, Guyer, 2014; Choudhury, Nestler, 2012). Other thermodynamic variables may be added to this model, such as multicomponent diffusion, displacement variables to represented elasticity and plasticity, and electromagnetic properties (Zhang, Steinbach, 2012; Welland, Wolf, Guyer, 2014; Choudhury, Nestler, 2012). In the case of coupling between phase transitions and another state variable such as concentration, it is required that equation 11) be coupled to the relevant transport equation; i.e. for a binary system undergoing solidification / melting, the Allen-Cahn equation should be coupled to a mass balance equation.

8.3. The Cahn-Hilliard model

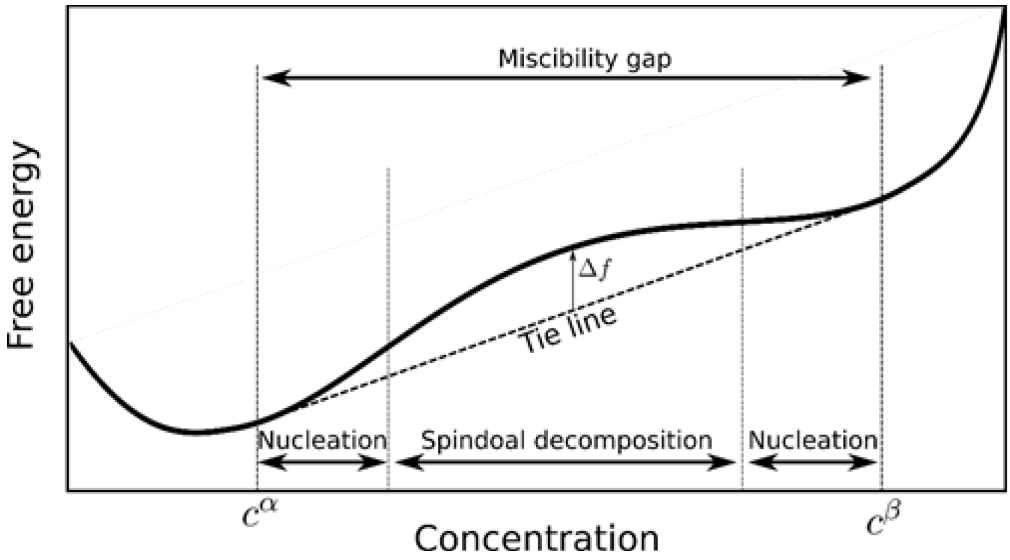
The Cahn-Hilliard model does not introduce an additional order parameter, but instead relies on a free energy potential which is concave in a state variable, commonly concentration, along with a gradient energy term for the same variable (Cahn, Hilliard, 1958). The state variable is usually conserved, and hence this model is also sometimes referred to as the conserved phase-field model, or Model B.

The free energy functional of a canonical Cahn-Hilliard equation is,

$$F = \int f(c, \dots) + \kappa |\nabla c|^2 \, dV \,, \tag{12}$$

where κ is the gradient energy coefficient which acts on the concentration field, in contrast to the phase-field variable as in the Allen-Cahn equation. The free energy density is represented by $f(c, \dots)$, an example of which is shown in Figure 8.5. which corresponds to a binary, two-phase system with a miscibility gap in the phase diagram between the stable bulk phases at concentrations c^α and c^β. The equilibrium solution varies continuously between c^α and c^β as indicated by the tie line, but in this case the material between the equilibrium phases still lies on the free energy curve. Therefore, there is an excess energy enclosed between the tie line and the free energy curve which serves the same function as the double-well potential in the Allen-Cahn equation.

Figure 8.5. Demonstrative free energy profile for a Cahn-Hilliard model.



Note: The tie line shows the equilibrium miscibility gap between concentrations c^{α} and c^{β}. The difference between the tie-line and the free energy curve is shown as Δf and penalises material between phases. Also labelled are the metastable region in which classical nucleation needs to occur to trigger the phase change, and the unstable region wherein infinitesimal fluctuation of composition can trigger spinodal decomposition.
Source: Welland, 2019

Since the excess energy described above depends on the free energy profile and is not of a standard form, the equilibrium profile across the interface does not in general have an analytical solution. Practically however, the double-well potential is a reasonable approximation and the concentration profile resembles the hyperbolic tangent solution in equation (6).

It is possible to obtain an estimate of the interfacial width, through consideration of the equilibrium solution to equation (12). Since this equation does not depend explicitly on the spatial co-ordinate, the equilibrium solution must satisfy the Bertrami Identity and it can be shown that,

$$\Delta f = \kappa |\nabla c|^2, \tag{13}$$

for which the integration constant can be eliminated, given that far from the interface both Δf and ∇c are zero. One can then estimate the interface width as (Cahn, Hilliard, 1958),

$$\begin{aligned} \mathrm{d} &= \frac{c^{\beta} - c^{\alpha}}{\max(\nabla c)}, \\ &= \left[c^{\beta} - c^{\alpha}\right] \sqrt{\frac{\kappa}{\max(\Delta f)}}. \end{aligned} \tag{14}$$

The governing equation for the evolution of the phase is simply that of the corresponding conserved quantity, in this case the concentration. However, since the free energy of the system is now a functional of that quantity, the thermodynamic driving force for mass flux must be the functional derivative,

$$\mu = \frac{\delta F}{\delta c}, \\ = \frac{\partial f}{\partial c} - 2\kappa\nabla^2 c. \tag{15}$$

The mass flux is then,

$$\vec{J} = -M\nabla\mu, \tag{16}$$

where the mobility M is related to the diffusivity of the species. The mass balance, which includes the movement of the interface, is then,

$$\frac{\partial \mathrm{c}}{\partial \mathrm{t}} = -\nabla \cdot \vec{J}. \tag{17}$$

The Cahn-Hilliard equation is a fourth order PDE, which can be problematic for the Finite Element Method numerical solvers. Therefore, it is common to split this equation into two second-order PDEs, equations (15) and (7), which is then solved as a coupled system.

8.4. Phase stability, decomposition and nucleation

A key difference between the Allen-Cahn and Cahn-Hilliard equations is in phase nucleation/decomposition. Mathematically, these processes are related to the stabilities of their solutions to infinitesimal fluctuations. Physically, these processes can be triggered by thermal fluctuations, radiation or other, typically stochastic processes which can be introduced into simulations to some degree.

Classical nucleation theory requires the embryo of a new phase to be of a sufficient size such that the liberation of energy from the transformation exceeds the interfacial energy incurred between the original and new phases. Since the embryo is larger than the scale of the interface it must be fully resolved within phase-field models. The random fluctuation must therefore be of a size comparable to the critical nucleus size. This may be implemented through Langevin noise terms in the governing equations, or ad hoc placement of nuclei according to some physical observation (Welland et al., 2009; Simmons et al., 2004; Jokisaari, Permann, Thornton, 2016).

This nucleation barrier allows both models to represent non-equilibrium scenarios such as undercooling. Although the formation of an embryo poses difficulty in phase-field models, the stability of an embryo is well captured and complicating phenomena such as latent heat evolution and depletion zones can be simulated once an embryo is introduced (Jokisaari, Permann, Thornton, 2016). While the standard Allen-Cahn model always requires an embryo of a critical size, this is only true in the Cahn-Hilliard model for the convex regions of the free energy density as shown in Figure 8.5.

For a system with a concentration in the concave region of Figure 8.5., spinodal decomposition can occur. The phases decompose spontaneously, splitting the average state variables between them which, due to the concavity of the free energy profile, results in a lower overall energy. Since the interfacial energy is related to the free energy profile, it does not prevent this decomposition. Mathematically, this result follows from Von Neumann stability analysis applied to the stationary points of equations (11) and (17) which shows that infinitesimal fluctuations, as may occur due to numerical error, are energetically favourable and therefore grow without a barrier.

8.4.1. Solution of the phase-field equations

Both the Allen-Cahn and Cahn-Hilliard equations are PDEs and, with the exception of a few simple cases, require numerical solutions. This may be accomplished through standard real-space numerical methods such as the Finite Element, or Finite Difference methods. Such real-space methods are a mature and common approach and offer adaptive mesh refinement techniques which are useful for resolving the interface over larger spatial length scales (Greenwood et al., 2018). Spectral methods, in which the solution is achieved in Fourier space, offer excellent convergence, however require a regular mesh and impose periodicity of the domain of interest although this may be overcome through the use of 'embedded domains' (Chen, 2002).

In order to resolve the interface shape, and accurately capture variation of the system on the scale of the interface, it is necessary to have a mesh be a fraction of the interface width d. The actual interface width of a system may be measured experimentally through, e.g. transmission electron microscopy, and is typically nanometers in scale. When structures of interest are many times larger than a realistic interfacial width, the simulations may become prohibitively computationally expensive. Therefore, it is common to exaggerate d to be larger than the real interface width (*capillary length*), but still small compared to the feature size of the problem. In some cases, this can result in exaggerated non-equilibrium effects such as solute trapping, which may be compensated for through the introduction of an anti-trapping term to the mass flux equation (Karma, 2001). While a larger interface width results in a more computationally tractable problem, the sensitivity of the solution to the choice of d then must be checked.

8.5. Connection to equilibrium thermodynamics

There is a close connection between phase-field models and CALPHAD (Computer Coupling of Phase Diagrams and Thermochemistry). CALPHAD includes databases of thermodynamic data for pure and solution phases from which phase diagrams may be obtained through free energy minimiser software. The results of CALPHAD calculations are typically for bulk systems and insensitive to mesoscale thermodynamic effects. Integration with phase-field models in a robust fashion rooted in the Theory of Irreversible Processes, introduces interfacial energy and kinetic / non-equilibrium effects and therefore provides an avenue for exploration of these phenomena.

Thermodynamic potentials are introduced directly into the free energy functional through the free energy density terms in equations (1) and (12). However, there remain two main issues to be addressed: how the model discovers the optimal state of a new phase during nucleation, and how the interfacial energy is implicitly related to the thermodynamics of the coexisting phases.

As discussed above, nucleation in either the Cahn-Hilliard or Allen-Cahn model involves an energetic barrier associated with interface creation, which must be overcome by the liberation of energy by the phase transformation. In this case, information about the state of the new phase must be known beforehand from the phase diagram or a coupled free energy minimiser. As an example, in a multicomponent system the species concentrations within the new phase must be known in order to know when the phase change should initiate.

For Cahn-Hilliard models within the spinodal decomposition region, the concavity of the thermodynamic potential allows the system to discover the optimal conditions by itself. The interfacial energy is determined from the thermodynamic potential as the distance

between the energy curve from the tie line. Through this model it is also straightforward to include other sources of interfacial energy, such as coherent elastic strain, wherein the strain varies continuously through the interface (Welland et al., 2015; Ulvestad et al., 2015). However, this definition of the interfacial energy is only strictly valid near the critical point of the miscibility gap where the concavity disappears (Steinbach, 2013).

In the Allen-Cahn model, excess energy is added explicitly through the double-well term. It may then be desirable to avoid implicit energy contributions so that they can be entirely controlled explicitly. This becomes challenging within the interfacial region where, e.g. the concentration is varying between equilibrium states. One approach to eliminating implicit energy contributions is the *interface dissipation method* pioneered by Steinbach et al. where the concentrations of each species in each phase is represented as unknown variables, and uses an 'interface dissipation' process to transfer mass between phases. This method has the drawback of an increasing number of unknown variables, but is able to accommodate a wide variety of expressions for the pure phase potentials (Steinbach, 2013; Zhang, Steinbach, 2012).

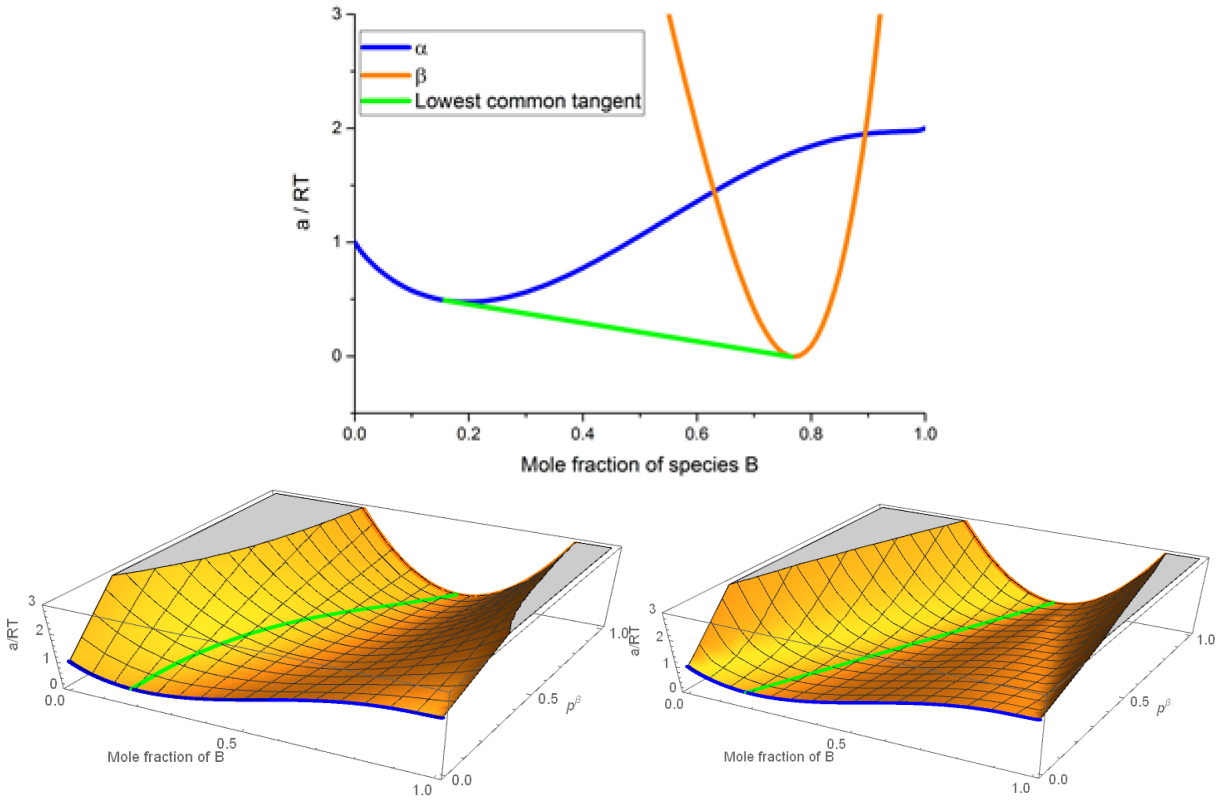
Alternately, a composite potential, $f(\phi, \dots)$, may be defined which combines the pure phase potentials weighted by their phase fractions. Early models considered the concentrations to be equal in all constituent phases throughout the interface (Wheeler, Boettinger, McFadden, 1992),

$$f(\phi, c) = [1 - p]f^{\alpha}(c) + pf^{\beta}(c). \tag{18}$$

In this model, the pure phases are being evaluated at concentrations away from their equilibria which implies variation in the chemical potential throughout the interface, and an excess energy associated with neither phase being in the equilibrium state. This is illustrated in Figure 8.6. where the free energy curves along with the lowest common tangent (tie line) are depicted, along with the composite potential showing a rise in energy along the tie line.

Figure 8.6. Demonstrative plots of a 2 phase binary system

Top: Free energy potentials with the global minimum indicated by the lowest common tangent construction. Bottom left: The composite profile with equal concentrations, showing the excess energy along the green lowest common tangent line. Bottom right: The Grand Potential formulation of the composite potential which is flat along the lowest common tangent line, indicating the lack of implicit interfacial energy.



Source: Welland et al., 2017

More recently, a technique, generalised as the Grand Potential approach, was derived on the premise that the species chemical potentials are equal through the interface, which implies each constituent phase is evaluated at its equilibrium state,

$$f(\phi, c) = [1 - p]f^{\alpha}(c^{\alpha}) + pf^{\beta}(c^{\beta}), \tag{19}$$

subject to the equality of the chemical potentials, $\frac{\partial f}{\partial c} = \frac{\partial f^{\alpha}}{\partial c^{\alpha}} = \frac{\partial f^{\beta}}{\partial c^{\beta}}$. This approach maintains a constant chemical potential through the interface and eliminates the excess energy. However, a problem with the approach is how to determine the phase-concentrations c^{α} and c^{β} that satisfy the chemical potential criterion. This may be done through a nested non-linear solver or through coupling with a free energy minimisation routine (Nestler, Garcke, Stinner, 2005; Choudhury, Nester, 2012). Alternately, if the equilibrium points are known *a priori*, the composite function may be evaluated approximately though, e.g. a local quadratic approximation of the free energy curve of the minimiser of the composite scheme (Choudhury, Nestler, 2012; Welland, Tenuta, Prudil, 2017).

8.6. Summary and future directions

Phase-field models have shown versatility and robustness in their applications to a wide variety of phenomena. Such models are effective in capturing interfacial energy and complex interface morphologies and can be coupled with other multiphysics simulations such as heat and mass transport, elasticity, fluid flow, etc, in a thermodynamically self-consistent fashion. Efficient integration with CALPHAD thermodynamic potential databases is an ongoing area of development, especially with regards to understanding the implicit effects on interfacial energy. Adaptive mesh refinement is also an active and important area of development which holds promise for scaling-up the reach of phase-field models by reducing computational expense. The diffuse interface concept has been applied to other models such as the included-phase model which offer efficient implementations for a subset of phenomena (Prudil, Welland, 2017). There exist a number of phase-field implementations in a variety of platforms and there have been recent attempts to create benchmark problems to compare these implementations (Steinbach, 2013; Jokisaari et al., 2017; Schwen et al., 2017).

Acknowledgements

The author would like to thank Dr Veena Tikare and Dr Nana Ofori-Opoku for reading this manuscript.

References

Ahmed, K. (2017), "Particle-grain boundary interactions: A phase field study", *Comput. Mater. Sci.*, Vol. 134, pp. 25–37.

Allen, S.M. and J.W. Cahn (1972), "Ground State Structures in Ordered Binary Alloys with Second Neighbour Interactions", *Acta Metall.*, Vol, 20, pp. 423–433.

Bellon, P. (2012), "Phase Field Methods", in *Comprehensive Nuclear Materials*, Elsevier, pp. 411–432.

Boettinger, W.J. et al. (2002), "Phase-Field Simulation of Solidification", *Annu. Rev. Mater. Res.*, Vol. 32, pp. 163–194.

Böhler, R. et al. (2014), "Recent advances in the study of UO_2-PuO_2 phase diagram at high temperatures", *J. Nucl. Mater.*, Vol. 448, pp. 330–339.

Borden, M.J. et al. (2012), "A phase-field description of dynamic brittle fracture", *Comput. Methods Appl. Mech. Eng.*, Vol. 217–220, pp. 77–95.

Bourdin, B., G.A. Francfort and J.-J. Marigo (2008), "The variational approach to fracture", *J. Elast.*, Vol. 91, pp. 5–148.

Cahn, J.W, and J.E. Hilliard (1958), "Free Energy of a Nonuniform System. I. Interfacial Free Energy", *J. Chem. Phys.*, Vol. 28, p. 258.

Chakraborty, P., M.R. Tonks and G. Pastore (2014), "Modeling the influence of bubble pressure on grain boundary separatino and fission gas release", *J. Nucl. Mater.*, Vol. 452, pp. 95–101.

Chen, L.-Q. (2002), "Phase-Field Models for Microstructure Evolution", *Annu. Rev. Mater. Res.*, Vol. 32, pp. 113–140.

Choudhury, A. and B. Nestler (2012), "Grand-potential formulation for multicomponent phase transformations combined with thin-interface asymptotics of the double-obstacle potentials", *Phys. Rev. E,* Vol. 85, 021602.

de Groot, S.R and P. Mazur (1984), *Non-Equilibrium Thermodynamics*, Dover Publications, New York.

Devanathan, R. et al. (2010), "Modeling and simulation of nuclear fuel materials", *Energy Environ. Sci.*, Vol. 3, p.1406.

Greenwood, M. et al. (2018), "Quantitative 3D phase field modelling of solidification using next-generation adaptive mesh refinement", *Comput. Mater. Sci.*, Vol. 142, pp. 153–171.

Hohenberg, P.C. and B.I. Halperin (1977), "Theory of dynamic critical phenomena", *Rev. Mod. Phys.*, Vol. 49, pp. 435–479.

Hu, S. et al. (2015), "Assessment of effective thermal conductivity in U-Mo metallic fuels with distributed gas bubbles", *J. Nucl. Mater.*, Vol. 462, pp. 64–76.

Jokisaari, A.M., C. Permann and K. Thornton (2016), "A nucleation algorithm for the coupled conserved-nonconserved phase field model", *Comput. Mater. Sci.*, Vol. 112, pp. 128–138.

Jokisaari, A.M. et al. (2017), "Benchmark problems for numerical implementations of phase field model", *Comput. Mater. Sci.*, Vol. 126, pp. 139–151.

Karma, A. and W.-J. Rappel (1996), "Phase-field method for computationally efficient modeling of solidification with arbitrary interface kinetics", *Phys. Rev. E*, Vol. 53, R3017–R3020.

Karma, A. (2001), "Phase-Field Formulation for Quantitative Modeling of Alloy Solidification", *Phys. Rev. Lett.*, Vol. 87.

Kennouche, D. et al. (2016), "Observing the microstructural evolution of Ni-Yttria-stabilized zirconia solid oxide fuel cell anodes", *Acta Mater.*, Vol. 103, pp. 204–210.

Li, Y. et al. (2010), "Phase-field modeling of void migration and growth kinetics in materials under irradiation and temperature field", *J. Nucl. Mater.*, Vol. 407, pp.119–125.

Li, Y. et al. (2017), "A review: applications of the phase field method in predicting microstructure and property evolution of irradiated nuclear materials", *Npj Comput. Mater.*, Vol. 3.

Millett, P.C. et al. (2012), "Phase-field simulation of intergranular bubble growth and percolation in bicrystals", *J. Nucl. Mater.*, Vol. 425, pp. 130–135.

Nestler, B., H. Garcke and B. Stinner (2005), "Multicomponent alloy solidification: Phase-field modeling and simulations", *Phys. Rev. E*, Vol. 71, 041609.

O'Connor, D.T et al. (2016), "Phase transformation and fracture in single Lix FePO4 cathode particles: a phase-field approach to Li-ion intercalation and fracture", *Model. Simul. Mater. Sci. Eng.*, Vol. 24, 035020.

Provatas, N. and K. Elder (2010), *Phase-Field Methods in Materials Science and Engineering*, Wiley-VCH, Weinheim.

Prudil, A.A. and M.J. Welland (2017), "A novel model of third phase inclusions on two phase boundaries", *Mater. Theory*, Vol. 1.

Schwen, D. et al. (2017), "Rapid multiphase-field model development using a modular free energy based approach with differentiation", *Comput. Mater. Sci.*, Vol. 132, pp. 36–45.

Simmons, J.P. et al. (2004), "Microstructual development involving nucleation and growth phenomena simulated with the phase field method", *Mater. Sci. Eng. A*, Vol. 365, pp. 136–143.

Steinbach, I. (2013), "Phase-Field Model for Microstructure Evolution at the Mesoscopic Scale", *Annu. Rev. Mater. Res.*, Vol. 43, pp. 89–107.

Tonks, M.R. (2014) "Demonstrating the Temperature Gradient Impact on Grain Growth in UO_2 Using the Phase Field Method", *Mater. Res. Lett.*, Vol. 2, pp. 23–28.

Tonks, M.R. et al. (2017), "Mechanistic materials modeling for nuclear fiel performance", *Ann. Nucl. Energy*, Vol. 105, pp. 11–24.

Ulvestad, A. et al. (2015), "Avalanching strain dynamics during the hydriding phase transformation in individual palladium nanoparticles", *Nat. Commun.*, Vol. 6, 10092.

Welland, M.J., B.J. Lewis and W.T. Thompson (2008), "A comparison of Stefan and phase field modeling techniques for the simulation of melting nuclear fuel", *J. Nucl. Mater.*, Vol. 376, pp. 229–239.

Welland, M.J., W.T.Thompson, B.J. Lewis and D. Menara (2009), "Computer simulations of non-congruent melting of hyperstoichiometric uranium dioxide", *J. Nucl. Mater.*, Vol. 385, pp. 358–363.

Welland, M.J. (2012), "Matter Transport in Fast Reactor Fuels", in *Comprehensive Nuclear Materials*, Elsevier, pp. 629–676.

Welland, M.J., D. Wolf and J.E. Guyer (2014), "Multicomponent phase-field model for extremely large partition coefficients", *Phys. Rev. E*, Vol. 89, 012409.

Welland, M.J., D. Karpeyev, D.T. O'Connor and O. Heinonen (2015), "Miscibility Gap Closure, Interface Morphology, and Phase Microstructure of 3D LixFePO4 Nanoparticles from Surface Wetting and Coherency Strain", *ACS Nano*, Vol. 9, pp. 9 757–9 771.

Welland, M.J., E. Tenuta and A.A. Prudil (2017), "Linearization-based method for solving a multicomponent diffusion phase-field model with arbitrary solution thermodynamics", *Phys. Rev. E*, Vol. 95.

Welland, M.J., M.H.A. Piro, S. Hibbins and N. Wang (2017), "A method of integrating CALPHAD data into phase-field models using an approximated minimiser applied to intermetallic layer growth in the Al-Mg system", *Calphad*, Vol. 59, pp. 76–83.

Wheeler, A.A., W.J. Boettinger and G.B. McFadden (1992), "Phase-field model for isothermal phase transitions in binary alloys", *Phys. Rev. A*, Vol. 45, pp. 7 424–7 439.

Zhang, L. and I. Steinbach (2012), "Phase-field model with interface dissipation: Extension to multi-component multi-phase alloys", *Acta Mater.*, Vol. 60, pp. 2 702–2 710.

Zhang, Q. et al. (2016), "Mechanisms of morphological evolution on faceted core-shell nanowire surfaces", *J. Mech. Phys. Solids*, Vol. 91, pp. 73–93.

9. A review of the rate theory of defect clustering in irradiated materials

Dariush Seif and Nasr M. Ghoniem
University of California, United States

Abstract

9.1. Introduction and background

Clustering of atomic defects leads to changes in the microstructure of materials, and hence induces drastic variations in their properties. We present here a review of the theory of atomic defect clustering under non-equilibrium conditions, particularly encountered during irradiation of materials with energetic particles. These conditions are met in a wide range of technical applications, ranging from nuclear and fusion energy to microelectronics and surface engineering. Rate theory methods, which dwell in the intermediate region linking the atomic and macro-scales, provide a useful tool in determining the evolution of irradiated microstructures at experimentally observable time and length scales.

In this chapter, we begin with a brief description of the cascade-damage environment and the physics of point defects. Classical nucleation theory is then presented in the context of point defect clustering. This method is shown to be adequate for obtaining nucleation rates for stable clusters with or without internal solute gas generation. To obtain average cluster densities, rate theory methods are utilised. Starting from a basic understanding of stochastic fluctuations in defect fields, it is shown that one can formulate master equations. When the transition probabilities can be replaced by average macroscopic reaction rates, it is shown that a deterministic set of rate equations for the concentrations of specific defect cluster sizes is sufficient. Following this idea, a reduced-set rate theory model is presented and shown to describe materials under irradiation with high helium generation. Next, a continuum Fokker-Planck approximation is shown for cases when atomic clusters grow or shrink by single atomic defects. When internal sink densities are high, the interaction of point defects with sink stress fields is known to cause a substantial imbalance in the loss of vacancies and interstitials to sinks. This bias is typically implemented in rate theory models to gaugege the effect of point defect segregation on void swelling and creep. We present an ‘effective medium’ rate theory approach to calculate dislocation bias factors in metals using point defect-dislocation interactions based on properties obtained from atomistic simulations. By incorporating point defect fluxes, spatial resolution is attained in the model. The chapter concludes with an assessment of the current state and future of rate-theory models.

Clustering of atomic species in materials is a process that embodies a rich variety of physical mechanisms, and which has significant implications to a wide range of material technologies. Strong deviations from thermodynamic equilibrium conditions drive atomic species into agglomeration, as an attempt to reduce the system’s free energy. Materials inevitably contain atomic defects, in the form of vacancies, self-interstitials, impurity atoms or insoluble gaseous species. In some situations, these defects are externally introduced into the material. Examples of this can be found in structures under neutron irradiation, or in materials that are processed by ion beams, plasmas, high-speed deformation or other high-energy sources. Apart from the valuable insight gained by understanding mechanisms of atomic clustering, advancing a multitude of cutting-edge technologies is dependent on

finding ways to modify the process of atom agglomeration. Energetic neutrons bombard a number of critical materials in nuclear fission reactors. Materials used as nuclear fuels, structural materials, the pressure vessel as well as instrumentation materials, are all subjected to the generation of non-equilibrium concentrations of atomic defects. Energetic neutron collisions with lattice atoms produce intrinsic lattice defects, as well as extrinsic impurities or insoluble gas atoms. Likewise, in a fusion energy system, many materials will be driven out of equilibrium as a result of intense neutron bombardment. Examples are the first wall and blanket structures, plasma interactive components, magnet materials, instrumentation and other special-purpose materials.

In this chapter, we present a review of the theoretical developments in the rate theory treatment of atomic clustering. The theory is applicable to atomic species, which broadly include lattice or surface atoms, molecules, impurities or point defects. The focus here will be to apply the theory to the study of nucleation and growth of precipitates by atomic processes and develop methods to solve for the evolution of their size distribution. To accomplish this task, we begin with a brief description of the cascade-damage environment and the physics of point defects. Classical nucleation theory is then presented in the context of point defect clustering to attain nucleation rates of stable clusters. From a basic derivation of stochastic fluctuations in defect fields, the development of master equations at the atomic level is achieved. A rate theory model is then presented for high helium generation rates using a reduced set of deterministic rate equations for defect species. The Fokker-Planck continuum approximation is then introduced for cases where the agglomeration of atomic clusters is primarily a result of simple atom absorption or emission. Lastly, the effects of stress are discussed in terms of internal point defect sinks and a spatial rate theory solution is developed to calculate dislocation bias factors in irradiated materials.

9.2. Physics of radiation-induced point defects

The generation of point defects in materials stems from two primary types of neutron interactions with bulk materials, known as elastic and inelastic interactions. These are caused by fundamentally different processes and lead to the production of different types of defects within materials. Elastic interactions occur when incoming neutrons transfer their kinetic energy to a host lattice atom. If the energy transferred by the neutron in this collision is greater than the displacement threshold energy, E_{th}, of the lattice atom, then the atom will be ejected from its lattice site. This initial ejected atom is called the primary knock-on atom (PKA). This process leads to the production of a vacancy and a self-interstitial atom (SIA) within the host lattice. Upon ejection from its lattice site, if the PKA still carries greater kinetic energy than E_{th}, it will lead to the ejection of a secondary knock-on atom, and so on. This successive chain of elastic collisions and scattering is referred to as cascade damage and has been studied extensively both experimentally and computationally. Inelastic interactions occur under high-energy irradiation conditions leading to (n, α) transmutation reactions within the host lattice. These interactions lead to the production of significant quantities of solute gas atoms within the bulk, namely helium and hydrogen. Helium is highly insoluble in metals and therefore highly mobile inside matrix materials.

The physics of point defects has been a topic of numerous studies within the field of radiation damage for some time. It is well known that both vacancies and SIA's agglomerate into clusters of different sizes. Vacancies combine to form three-dimensional (3D) voids in the vicinity of the primary damage site while SIA's typically combine to form two-dimensional (2D) glissile or sessile dislocation loops on the periphery. Studies have shown

that glissile loops are highly mobile and perform one-dimensional (1D) random motion in their slip direction. The production of Frenkel pairs and resulting clusters populate the microstructural atmosphere are known to lead to hardening, swelling and creep phenomena in irradiated metals.

In any rate theory study, it is of prime importance to properly identify all critical processes that contribute to the production or loss. In radiation damage rate theory, production terms may be due to collision cascades, emission from clusters, etc., while loss terms may include recombination, absorption into clusters, etc. With the exception of the cascade production term, each process is given a rate proportional to the Boltzmann factor for that process. Thus, it is required that the energetics of all processes be known *a priori*.

The formation energy of a defect is typically defined as the difference in the total energy (at zero temperature) of the system containing the defect and that of a system of the same number of atoms in their lowest energy configurations. These energies are commonly used to determine equilibrium concentrations of defects. As an example, the formation energy of a carbon impurity atom in an iron crystal containing N atoms, can be calculated as

$$E_C^f = E(N+1) - N\mu_{Fe} - \mu_C. \tag{1}$$

Here, $E(N+1)$ is the energy of the system containing the defect and μ_{Fe} and μ_C represent the chemical potentials of iron and carbon, respectively. The chemical potential represents the energy per atom in a perfect, infinite lattice. Calculations of formation energies are best performed using first-principles techniques.

The migration energy, E^m, of a defect represents the activation energy of diffusion within the crystal lattice. Lattice vibrations may result in an atomic jump to a neighbouring site only if the diffusing defect is carrying kinetic energy equal to or greater than the migration energy. Accurate calculations of migration energies may be obtained using first-principles calculations with minimum energy pathway (MEP) search methods such as the nudged elastic band (NEB) method and the dimer method. These methods allow relaxations of the lattice around the defect as the defect treks the reaction co-ordinate to the final state. The point of maximum energy along the path represents the saddle point and the migration energy can be calculated as the difference in energy between the saddle point configuration and the initial state.

The aggregation of point defects following collision cascades leads to the formation of defect clusters in materials. The stability of radiation-induced defect clusters against dissociation is governed by the binding energy, E^B. The binding energy is defined as the energy required to dissociate a bound system into its constituent parts. These energies can be calculated in terms of formation energies under the assumption that the products of dissociation are infinitely separated. For a dissociative reaction of type C→A+B, the binding energy can be calculated as

$$E_C^B = E_A^f + E_B^f - E_C^f. \tag{2}$$

The activation energy required for the emission of a point defect from a cluster is taken as the binding energy plus the migration energy of the point defect. A representative set of values for the formation and migration energies of defects is shown in Table 9.1. Values in bold represent results from first-principles calculations.

Table 9.1. Relevant energy barriers (in eV) for point defect processes in four materials.

Values in bold represent results of first-principles studies.

Process	Fe	W	Cu	Ni
Vacancy Formation	**2.12**[a], **2.15**[b]	**3.11**[c], **3.16**[d]	1.28[e], 1.3[f]	1.62 [g], 1.60 [h,i]
Vacancy Migration	**0.67**[j], **0.67**[b], 0.87[k]	**1.66**[c], **1.78**[b], 1.8[l]	0.72[f]	1.30[h], 1.29[i]
Divacancy Formation	3.18[m]	6.70[n]	2.0[o]	3.02[g]
Divacancy Migration	**0.62**[j]	1.95[p]	0.55[k]	0.59[q]
Interstitial Type	[110]	[111]	[100]	[100]
Interstitial Formation	**3.86**[r], **3.93**[b]	**9.82**[c], **9.548**[b]	2.73[s]	4.91[i], 4.93[s]
Interstitial Migration	**0.34**[j], 0.30[t], 0.167[k]	**0.013**[b]	0.078[u], 0.13[k]	0.16[s]
He Interstitial Form.	**4.40**[v], **3.37**[w]	**6.19**[d], 6.16[c]	1.97[x]	4.06[y]
He Interstitial Mig.	**0.06**[v]	**0.07**[d]	0.5[z], 0.63[α]	0.66[α], 0.81[β]
He Subst. Form.	**4.23**[v], **4.08**[w]	**4.77**[d], 4.70[c]	**2.58**[γ]	3.13[g], **3.23**[γ]

Source: Corresponding sources for the values represented in the table can be found in the References section on page 158 (see the corresponding superscripts following each reference).

9.3. Classical nucleation theory

9.3.1. Becker-Döring theory

Classical nucleation theory, as developed by Becker and Döring in 1935, is a method to develop a set of equations to model phenomena such as condensation, evaporation and in general, any two-phase system (gas-liquid, crystalline-amorphous, etc.). It is especially useful in computations of equilibrium size distributions of phases. In the case of vacancy clustering, assumptions of the theory are as follows:

1. Atomic clusters originate when two atoms come together.
2. Clusters can grow or shrink by the absorption or emission of single vacancies, respectively.
3. Small clusters may dissociate into monomers, however once a critical size is reached, the cluster will grow to a macroscopic size.

If we let C(m,t) represent the number of clusters containing m vacancies at time t per volume, then we can express their rate of change as

$$\frac{\partial C}{\partial t} = \int_1^\infty \{C(m',t)R(m\,|\,m') - C(m,t)R(m'\,|\,m)\}\mathrm{d}m', \tag{3}$$

where $R(m|m')$ is the transition probability rate that a cluster of size m' will grow to size m. We can next define the impingement frequency $\beta_v(m)$ and emission frequency $\alpha_v(m)$ of single vacancies to clusters. These quantities are functions of the vacancy concentration

and diffusion coefficient, as well as the cluster surface area, *A(m)*. Under the second assumption above, the transition probabilities can then be expressed in terms of Dirac delta functions, δ, as

$$R(m \mid m') = \beta_v(m')\delta(m-(m'+1)) + \alpha_v(m'-1)\delta(m-(m'-1)), \tag{4}$$

$$R(m' \mid m) = \beta_v(m)\delta(m'-(m+1)) + \alpha_v(m-1)\delta(m'-(m-1)). \tag{5}$$

Substituting these into equation (3) and evaluating the integral, we obtain

$$\frac{\partial C}{\partial t} = \beta_v(m-1)C(m-1,t) + \alpha_v(m)C(m+1,t) - \beta_v(m)C(m,t) - \alpha_v(m-1)C(m,t). \tag{6}$$

At equilibrium, $\partial C/\partial t=0$, we can invoke the principle of thermodynamic reversibility, i.e. that $R(m\pm1|m)=R(m|m\pm1)$. The corresponding concentration is called the constrained equilibrium function, C_0. This allows $\alpha_v(m)$ to be solved for as

$$\alpha_v(m+1) = \beta_v(m)\frac{C_0(m)}{C_0(m+1)}, \tag{7}$$

which is assumed to be valid in non-equilibrium conditions as well. Plugging the expression for $\alpha_v(m)$ into equation (6) gives

$$\frac{\partial C}{\partial t} = \beta_v(m)C_0(m)\left[\frac{C(m+1,t)}{C_0(m+1)} - \frac{C(m,t)}{C_0(m)}\right] - \beta_v(m-1)C_0(m-1)\left[\frac{C(m,t)}{C_0(m)} - \frac{C(m-1,t)}{C_0(m-1)}\right]. \tag{8}$$

This is known as the difference differential equation for nucleation. Expanding each of the terms in the brackets in Taylor series about n, it can be shown that equation (8) can be written as

$$\frac{\partial C}{\partial t} = \frac{\partial}{\partial n}\left\{\beta_v(m)C_0(m)\frac{\partial}{\partial n}\left[\frac{C(m)}{C_0(m)}\right]\right\}. \tag{9}$$

We note this expression has the form of Fick's second law in a force field with a variable diffusion coefficient. This is consistent with the view that nucleation is random walk process through cluster size phase space. Viewing $\beta_v(m)$ as an effective diffusion coefficient in size space and applying the continuity equation, the cluster nucleation current, I(m) can be taken as

$$I(m) = -\beta_v(m)C_0(m)\frac{\partial}{\partial n}\left[\frac{C(m)}{C_0(m)}\right], \tag{10}$$

and the continuity equation in an *m*-dimensional size space can be given as

$$\frac{\partial C}{\partial t} + \vec{\nabla}\cdot\vec{I}(m) = 0. \tag{11}$$

Thus we have developed a rate equation for the concentration of m-sized voids.

9.3.2. Extension to radiation damage

Following the observations by Cawthorne and Fulton (1967) that the formation of voids occurs in metals exposed to fast neutrons, interest in the application of classical nucleation theory to radiation damage grew rapidly in the early 1970s. The first such undertakings were done independently by Katz and Wiedersich (1971) and Russell (1971). These studies extended the classical theory to irradiation cases, where in addition to vacancies, interstitials can also contribute to growth and shrinkage processes. Also, radiation-induced vacancies and interstitials are produced continuously, leading to the supersaturation of their concentrations beyond equilibrium, ultimately changing the size distribution evolution. The Gibb's free energy (GFE) of a system containing a concentration distribution, $N^{eq}(m)$, of clusters of sizes m, can be expressed as

$$G = G_0 + \sum_m N^{eq}(m) g_m - kT \sum_m \ln W_m, \tag{12}$$

where G_0 is the free energy of the perfect lattice, g_m is the GFE required to form a void of size m, and the last term is the absolute temperature multiplied by the configurational entropy. We note that g_m is typically estimated as the void formation energy (E^f_m) under the assumption of negligible volume and entropy changes in the formation process. Also, W_m represents the number of possible ways in which a void of size m can be distributed in the lattice, which can be shown to be

$$W_m = \frac{m^{N^{eq}(m)} [(\Omega m)^{-1}]!}{[(\Omega m)^{-1} - N^{eq}(m)]! [N^{eq}(m)]!}. \tag{13}$$

The chemical potential of a void of size m can be defined as the partial derivative of G with respect to the equilibrium concentration of size m voids, or

$$\mu_m = \frac{\partial G}{\partial N^{eq}(m)} = E^f_m + kT \ln\left[\Omega N^{eq}(m)\right]. \tag{14}$$

The chemical potential of a single vacancy can therefore be taken as $\mu_v = E_v^f + kT\ln[\Omega C_v]$, where the vacancy formation energy can be substituted from the equilibrium vacancy concentration, given as $C_v^{eq} = \Omega^{-1}\exp(-E_v^f/kT)$. The vacancy chemical potential can therefore be expressed as

$$\mu_v = -kT \ln\left[\Omega C_v^{eq}\right] + kT \ln\left[\Omega C_v\right] = kT \ln S_v, \tag{15}$$

where $S_v = C_v/C_v^{eq}$, is the vacancy supersaturation in the lattice. For the vacancies and voids to be in chemical equilibrium, the chemical potential of voids of size m must be equal to m times the chemical potential of an individual vacancy. Therefore, the equilibrium void distribution function can be solved for from equations (14) and (15) as

$$N^{eq}(m) = \Omega^{-1} \exp(m \ln S_v - E^f_m / kT). \tag{16}$$

Under non-equilibrium conditions, we can define the nucleation current (in the phase space of cluster size) as the rate at which clusters of size m grow to m+1. This can be expressed as

$$I = \beta_v(m)N(m) - \alpha_v(m+1)N(m+1) - \beta_i(m+1)N(m+1). \tag{17}$$

Here, the emission of interstitials as a void growth process is neglected due to the large formation energy of interstitials. By the condition of thermodynamic reversibility, the rate of capture of vacancies to clusters of size m should equal the rate of emission of vacancies of clusters of size m+1, and therefore α_v can be solved for from equation (7), with $C_0(m)$ replaced with $N^{eq}(m)$. Plugging this back into equation (17), we obtain

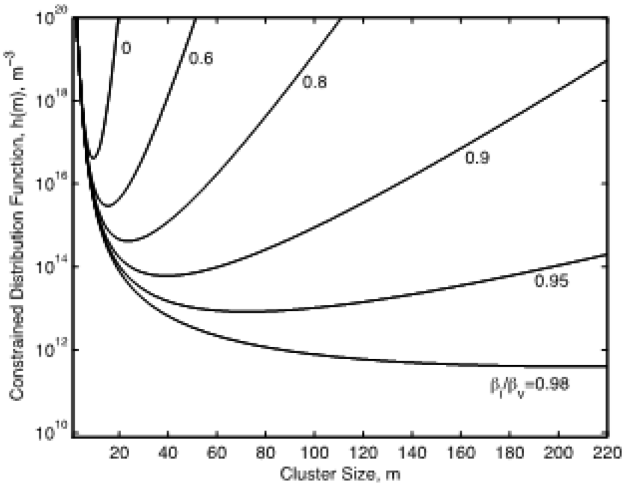
$$I = -\beta_v(m)h(m)\left[\frac{N(m+1)}{h(m+1)} - \frac{N(m)}{h(m)}\right], \tag{18}$$

where the function *h(m)* is defined in

$$\frac{h(m)}{h(m+1)} = \frac{N^{eq}(m)}{N^{eq}(m+1)} + \frac{\beta_i}{\beta_v}. \tag{19}$$

The ratio of impingement frequencies β_i/β_v is referred to as the arrival-rate ratio and plays a large role in determining the critical void size, m_c, which occurs at the minimum of h(m). Figure 9.1. shows a plot of the constrained distribution function, h(m), versus cluster size for a range of arrival-rate ratios. This figure illustrates the fact that nucleation theory is heavily dependent on the parameters used. This is a major disadvantage since in most cases, many parameters are not known to a high degree of accuracy.

Figure 9.1. Constrained distribution function, h(m), versus cluster size, m.



Source: Russell, 1971

Approximating the difference in the brackets of equation (18) as the differential d(N/h)/dm, the resulting differential equation can be solved under the boundary conditions that N/h goes to 1 as m goes to 1 and goes to 0 for infinitely large cluster sizes. The nucleation current can then be calculated as

$$I = \left[\frac{1}{2\pi}\left(\frac{\mathrm{d}^2 \ln h}{\mathrm{d}m^2}\right)_{m_c}\right]^{1/2} \beta_v(m_c) h(m_c). \tag{20}$$

where the first term on the right hand side is called the Zeldovich factor, and corrects for the fact that some fraction of clusters, which have grown past the critical size, *will still shrink to smaller sizes.*

The extension of the theory to include the effects of gas atoms resulting from transmutation reactions was also accomplished by Katz and Wiedersich (1973) and Russell (1972). The derivation of these heterogeneous nucleation models follows the same structure as the preceding model, under the assumption that gas atoms are static nucleation sites for voids. The number of clusters with j gas atoms per unit volume is taken as M_j, making the total concentration of gas atoms, M, equal to $\Sigma_{j=1} jM_j$. The GFE expression takes the modified form:

$$G = G_0 + \sum_j \sum_m \left[N_j^{eq}(m) g_{mj} - kT \sum_m \ln W_{mj} \right], \tag{21}$$

where N_j^{eq} is the equilibrium distribution of gas-vacancy clusters per unit volume containing m vacancies and *j* gas atoms. The configurational term can be calculated as

$$W_{mj} = \frac{M_j!}{[M_j - N_j^{eq}(m)]![N_j^{eq}(m)]!}. \tag{22}$$

The presence of gas atoms in voids leads to an internal pressure that aids in stabilising voids against dissociation. In a study of cavity growth in the presence of helium, Odette and Langley (1975) found that cavities would not grow above a critical size if they contain less than a certain number of helium atoms. Furthermore, a bimodal distribution of cluster sizes was found under irradiation in the presence of gas atoms. The Calculations of Hayns and Wood (1979) also found initial unimodal distributions evolve into bimodal for both diffusion and surface reaction limited growth kinetics under irradiation. Typically observed bimodal distributions consist of one class with cluster radii of a few nanometers and another class with much larger size (Mansur and Coghlan, 1983). The presence of helium has the effect of leading to smaller values of the critical size, as the pressure reduces the effect of surface tension against collapse. Thus, the reversible work of formation g_{mj} will contain a positive contribution due to surface tension and a negative contribution due to the gas pressure. Assuming the gas atoms in the matrix are at an effective pressure p^{eq}, corresponding to the temperature and their concentration, the reversible work of adding gas atoms to a cluster at pressure p is $jkT\ln(p^{eq}/p)$. Assuming the ideal gas law holds, the pressure can be determined from pV=jkT, where V=mΩ. Estimates for p^{eq} can be obtained from statistical mechanics models. The reversible work of cluster formation can then be expressed as

$$g_{mj} = 4\pi R^2 \gamma - jkT \ln\left(p_{eq}/p\right), \tag{23}$$

where γ is the surface energy and the cluster radius R can be taken as $(3m\Omega/4\pi)^{1/3}$. For chemical equilibrium to be satisfied, the condition $m\mu_v=\mu_{mj}$ must be satisfied, where the chemical potential of gas-filled clusters can be taken as

$$\mu_{mj} = \frac{\partial G}{\partial N_j^{eq}(m)} = g_{mj} + kT\ln\left[N_j^{eq}(m)/M_j\right]. \tag{24}$$

Using equation (15) and assuming chemical equilibrium, the equilibrium distribution of gas atoms in voids, can be solved for as

$$N_j^{eq}(m) = M_j \exp(m\ln S_v - \xi m^{2/3} + j\ln(p^{eq}/p)). \tag{25}$$

where $\xi=(4\pi/kT)(3\Omega/4\pi)^{2/3}\gamma$. The exponential term represents the free energy of cluster formation. Whereas in the homogeneous void nucleation case this was a function of m alone, this term represents an energy surface dependent on both m and j. Solution of the nucleation current I_j follows equation (20), with the modification that the critical cluster size and $h(m)$ are now a function of j as well. For more information, Mansur and Coghlan (1983) provide an extensive review of the effect of helium on the critical void size and develop analytical solutions for various conditions.

9.4. Rate theory models of nucleation and growth

9.4.1. General clustering rate theory

Under the non-equilibrium conditions that drive atomic defect clustering, the size of each individual cluster grows or shrinks by a process of accretion or emission. The "step" which changes the cluster size is controlled by fluctuations in the neighbourhood of the cluster. Thus, an appropriate framework for these events is the theory of stochastic processes. Defect clusters are either "born" or they "die" with each event. Many physical phenomena have similar characteristics, and are generally described as birth-and-death processes. In this description, the size of clusters is a randomly distributed variable with each size containing any number of atomic species.

If the size of the transition step is by single atomic defects, then $\Delta x << x$. Here, Δx is the step size and x is a vector of components representing the cluster constituents. Clustering may occur via multiple atom aggregation, and Δx is still smaller than x. Finally, the typical case of coalescence is associated with a transition, where Δx is of the same order of x. In the master equation description, one assigns to the system a set of transition probabilities describing the process of cluster size variation in size space. Consider that the cluster size, x, is discrete, and is represented by a sequence of numbers, each describing the number of a particular specie. Thus, instead of the vector notation, we use the index notation, $C_{ijk...}$, representing the fractional concentration of defect clusters containing i defects of type 1, j defects of type 2, and so on. The types (i,j,k,…) may be lattice atoms, vacancies, gas atoms, etc. The notation can also be extended to more complex defect phases in the solid. Since $C_{ijk...}$ is fractional, it is precisely the probability of finding a defect cluster of these constituents. The transition probabilities, ω, depend on a set of integers, $\Delta_{ijk...}$, which may be positive, negative or zero, and they describe the step size, Δx.

The conditional probability of a transition describing the change of the number (i,j,k,…) due to a step, $\Delta_{ijk\ldots}$, is given by

$$\Delta C_{ijkl\_}(t) = \Delta t \sum_{\Delta} \omega(\{(i,j,k,\ldots) - \Delta_{ijk\_}\} \rightarrow \{i,j,k,\ldots\}) C(\{(i,j,k,\ldots) - \Delta_{ijk\_}\}, t) - \sum_{\Delta} \omega(\{i,j,k,\ldots\} \rightarrow \{(i,j,k,\ldots) + \Delta_{ijk\_}\}) C(\{(i,j,k,\ldots)\}, t) \quad (26)$$

The fundamental relationship represented by equation (26) is the Smoluchowski-Chapman-Kolmogorov (SCK) equation, applied to defect clustering. Taking the limit as $\Delta t \rightarrow 0$, the master equation is obtained, which describes the birth-and-death process of the cluster. Thus, the most general case of defect clustering can be described by the master equation.

It has been shown that if a Poisson distribution describes the transition probabilities ω, their microscopic statistical averages become identical to the macroscopic system averages used in reaction rate theories. Thus, the equilibrium laws of reaction rate theory become exact consequences of the master equation. Averaging over transition probabilities leads to the popular reaction rate theory, which now discards important aspects of defect fluctuations. These are related to the system size, range over which fluctuations extend, and correlation length over which two parts of the system can feel each other. It can be shown that in a system that is macroscopically homogeneous, locally, fluctuations break this homogeneity and lead to the emergence of spatial defect patterns. As soon as the system deviates from the uniform state described by the spatially averaged rate equations, non-linear couplings between neighbouring volume elements take place. As a result, a spatially uniform description of defect concentrations becomes inadequate.

9.4.2. Reduced-set rate theory for high helium generation

The starting point of the theory of microstructure evolution is the SCK equation for a Markovian process. In the continuum limit in size space, equation (26) can be expressed as

$$\frac{\partial C}{\partial t} = \int [\omega(x', x, t') C(x', t') - \omega(x, x', t) C(x, t')] dx' \quad (27)$$

where C(x,t) is the probability distribution for the stochastic variable x at time t and ω(x,x',t) is the transition probability per unit time from state x to state x' at time t. If the transition probabilities can be replaced by average macroscopic reaction rates, then (27) can be reduced to a deterministic set of rate equations for the concentrations of a specific defect cluster sizes. This mean-field approximation does not take into account the statistical nature of defect production, cascade effects, and the arrival and absorption of single and multiple defects at defect clusters. Since defect clustering in irradiated solids is driven by the concentration of three types of monomers (vacancies, interstitials, and helium), we would have a coupled hierarchy of discrete equations for the probability distribution, C, using rate or master equations. Each equation equates the concentration rates to the sum of all production and loss rates of the species, stemming from the known physical phenomena (trapping/de-trapping, recombination, etc). For the case of radiation-induced helium-vacancy (HV) clustering, rate theory models have in the past been developed. Following a study by Ghoniem et al. (1985), rate equations for the concentrations of vacancies (C_v), self-interstitials (C_i), helium interstitials (C_g), helium substitutionals (C_{gv}), di-helium interstitials (C_{2g}), di-helium mono-vacancy (C_{2gv}) and the critical nucleus (C^*) can be expressed as

$$dC_v/dt = fG + (\beta e_1 + \delta)C_{gv} - [\alpha C_i + \beta C_g + \gamma(C_s^v + C_{gv} + 2C_{2g} + 2C_{2gv} + 3C^*)]C_v, \quad (28)$$

$$dC_i/dt = fG - \alpha(C_v + C_{gv} + 2C_{2gv} + 3C^* - C_s^i)C_i, \quad (29)$$

$$dC_g/dt = G_H + (\beta e_1 + \delta + \alpha C_i)C_{gv} + (\beta e_2 + 2\delta)C_{2gv} + 3(\delta + \alpha C_i)C^* + 4\delta C_{2g} + 4\alpha C_i C_{2gv} + m\delta_{tot} + \delta M_{GB} + \delta M_{ppt} - \beta[C_v + 4C_g + C_{gv} + 2C_{2g} + 2C_{2gv} + C_{GB} + \varepsilon_{ppt} C_{ppt}]C_g, \quad (30)$$

$$dC_{gv}/dt = \beta C_g C_v + (e_2 + 2\delta)C_{2gv} - [e_1 + \beta C_g + \delta(1 + C_v) + \alpha C_i]C_{gv}, \quad (31)$$

$$dC_{2g}/dt = 2\beta C_g^2 + 3\alpha C_i C^* - 2(\gamma C_v + \beta C_g + \delta)C_{2g}, \quad (32)$$

$$dC_{2gv}/dt = \beta C_g C_{gv} + 3\delta C^* + 2\gamma C_v C_{2g} - [2\beta C_g + 2\delta(1 + C_v) + \beta e_2 + 2\alpha C_i]C_{2gv}, \quad (33)$$

$$dC^*/dt = 2\beta(C_{2g} + C_{2gv})C_g - 3(\beta C_g + \gamma C_v + \alpha C_i + \gamma)C^*, \quad (34)$$

respectively. Here, the critical nucleus is taken as any cluster containing three helium atoms. The production of Frenkel pairs is taken as the displacement damage rate, G, times the average fraction of surviving pairs, f. Helium atoms are generated at a rate G_H. Annihilation at sinks occurs at rates proportional to the concentrations of sinks, which for grain boundaries and precipitates, are C_{GB} and $C_{ppt,}$ respectively. Absorption of interstitials and vacancies at dislocations is accounted for in the equivalent sink concentrations C_s^i and C_s^v, respectively. Expressions for the concentration rates of helium atoms at precipitates (M_{ppt}) and grain boundaries (M_{GB}) are given as

$$dM_{ppt}/dt = \varepsilon_{ppt}\beta C_{ppt} C_g - \delta M_{ppt}, \quad (35)$$

$$dM_{GB}/dt = \beta C_{GB} C_g - \delta M_{GB}. \quad (36)$$

The reaction frequencies of self-interstitials (α), helium (β) and vacancies (γ) and the thermal emission probabilities of vacancies (e_1,e_2) in the above equations are given as

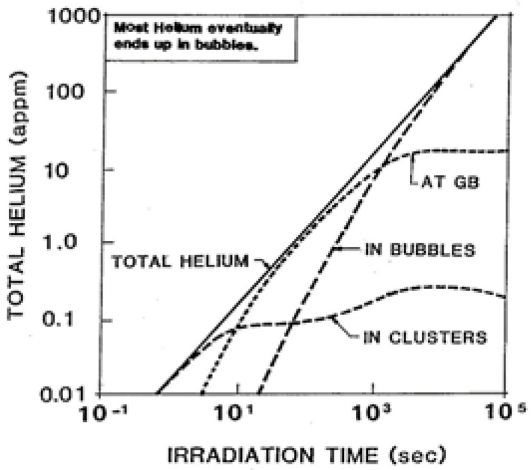
$$\begin{aligned} \alpha &= 48\nu_i \exp(-E_i^m/kT), & e_1 &= \exp(-E_{v,g}^b/kT), \\ \beta &= 48\nu_g \exp(-E_g^m/kT), & e_2 &= \exp(-E_{v,2g}^b/kT), \\ \gamma &= 48\nu_v \exp(-E_v^m/kT), & \delta &= bG, \end{aligned} \quad (37)$$

where the factor of 48 is the assumed combinatorial number, ν is the atomic jump frequency and E^m and E^b represent migration and binding energies, respectively. The factor δ is called the resolution frequency and represents the probability per displacement for dissolving a helium atom back into the matrix, b, times the displacement damage rate.

Solving the above coupled rate equations in time gives insight into the roles of various mechanisms on cluster nucleation rates and their effect on helium transport to grain boundaries. During the initial stages of irradiation, helium atoms are generated as interstitials, but soon are trapped when vacancies become available. These calculations reveal the primary mechanism for reducing helium transport to grain boundaries is the growth of bubbles as available sinks. This is shown in Figure 9.2., where the rate of helium

absorption at grain boundaries reaches saturation while the concentration of helium in matrix bubbles continues to grow.

Figure 9.2. Helium distribution at various traps versus time



Source: Ghoniem et al., 1985

In a study by Stoller and co-workers (2008), a comparison between similar rate theory studies and kinetic Monte Carlo methods was presented. Their findings expose the strong parameter dependence of rate theory methods while arguing that it remains more independent and resistant to errors due to simulation cell sizes.

9.5. Evolution of the size distribution and the Fokker-Planck approximation

In certain applications, the spacing between the states is small, and the hierarchy of discrete master equations can be replaced by an equivalent continuum Fokker-Planck (F-P) equation. It is interesting to note that except for small size interstitial loops, large defect clusters have low mobility inside irradiated solids. On the other hand, when we consider the formation of surface atomic clusters during plasma or ion beam deposition, the mobility of large clusters becomes significant. The transition probability from a defect cluster of size x to x' may be re-defined as

$$\omega(x',x,t') \equiv W(x',x-x',t') \equiv W(x-r,r,t,\tau) \tag{38}$$

We may distinguish between slow and fast variables in the transition probabilities, ω, in the SCK equation, as (x, t) and (r=x-x', τ), respectively. By expanding the function $W(r$=x-x', $\tau)C(r$=x-x', $\tau)$ in a Taylor series for the slow variable, truncating to second order, and integrating over an appropriate correlation time, T, we obtain the F-P equation

$$\frac{\partial C(x,t)}{\partial t} = \frac{1}{2}\sum_{i,j}\frac{\partial^2}{\partial x_i \partial x_j}(\langle r_i r_j \rangle C) - \sum_i \frac{\partial}{\partial x_i}(\langle r_i C \rangle), \tag{39}$$

where $\langle r_i(x,t) \rangle = \frac{1}{T}\int_0^T \int_1^\infty \bar{r}_i \omega(\bar{x},\bar{r},t,\tau) d\bar{r} d\tau$ and $\langle r_i(x,t) r_j(x,t) \rangle = \frac{1}{T}\int_0^T \int_1^\infty \bar{r}_i \bar{r}_j \omega(\bar{x},\bar{r},t,\tau) d\bar{r} d\tau$ represent the first and second moments of the transition probability, respectively. The first moment

represents the elements of a drift vector, while the second is used to derive the elements of a dispersion tensor. The indices i and j represent the dimension in cluster size space (1D for single specie cluster, 2D for a two specie cluster, etc.). The correlation time, T, is chosen to represent the appropriate physics of the relevant transition (for example, inverse of arrival frequency for single-atom transitions or cascade production frequency for cascade-induced transitions). A review of solution methods for the F-P equation can be found in Ghoniem (1998).

9.6. Stress effects in rate theory models and the dislocation bias

The effect of stress on point defect diffusion has been a topic of great interest for some time. It is well known that the presence of internal point defect sinks in irradiated materials plays a significant role in the determining the size distribution. Point defect sinks include dislocations, voids, grain boundaries, precipitates, and so on, each carrying their own associated stress fields. Internal stresses in materials alter the energy landscape for point defect diffusion, leading to various macroscopic changes in material properties over time. A comprehensive theory of sink strengths was presented by Brailsford and Bullough (1981) for various sinks in periodic and random arrays. In general, for a point defect of type η, the associated sink strength for a sink of type α (dislocation, void, etc.) can be defined as

$$k_{\alpha}^{2} = \rho_{\alpha} Z_{\eta}^{\alpha} \tag{40}$$

where ρ_{α} represents the sink density in the material and $Z_{\eta}{}^{\alpha}$ is the capture efficiency of η to α. The capture efficiency represents a ratio of the point defect flux into the sink with the associated drift interaction, to the flux into the core due only to random walk diffusion. Essentially, it is a measure of the effect of drift in the diffusion equation.

It is well known that swelling in irradiated metals is greatly enhanced by the preferential absorption of interstitials rather than vacancies, to dislocations. This preference arises from the stronger interactions interstitials have with dislocation stress fields combined with their high mobility in the matrix. A measure of this preference for interstitials is called the dislocation bias factor and is computed from the ratio of capture efficiencies of dislocations to interstitials and vacancies. The bias can be defined as

$$B = (Z_i - Z_v)/Z_v \tag{41}$$

where Z_i and Z_v are the capture efficiencies of an interstitial and vacancy to a dislocation, respectively. These parameters are of utmost importance specifically in radiation damage rate theory (RT) codes simulating void swelling and radiation-creep. In the past, many authors have developed analytical solutions for the dislocation bias. However, obtaining solutions to diffusion equations with non-spherically symmetric sinks has proven a difficult task. Wolfer and Ashkin (1975) developed a solution for spherical sinks under the assumption that the point defect-sink interaction is strictly radially dependent. This came after Ham's solution (1959) to the steady-state diffusion equation near screw and edge dislocations with point defects modelled as centres of dilatation. For interactions with both radial and angular dependence, Wolfer and Ashkin (1976) developed a perturbation method, extending Ham's work to also include the modulus interaction and externally applied stresses. In all of these studies solutions were reached under numerous assumptions, including approximate treatments of both the point defect fields and interaction energy.

9.7. Space-time rate theory

Recently, we have developed a finite element rate theory model to calculate capture efficiencies and bias factors in irradiated metals. The model implements the diffusive flux term allowing for two-dimensional spatial resolution which is not typically observed in mean-field rate theory calculations. Additionally, the use of atomistically obtained point defect representations (dipole tensors) lead to a much more accurate interaction energy expression than has been used in past studies. Departing from traditional mean-field rate theory and the assumption of continuous sink densities, we consider the 'effective medium' approach. In this method, we may consider a single dislocation centred in a finite medium, with vacancy and interstitial concentrations maintained constant at an outer boundary. These boundary concentrations can be obtained as steady-state solutions using a mean-field rate theory approach. In the vicinity of a defect with an associated strain field, the flux of a diffusing species, η, is given by Fick's first law as

$$\vec{j}_\eta = -\vec{\nabla}(D_\eta C_\eta) - \beta D_\eta C_\eta \vec{\nabla} E_{int} \tag{42}$$

where D_η and C_η are the diffusion coefficient and concentration of species η, respectively, $\beta = (kT)^{-1}$, and E_{int} is the elastic interaction energy of the point defect with the strain field. The capture radius of a species can then be defined such that the condition, $\|\beta \nabla E_{int}\| = b^{-1}$, is satisfied. The capture radius defines the approximate distance from the core at which the drift force due to the interaction energy matches the thermal motion of the defect. It is at this radius that the calculations of the flux are computed, and the diffusion coefficient can be taken as a constant. It is important to note that the capture radius is a function of temperature, material properties, and the defect's dipole tensor, which will be defined in what follows. By applying the continuity equation to equation (42), we obtain the rate equation for the concentration as

$$\frac{\partial C_\eta}{\partial t} + \vec{\nabla} \cdot \vec{j}_\eta = P - L, \tag{43}$$

where P represents production terms for defect η (Frenkel-pair production rate, mission from clusters, etc) and L represents loss terms (recombination, absorption at sinks, etc.). In the effective medium approach, the effect of production and loss terms are accounted for in the outer boundary concentrations, and thus equation (43) can be expressed as

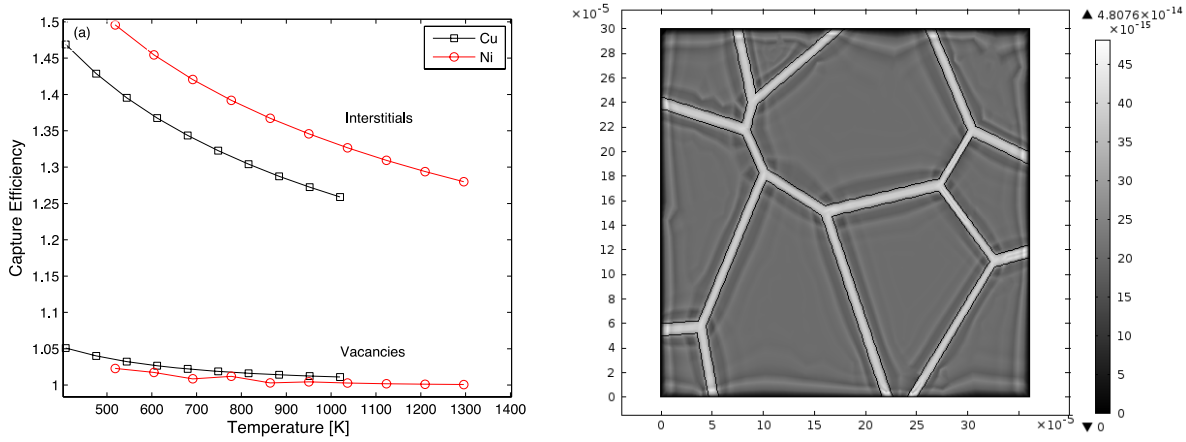
$$\frac{\partial C_\eta}{\partial t} = D_\eta \nabla^2 C_\eta + \beta D_\eta \vec{\nabla} \cdot \left[C_\eta \vec{\nabla} E_{int} \right]. \tag{44}$$

Here, the interaction energy term represents the elastic interaction between the diffusing species and a single edge dislocation. To calculate this interaction, we have recently developed an atomistic-continuum coupling method in which an elastic representation of point defects is made possible by the dipole tensor, which is calculated from the results of molecular statics calculations using interatomic potentials. To first order, the k^{th} component of the continuum displacement field can be calculated in terms of the Green's tensor function G(x) and the defect's dipole tensor $P^{(1)}$ as

$$u_k(\vec{x}) = -G_{kj,i}(\vec{x}) P_{ij}^{(1)}, \tag{45}$$

where repeated indices imply summation. Implementing the anisotropic elastic Green's function and displacements from molecular statics calculations, a linear least squares solution can be constructed to solve for the components of the dipole tensor from (45). In the strain field of an edge dislocation, the elastic interaction energy can be expressed as $P_{ij}^{(1)}u_{j,i}^{\perp}$, where $u_{j,i}^{\perp}$ represents the displacement gradient tensor due to the dislocation. Upon solving for the concentration rates and fluxes, the dislocation bias factor defined by equation (41) can be calculated by first computing the capture efficiencies for the interstitial and vacancy. This is accomplished via line integrations of the dot product of the flux and unit normal along the capture radius. In Figure 9.3.a., the results of these calculations show a decrease in capture efficiencies (and bias) as the temperature is increased. This is a result of thermal fluctuations increasingly competing with drift forces at higher temperatures. In Figure 9.3.b., the results of a separate spatial rate theory model simulating polycrystalline tungsten under high heat flux is shown (Crosby and Ghoniem, 2011). Here, a heat transfer analysis is coupled with the rate equations, and is solved within the framework of the finite element method using a transient segregated solver. By implementing finite element techniques, we have shown that spatial resolution can be obtained via the flux term of interstitials and vacancies.

Figure 9.3. (a) Dislocation capture efficiencies versus Temperature for copper and nickel, (b) Distribution of helium atoms in polycrystalline tungsten under high heat flux



Source: (a) Seif and Ghoniem, 2011, (b) Crosby and Ghoniem, 2011

9.8. Conclusions and research needs

The theory of defect clustering is deeply rooted in statistical physics, and shows fruitful results for the study of a variety of interesting physical phenomena. Under irradiation, the nucleation rate of critical-size voids is shown to be adequately described using classical nucleation theory. To obtain cluster densities, however, rate theory methods need to be utilised. Starting from a basic understanding of stochastic fluctuations in defect fields, it is shown that one can formulate master equations at the microscopic or atomic level. These equations can be treated in a more macroscopic sense by equivalent rate equations, in the limit of Poisson probability distributions for transitions between states. When the transition probabilities can be replaced by average macroscopic reaction rates, then the SCK equation can be reduced to a deterministic set of rate equations for the concentrations of specific defect cluster sizes. A reduced-set rate theory model was presented and shown to describe

materials under irradiation with high helium generation. When the diffusion of large defect clusters in the bulk is slow, and the transitions between states are just between nearest neighbours, a useful approximation is obtained, which is shown to be the Fokker-Planck continuum theory. In materials with large dislocation densities, the interaction of point defects with internal stress fields is known to cause an imbalance in the loss of vacancies and interstitials to dislocation sinks. This dislocation bias is implemented in rate theory models to gauge the bias effect on void swelling and creep. We have shown a novel 'effective medium' rate theory approach to calculate bias factors in metals using point defect-dislocation interactions based on properties obtained from atomistic simulations. By incorporating the flux term, a finite element method was developed which allowed for spatial resolution in the model.

In the decades following early radiation damage rate theory studies, the availability of high-performance computational facilities has grown significantly, enabling large-scale atomistic simulations. The rate theory approach, however, remains unique and useful in multiple ways. With parameters obtained from atomistic methods such as density functional theory and molecular dynamics, rate theory models allow for simulation at the experimentally observable time and length scales unlike atomistic methods. In many cases this allows for additional parameters to be taken directly from known experimental data. Models using the mean field assumption allow for various system statistics to be obtained, however, we have shown that spatial rate theory solutions are also attainable which may be applied to simulations of heterogeneous microstructures. For the continued progress and relevance of rate theory models there are several aspects that need to be addressed and improved upon. With such high sensitivity to physical parameters, emphasis must be placed on developing a large database of defect energetics obtained using the most accurate first-principles methods. Activation barriers for important processes must be known to a high degree of accuracy to ensure the reliability of these models. An important next step is to couple spatial rate theory with other mesoscale methods such as dislocation dynamics, where phenomena such as the embrittlement of nuclear materials can be investigated in terms of dislocation motion processes. Additionally, efforts are needed to develop methods that can incorporate the evolution of the cluster size distribution as a function of space and time in an efficient computational manner that does not compromise the basic physics underlying the rate theory.

References

References for the values from Table 9.1. on page 146 are given as superscripts following each reference.

Ackland, G.J. et al. (2004), "Development of an interatomic potential for phosphorus impurities in α-iron", *J. Phys.: Condens. Matter*, Vol. 15, S2629–S2642. [(m)]

Adams, J.B. and W.G. Wolfer (1989), "Formation energies of helium-void complexes in nickel", *J. Nuc. Mat.*, Vol. 166, pp. 235–242. [(g)]

Balluffi. R.W (1978), "Vacancy defect mobilities and binding energies obtained from annealing studies", *J. Nuc. Mat.*, Vol. 69-70, pp. 240–263. [(l)]

Becker, R. and W. Döring (1935), "Kinetische Behandlung der Keimbildung in übersättigten Dämpfen", *Ann. Phys.*, Vol. 24, p. 719.

Becquart, C.S. and C. Domain (2007), "Ab initio calculations about intrinsic point defects and He in W", *Nuc. Instr. and Meth. in Phys. Res. B*, Vol. 255, pp. 23–26. [(c)]

Brailsford, A.D. and R. Bullough (1981), "The theory of sink strengths", *Philos. Trans. R. Soc. London A*, Vol. 302, p. 87.

Caturla, M.J., C.J. Ortiz and C-C.Fu (2008), "Helium and point defect accumulation: (ii) kinetic modelling", *C.R. Physique*, Vol. 9, pp. 401–408. (v)

Caturla, M.J., et al (2000), "Comparative study of radiation damage accumulation in Cu and Fe", *J. Nuc. Mat.*, Vol. 276, pp. 13–21. (k)

Cawthorne, C. and E.J. Fulton (1967), "Voids in Irradiated Stainless Steel", *Nature,* Vol. 216, p. 575.

Chen, Z. et al. (2010), "Strain-field effects on the formation and migration energies of self interstitials in α-Fe from first principles", *Phys. Rev. B*, Vol. 81. (r)

Crosby, T. and Ghoniem, N.M. (2013), "Multiphysics model of thermomechanical and helium-induced damage of tungsten during plasma heat transients"*, J. Nucl. Mat*., Vol. 442, p. 261.

Derlet, P.M., D. Nguyen-Manh, and S.L. Dudarev (2007), "Multiscale modelling on crowdion and vacancy defects in body-centered-cubic transition metals", *Phys. Rev B*, Vol. 76. (b)

Dublek, G., et al (1977), "A study of mono and divacancies in Cu and Au by positron annihilation", *Appl. Phys.*, Vol. 13, pp. 67–70.(o)

Foiles, S.M., M.I. Baskes and M.S. Daw (1986), "Embedded-atom-method functions for the fcc metaks Cu, Ag, Au, Ni, Pd, Pt, and their alloys", *Phys. Rev. B*, Vol. **33**(12). (e)

Fu, C.C., and F. Willaime (2008), "First principles calucations in iron: structure and mobility of defect clusters and defects complexes for kinetic modelling", *C.R. Physique*, Vol. 9, pp. 335–342. (j)

Fu, C.C., and F. Willaime (2005), "*Ab initio* study of helium in α−Fe: Dissolution, migration, and clustering with vacancies", *F., Phys. Rev. B*, Vol. 72.(a)

Ghoniem, N.M., J.N. Alhajji, and D. Kaletta (1985), "The Effect of Helium Clustering on Its Transport to Grain Boundaries", *J. Nuc. Mat.*, Vol. 136, p. 192.

Ghoniem, N.M. (1999), "Clustering theory of atomic defects", *Rad. Eff. and Def. in Sol.*, Vol, 148:1-4, p. 269.

Golubov, S.I., B.N. Singh, and H. Trinkaus (2000), "Defect accumulation in fee and bcc metals and allows under cascade damage conditions – Towards a generalisation of the production bias model", *J. Nuc. Mat.*, Vol. 276, pp. 78-89.

Ham, F.S. (1959), "Stress-Assisted Precipitation on Dislocations", *J. Appl. Phys.*, Vol. 30, p. 915.

Hayns, M.R. and M.H. Wood (1979), "A numerical study of evolving void-size distribution functions", *J. Nuc. Mat.*, Vol. 87, p. 97.

Johnson, R.A. (1983), "Point-defect calculations for tungsten", *Phys. Rev. B*, Vol. 27(4). (p)

Katz, J.L. and H. Wiedersich (1971), "Nucleation of Voids in Materials Supersaturated with Vacancies and Interstitials", *J. Chem. Phys.*,Vol. 55, p. 1 414.

Katz, J.L. and H. Wiedersich (1973), "Effect of insoluble gas molecules on nucleation of voids in materials supersaturated with both vacancies and interstitials", *J. Nuc. Mat.*, Vol. 46, p. 41.

Kuhalainan, J., M. Manninen and E. Kautto (1996), "Helium, neon and argon in copper studied with effective medium theory", *J. Phys.: Cond. Matt.*, Vol. 8:10317. (z)

Mansur, L.K. and W.A. Coghlan, "Mechanism of helium interaction with radiation effects in metals and alloys: A review", *J. Nuc. Mat.*, Vol. 119, pp. 1-25.

Megchiche, E.H., C. Mijoule and M. Amarouche (2010), "First principles of calculation of vacancy-vacancy interactions in nicel: thermal expansion effects", *J. Phys.: Cond. Matt*, Vol. 22:485502. (q)

Mhaskar, S.P., and Naik, P.K., "Effect of surface relaxation on point defects in the vicinity of W (100) surfaces", *Phys. Stat. Sol. (b)*, Vol. 105, pp. 685–689, 1981. (n)

Mishin, Y. et al (1999), "Interatomic Potentials for Al and Ni from Experimental Data and Ab initio Calculations", in *Multiscale Modelling of Materials. Symposium*, Cambridge University Press, Cambridge, pp. 535–540. (i)

Morishita, K. et al. (2000), "A molecular dynamics simulation study of small cluster formation and migration in metals", *J. Nuc. Mat.*, Vol. 283-287, pp. 753–757. (u)

Odette, G.R. and S.C. Langley (1975), "Modeling of synergistic effects of displaced atom of transmutant damage in fission and fusion enrivonments" in J.S. Watson, and F.W. Wiffen (eds) *Proc. Int. Conf. on Rad. Eff. and Tritium Tech. for Fusion Reactors*, Gatlinburg, Tennessee, CONF-750989.

Phillips, V., K. Sonnenberg and J.M. Williams (1982), "Diffusion of helium in nickel", *J. Nuc. Mat*, Vol. 107, pp. 271–279. (β)

Russell, K.C. (1971), "Nucleation of voids in irradiated metals", *Acta Met.*, Vol. 19, p. 753.

Russell, K.C. (1972), "Thermodynamics of gas-containing voids in metals", *Acta Met.*, Vol. 20, p. 899.

Schultz, H. (1991), "Atomic Defects in Metals", in *Atomic Defects in Metals - Fe*. Landolt-Börnstein - Group III Condensed Matter, Vol. 25, H. Ullmaier (ed.). (t)

Seif, Dariush, and Nasr M. Ghoniem (2013), "Effect of anisotropy, SIA orientation, and one-dimensional migration mechanisms on dislocation bias calculations in metals", *J. Nucl. Mat.*, Vol. 442.1-3, p. 633.

Seletskaia, T. et al. (2005), "Magnetic Interactions Influence the Properties of Helium Defects in Iron", *Phys. Rev. Lett.*, Vol. 94. (w)

Stoller, R.E., A.S. Kumar and D.S. Gelles (1992), *Eff. of Rad. on Mat.:15th Int'l Symp,* American Society for Testing and Materials (ed.), Philadelphia, pp. 517-529. (α)

Stoller, R.E. et al. (2008), "Mean field rate theory and object kinetic Monte Carlo: A comparison of kinetic models", *J. Nuc. Mat.*, Vol. 382:2-3, pp. 77-90.

Wilson, W.D, M.I. Baskes and C.L. Bisson (1976), "Atomistics of helium buddle formation in a face-centered-cubis metal", *Phys. Rev. B*, Vol. 13(6). (x)

Wolfer, W.G. and M. Ashkin (1975), "Stress-induced diffusion of point defects to spherical sinks", *J. Appl. Phys.*, Vol. 46, p. 547.

Wolfer, W.G. and M. Ashkin (1976), "Diffusion of vacancies and interstitials to edge dislocations", *J. Appl. Phys.*, Vol. 47, p. 791.

Wycisk, W. and M. Feller Kniepmeier (1978), "Quenching experiments on high-purity nickel", *J. Nuc. Mat*, Vol. 69-70, p. 616. (h)

Yang J. et al. (2006), "The Formation Energies and Binding Energies of Helium Vacancy Cluster: Comparative Study in Ni and Pd", *J. of Phys.: Conf. Series*, Vol. 29, pp. 190–193. (y)

Yang, L., H. Liu and X. Zu (2009), "First-principles study of the migration of helium in tungsten", *Intl. J. of Mod. Phys. B*, Vol. 23(8). (d)

Zhao, P. and Y. Shimomura (1999) "Molecular dynamics calculations of properties of the selfinterstitials in copper and nickel", *Comp. Mat. Sci.*, Vol. 14, pp. 84–90. (s)

10. Finite element based material modelling of nuclear fuel

Joseph Y. R. Rashid
ANATECH Corp., United States

10.1. Introduction

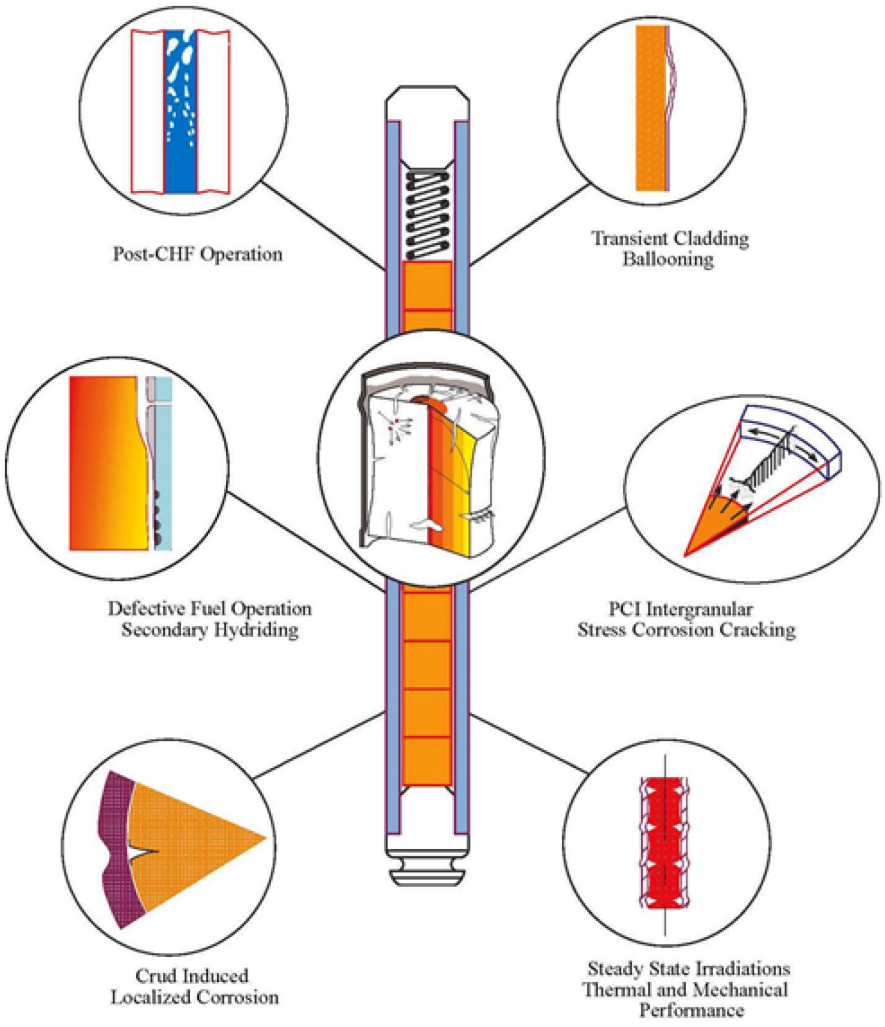
In light water reactors (LWRs) technology, the term "nuclear fuel" is understood to apply to a system of fuel rods each of which is composed of two major sub-elements: the ceramic UO2 fuel pellets and the zirconium-based alloy cladding. The former is the fission energy source, and the latter is the heat transfer medium, which also serves as the primary containment of fission products. Modelling of the thermo-mechanical behaviourr of both materials, individually and as interacting components of a fuel rod, is one of the most important elements governing reactor performance. In a full-length LWR fuel rod, the fuel column consists of several hundred cylindrical pellets, which, during normal operation, experience fracturing and relocation, swelling, densification, and viscoplastic hot pressing. As contiguous bodies, the fuel pellets are subject to rigid body motion and cannot be isolated, geometrically and functionally, from the cladding. Consequently, they derive their coherency from cladding confinement and, as such, their thermo-mechanical modelling requires special numerical techniques to transform a geometrically discontinuous behaviourr to a computationally stable continuum. The finite element modelling and simulation of this thermo-mechanical behaviourr produces a non-linearlinear system of equations, where the fundamental solution variables are the strongly coupled displacement and temperature fields. For the coupled displacement and temperature solutions, the temperature response will depend on the displacements through the pellet-cladding gap thermal conductance, and the mechanical response variable – the stress field – will depend on the temperature through the material properties and behaviouralral models. This dependency is reflected in both, the constitutive relations and the nodal force-displacement-temperature system of governing equations.

A recent review of the current state of the art in LWR nuclear fuel modelling is presented by Rashid, Yagnik, Montgomery (2011), placing special emphasis on pellet-clad mechanical interaction (PCMI) which governs fuel rods integrity during normal operations and under accident conditions. The purpose of the present paper is to describe the materials modelling of fuel viscoplasticity and fracture (Rashid, Tang, Johansson, 1974; Rashid, 1974), which are the behavioural regimes that determine the intensity of PCMI and cladding failure potential in LWRs. The computational implementation via the finite element method (FEM) has achieved a mature level of practical application, but is confined to two-dimensional simulation only (Rashid, Montgomery, Dunham, 2004), with three-dimensional modelling limited to a few special effects studies (Brochard, Bentejac, Hourdequin, 1997). However, full three-dimensional FEM simulations of general utility have recently been undertaken (Larzelere, 2010, CASL), but still remain in the formative stage.

The nuclear fuel rod material and spatial domains are: the ceramic fuel pellets, individually as contiguous cylindrical bodies or as a continuous fuel column; the void volume and pellet-cladding gap interface; the metal cladding; and the surrounding coolant channel. The thermo-mechanical response of such fuel element is intimately connected with the surrounding material; therefore, to convey to the reader a coherent picture of fuel behaviour

modelling from which an assessment of adequacy or future improvements can be made, it is instructive to describe the physical environment in which nuclear fuel operates. This is depicted pictorially in Figure 10.1., which shows the various physical processes and behavioural regimes that result from, and are driven by, the response of the fuel. These regimes are depicted through inserts in Figure 10.1., and the interaction regimes represented in this figure are greatly influenced by the fuel's constitutive behaviour and thermo-mechanical response.

Figure 10.1. Interfaces and behavioural regimes in LWR fuel



Source: Rashid, Y.R., S.K. Yagnik and R.O. Montgomery, 2011.

The mechanical response of the fuel pellets to the temperature spatial distribution causes the pellet to hourglass and fracture radially and axially in a semi-random pattern, with radially oriented cracks as dominant feature. Circumferential cracks can occur under cooling during power reduction, and the resulting fracture pattern becomes a three-dimensional (3D) fragmentation pattern. As the pellet fragments relocate radially outward under time varying power history, the resulting change in the fuel-cladding gap can produce a change in the gap conductance, with a corresponding change in the heat transport to the

cladding and the surrounding coolant. Pellet-cladding contact eventually begins and the resulting mechanical interaction (PCMI) can exert localised hoop stress at the cladding inner surface at radial pellet cracks, which, in the presence of fission products can lead to stress corrosion cracking (SCC) failure of the cladding. This failure mechanism came to be known by the acronym PCI, which characterises the chemical effect of PCMI, and is the primary failure mechanism under normal operations. The PCMI/PCI effects are further aggravated by the combined damage of pellet radial cracks and manufacture-induced chipped pellet corners, known as missing-pellet-surface (MPS). Under transient conditions, cladding heat-up can occur due to inefficient heat transfer to the coolant caused by one of several transient regimes such as power-coolant mismatch, reactivity insertion accident (RIA) or loss of coolant accident (LOCA). The resulting cladding deformations during such events lead to a loss of confinement of the fractured fuel pellets and induce rigid-body like relocation motion. A high fidelity fuel model must have the capabilities to account for all of these behaviour regimes (Rashid, Montgomery, Dunham, 2004).

Under the thermal-hydraulic and the thermo-mechanical behaviour regimes represented in Figure 10.1., both fuel and cladding material elements are subject to viscoplasticity and fracture, evolving differently in the two materials, with the ceramic fuel exhibiting the more complex constitutive behaviour. In the present mathematical construct, emphasis is given to the ceramic fuel pellet, which, with appropriate restrictions on certain model parameters, can be generalised to the cladding material. In this respect, we identify two unique features that distinguish the constitutive behaviour of fuel pellets from that of the cladding, as well as metallic materials in general. The first unique feature is that the ceramic pellet, being a porous compact of sintered uranium oxide powder, exhibits non-zero volumetric viscoplastic dilatation, as compared to zero plastic volume change in metals. The second unique feature is that fracture in the fuel pellet is an inherent behaviour regime under normal operating conditions, and consequently is modelled as a constitutive property in the stress-strain relations. This results in non-linear time-dependent force-displacement equations during all deformation stages, including small deformations where linear behaviour is expected in other materials. In contrast, fracture in the cladding is calculated as an end limit state, for comparison with a failure criterion, and, unlike the fuel pellet, is seldom treated during the course of the computations as a propagating fracture.

FEM is uniquely suited for the computational simulation of the above described thermo-mechanical behaviour, especially in the treatment of pellet fracture, which is modelled as multi-axial direction-dependent distributed damage. This type of modelling fracture frees the computational simulation of crack evolution and propagation from the rigid constraints of the computationally averse classical fracture mechanics. The constitutive formulation presented below reflects this unique adaptability of the FEM to the treatment of the interdependency between the material behaviour and the numerical simulation of the overall global response.

10.2. Modelling of UO_2 fuel viscoplasticity

Fuel viscoplastic behaviour under plastic flow and creep is modelled by first prescribing a yield potential function and flow rule that admit plastic volume change, which was characteried experimentally as hot pressing (Kaufman, 1961). A yield potential function is chosen that is essentially an extension of the von Mises yield condition to include a dependence on the first invariant of the stress tensor. Similarly, a creep flow rule is postulated in which a creep volumetric compression is assumed to be related to a Poisson's ratio type parameter. Further more, the yield surface is non-symmetric in the stress space,

as will be shown later, and the material behaves elastically in tension until it fractures. The yield function has the form

$$F = J_2 + \alpha J_1^2 - k^2 = 0, \tag{1}$$

The associated flow rule of this function is

$$\dot{\varepsilon}_{ij}^{p} = \lambda\left(S_{ij} + 2\alpha J_1 \delta_{ij}\right) \tag{2}$$

And the plastic volumetric incremental change is

$$\dot{\varepsilon}_{ii}^{p} = 6\lambda\alpha J_1 \tag{3}$$

Equation (3) states that the plastic volume change is proportional to and of the same sign as J_1, hence, equation (1) provides for plastic densification under isotropic compression. The quantities J_1 and J_2 in Yield function (1) are defined as follows:

$$J_1 = \sigma_{ii}, \tag{4}$$

$$J_2 = \frac{1}{2} S_{ij} S_{ij}, \tag{5}$$

where σ_{ij} are the co-ordinate stresses, S_{ij} are the deviatoric stresses given by,

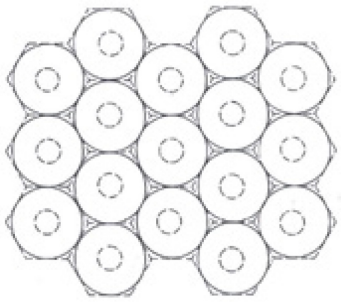
$$S_{ij} = \sigma_{ij} - \frac{1}{3}\sigma_{kk}\delta_{ij}. \tag{6}$$

10.2.1. Determination of α

An assessment of this parameter is made by considering an idealised model of porous material. In doing so, the following assumptions are made:

1. Macroscopically, the porous material is viewed as a homogeneous continuum. This continuum consists of numerous cells structured as spherical pores surrounded by spherical shells.
2. These shells are closely packed as shown in Figure 10.2. and the interstitial material between the shells transmits forces but does not contribute to the strength of the material.
3. The cell is much larger than a microscopic element; therefore, its stress and strain fields can be analysed within the framework of continuum mechanics.

Figure 10.2. Plane cross section of the idealised model of porous material (not to scale)



Source: Rashid, Y.R., H.T. Tang and E.B. Johansson (1974).

For the array shown in Figure 10.2., 74% of the material volume is in the spherical shells and 26% is in the interstitial material. Under hydrostatic compression, the spherical shell will yield if

$$P = 2Y_s \ln(r_2/r_1), \tag{7}$$

where, Y_s is the yield strength of the solid material surrounding the pores; r_2 and r_1, are outer and inner radii of the shell; and P is the pressure required to yield the shell. Equation (7) is obtained straightforwardly by solving the equilibrium equation subject to the von Mises yield criterion (Hill, 1950). The volume fraction of the pore referred to the spherical shell is $(r_1/r_2)^3$ and referred to the total volume is $0.74(r_1/r_2)^3$. If we let ρ be the theoretical density fraction referred to the total volume, such that $1-\rho = 0.74(r_1/r_2)^3$, and substituting for r_1/r_2 into equation (7) gives,

$$P = \frac{2}{3} Y_s \ln\{0.74/(1-\rho)\} \tag{8}$$

This relation can be used together with the yield criterion to determine a value for α. In the case of hydrostatic compression, $J_2 = 0$, $J_1 = 3P$, and equation (1) has the form

$$9\alpha P^2 = K^2. \tag{9}$$

At initial yielding $K = (1/\sqrt{3})\ Y(\rho_0)$ and eq. (9) gives

$$\alpha = \left\{Y(\rho_0)/3\sqrt{(3)}P\right\}^2, \tag{10}$$

where, ρ_0 is the initial density fraction, and $Y(\rho_0)$ is the uniaxial yield stress of the porous material at density fraction ρ_0. Substituting equation (8) into equation (10) and setting $\rho = \rho_0$ yields

$$\alpha = \left[\frac{Y(\rho_0)/Y_s}{2\sqrt{3}\ln\{0.74/(1-\rho_0)\}}\right]^2. \tag{11}$$

The numerator in equation (11) is the ratio of the flow stress at the initial density to the flow stress of the solid material surrounding the pores. It is seen that as ρ_0 becomes larger, $Y(\rho_0)$ approaches Y_s, and if $\rho_0 = 1$, $\alpha = 1$, which corresponds to von Mises material. In the above,

ρ_0:	Initial (as-manufactured) fractional density of the fuel
$Y(\rho_0)$:	Initial yield strength of the fuel
Y_S:	Yield strength of solid (100% dense) UO_2

Except for Y_S, the parameters in equation (11) are known for a typical fuel design. Because it is virtually impossible to determine Y_S experimentally, and porosity dependent yield strength data is not available for extrapolation to zero porosity, it is necessary to estimate Y_S in order to make use of the above expression for α. However, it is possible to avoid this process by allowing $Y(\rho_0)$ to approach Y_S as ρ_0 approaches unity and then calculate an asymptotic value for α. Using this procedure α is calculated as follows:

$$\alpha = \lim_{\rho_0 \to 1} \left[1/\left(2\sqrt{3}\ \ln\left(0.74/\left(1-\rho_0\right)\right)\right)\right]^2 \doteq 0.005 \tag{12}$$

This value of α was found to give best results with respect to observed behaviour of PCMI.

10.2.2. Creep law

Utilising a creep flow rule, the creep strain rates are determined from the following pair of flow rules:

$$\dot{e}_{ij}^{c} = JS_{ij}, \tag{13}$$

and

$$\dot{\varepsilon}_{ii}^{c} = L\sigma_{ii}, \tag{14}$$

where, J and L are general functions of stress, strain, temperature and $\dot{\varepsilon}_{ij}^{c}$ are the deviatoric strain rate tensor defined by

$$\dot{e}_{ij}^{c} = \dot{\varepsilon}_{ij}^{c} - \frac{1}{3}\dot{\varepsilon}_{kk}^{c}\delta_{ij}. \tag{15}$$

Combining equations (13), (14), and (15) yields

$$\dot{\varepsilon}_{ij}^{c} = JS_{ij} + \frac{1}{3}L\sigma_{kk}\delta_{ij}. \tag{16}$$

Contracting both sides of equation (13) and substituting the following effective strain and effective stress expressions:

$$\dot{\varepsilon}_{e}^{c} = \left(\frac{3}{2}\dot{e}_{ij}^{c}\dot{e}_{ij}^{c}\right)^{1/2} \tag{17}$$

$$\sigma_e = \left(\frac{3}{2} S_{ij} S_{ij}\right)^{1/2}, \tag{18}$$

gives

$$J = \dot{\varepsilon}_e^c / \sigma_e. \tag{19}$$

where J is seen to be the specific creep rate in shear, which is generally a function of time, stress, temperature and neutron flux.

Returning to the second flow rule which defines the dilatational creep rate, equation (14), we observe that the function L is the specific creep rate under hydrostatic stress, which, like J can be derived from creep tests generated under 3D stress conditions. Such tests, however, are limited and therefore assumptions have to be made to permit stress analysis. In order to make use of a single creep rate function, namely J, the dilatational creep response is assumed to have the same time dependence as the creep in shear, implying that L would differ from J by a constant. This leads to the assumption that Poisson's ratio for creep remains constant with time but may take on a different value for creep from the elastic value. On this basis, the combined flow rule, equation (16), becomes:

$$\dot{\varepsilon}_{ij}^c = J B_{ijkl} \sigma_{kl}, \tag{20}$$

where,

$$B_{ijkl} = (0.5 + \nu_0)\delta_{ik}\delta_{jl} - (0.5 - \nu_0)\delta_{ij}\delta_{kl} \tag{21}$$

and ν_0 is a new parameter which governs the dilatational creep response. The value of this parameter falls in the range

$$0 \le \nu_0 \le 0.5 - \nu \tag{22}$$

where, ν is the Poisson ratio for creep. From equations (21) and (22) it is clear that if the creep volumetric compression parameter ν_0 is zero then the material is incompressible under creep, in which case the combined flow rule, equation (20), reduces to the usual flow rule, equation (13).

10.2.3. Determination of ν0

The value of ν_0 ranges from zero for incompressible material to (0.5 – ν), where ν is the Poisson ratio for creep of the porous material. In order to establish a value for this parameter, we employ the same idealised material model used in the determination of the plastic compression parameter α. Assume that the material surrounding the pores obeys a creep law of the form

$$\dot{\varepsilon} = A\sigma^n, \tag{23}$$

where, A and n are material constants. Under hydrostatic compression, the tangential creep rate at the outer surface of the spherical shell is solved in Hill (1950) and has the form

$$\dot{\varepsilon}_t = \frac{A}{2}\left(\frac{3}{2n}\right)^n \left\{\frac{1}{\left(r_2/r_1\right)^{3/n}-1}\right\}^n P^n \tag{24}$$

where, r_2, r_1 and P are defined in equation (7). Substituting the expression $1\text{-}\rho = 0.74(r_1/r_2)^3$ for r_1/r_2 into equation (24), we obtain

$$\dot{\varepsilon}_t = \frac{A}{2}\left(\frac{3}{2n}\right)^n \left[\frac{1}{\left\{0.74/\left(1-\rho\right)\right\}^{1/n}-1}\right]^n P^n \tag{25}$$

Equation (25) is equal to $\dot{r}/r_2$ at $r = r_2$ with $\dot{r}$ being defined as the rate of the radial displacement of the spherical shell. Viewing the spherical shell as the basic element composing the porous material, $\dot{\varepsilon}_t$ of equation (25) is then the uniaxial strain rate of the porous material. Consequently, the volumetric strain rate is

$$\dot{\varepsilon}_p = 3\dot{\varepsilon}_t \tag{26}$$

From equation (23) and the modified creep law, the volumetric strain rate under hydrostatic compression is

$$\dot{\varepsilon}_n = 6\nu_0 A P^n \tag{27}$$

Thus, treating the porous material as a homogeneous material, the creep parameter ν_0 can be determined by substituting equation (25) into equation (26) and setting $\dot{\varepsilon}_p = \dot{\varepsilon}_n$, which yields

$$\nu_0 = \frac{1}{4}\left(\frac{3}{2}\right)^n \left[\frac{1}{\left\{0.74/\left(1-\rho\right)\right\}^{1/n}-1}\right]^n \tag{28}$$

As ρ approaches 1 ν_0 approaches 0, and at 100% density the material becomes creep incompressible.

10.2.4. Derivation of constitutive equations

Consider a material element where it is assumed that the strain rate tensor $\dot{\varepsilon}_{ij}$ can be expressed in incremental form as the sum of four components: elastic, time-independent plastic, creep and free expansion; namely,

$$d\varepsilon_{ij} = d\varepsilon_{ij}^{e} + d\varepsilon_{ij}^{p} + d\varepsilon_{ij}^{c} + d\varepsilon_{ij}^{f} \tag{29}$$

The elastic strains are related to the incremental stress by

$$d\varepsilon_{ij}^{e} = d\left(C_{ijkl}\,\sigma_{kl}\right) = C_{ijkl}\,d\sigma_{kl} + dC_{ijkl}\,\sigma_{kl} \tag{30}$$

or by the inverse relation

$$d\sigma_{ij} = d\left(D_{ijkl}\,\varepsilon^{e}_{kl}\right) = D_{ijkl}\,d\varepsilon^{e}_{kl} + dD_{ijkl}\,\varepsilon^{e}_{kl} \tag{31}$$

Equations 30 and 31 admit the variation of the material tensors C_{ijkl} and D_{ijkl} with temperature and neutron flux. For isotropic material, which characterises UO_2 pellets, they reduce to

$$C_{ijkl} = \frac{1-\upsilon}{E}\,\delta_{ik}\,\delta_{jk} - \frac{\upsilon}{E}\,\delta_{ij}\,\delta_{kl} \tag{32}$$

$$D_{ijkl} = \frac{E}{1-\upsilon}\,\delta_{ik}\,\delta_{jl} + \frac{\upsilon E}{(1+\upsilon)(1-2\upsilon)}\,\delta_{ij}\,\delta_{kl} \tag{33}$$

where, E is Young's modulus and υ is Poisson's ratio.

The plastic strain increments must obey a flow rule of the form

$$d\varepsilon^{p}_{ij} = d\lambda\left(\partial F/\partial\sigma_{ij}\right) \tag{34}$$

where, F is defined in equation 1 and $d\lambda$ is a positive scalar function which depends on the state of stress, strain, temperature, hardening and neutron flux. Combining equations 20, 21, 29, 30, 31 and 34 the incremental stress-strain relations may be written as

$$d\sigma_{ij} = D_{ijkl}\left(d\varepsilon_{kl} - d\lambda\frac{\partial F}{\partial\sigma_{kl}} - dtB_{klmn}J\,\sigma_{mn} - dC_{klmn}\,\sigma_{mn} - d\varepsilon^{f}_{kl}\right) \tag{35}$$

In the above equation, the scalar function $d\lambda$ remains to be determined. This is done with the aid of the consistency condition which ensures that the stresses remain on the yield surface during plastic flow, and is given by

$$dF = \frac{\partial F}{\partial\sigma_{ij}}\,d\sigma_{ij} + \frac{\partial F}{\partial\varepsilon^{p}_{ij}}\,d\varepsilon^{p}_{ij} + \frac{\partial F}{\partial K}\,dK = 0 \tag{36}$$

This equation admits the dependence of the yield function on the temperature and neutron flux through the parameter *K*. Making use of equations 34 through 36, and after some manipulation, the final form of the incremental stress-strain law is found to be

$$d\sigma_{ij} = D_{ijkl}\left(d\varepsilon_{kl} - \eta\frac{\partial F}{\partial\sigma_{kl}}\,\Lambda_{mn}\,d\varepsilon^{*}_{mn}\right) - d\psi_{ij} \tag{37}$$

where,

$$d\psi_{ij} = \eta\,D_{ijkl}\frac{\partial F}{\partial\sigma_{kl}}\frac{\partial F}{\partial K}\,dK\Big/\left(D_{mnpq}\frac{\partial F}{\partial\sigma_{mn}}\frac{\partial F}{\partial\sigma_{pq}} - \frac{\partial F}{\partial\varepsilon^{p}_{mn}}\frac{\partial F}{\partial\sigma_{mn}}\right) \tag{38}$$

$$d\varepsilon^{*}_{kl} = d\varepsilon_{kl} - dtB_{klmn}\,J\sigma_{mn} - dC_{klmn}\,\sigma_{pq} - d\varepsilon^{f}_{kl} \tag{39}$$

$$\Lambda_{kl} = D_{ijkl}\frac{\partial F}{\partial \sigma_{ij}}\left(D_{mnpq}\frac{\partial F}{\partial \sigma_{mn}}\frac{\partial F}{\partial \sigma_{pq}} - \frac{\partial F}{\partial \varepsilon^{p}_{mn}}\frac{\partial F}{\partial \sigma_{mn}}\right) \tag{40}$$

$$\eta = 0 \quad \text{when} \quad F < 0$$

$$\eta = 1 \quad \text{when} \quad F = 0 \quad \text{and} \quad \frac{\partial F}{\partial \sigma_{ij}}\, d\sigma_{ij} + \frac{\partial F}{\partial K}\, dK \geq 0 \tag{41}$$

Rather than continue with the cumbersome indicial notation, we switch to the more convenient matrix notation, which is highly adaptable to implementation in the finite element method, as well as for formulating pellet fracture as will be shown later. After substituting for the various derivatives in equations 37 through 40, and casting the results in matrix notation, we obtain the following final form of the incremental stress-strain matrix equations:

$$\begin{aligned}\Delta \underset{\sim}{\sigma}_n = \underset{\approx}{H}\left\{\Delta \underset{\sim}{\varepsilon}_n - \Delta \underset{\sim}{\varepsilon}^{f}_n - \left(J_n \underset{\approx}{B} + \Delta \underset{\approx}{C}\right)\underset{\sim}{\sigma}_{n-1}\right\} \\ + 2\eta \frac{K \Delta K}{Q}\left(\underset{\approx}{I} + J_n \underset{\approx}{M}\,\underset{\approx}{C}\right)^{-1}\left(\underset{\sim}{S}_n + 2\beta I_{1n}\,\underset{\sim}{\delta}\right)\end{aligned} \tag{42}$$

in which

$$\underset{\approx}{M} = \left[\underset{\approx}{I} + \underset{\approx}{D}\left(J_n \underset{\approx}{B} + \Delta \underset{\approx}{C}\right)\right]^{-1}\underset{\approx}{D} \tag{43}$$

$$\underset{\approx}{H} = \underset{\approx}{M} - (\eta/Q)\,\underset{\approx}{M}\left(\underset{\sim}{S}_n + 2\beta I_{1n}\,\underset{\sim}{\delta}\right)\left(\underset{\sim}{S}_n + 2\beta I_{1n}\,\underset{\sim}{\delta}\right)^{T}\underset{\approx}{M} \tag{44}$$

$$Q = \left(\underset{\sim}{S}_n + 2\beta I_{1n}\,\underset{\sim}{\delta}\right)^{T}\underset{\approx}{M}\left(\underset{\sim}{S}_n + 2\beta I_{1n}\,\underset{\sim}{\delta}\right) + \frac{2}{3}E'\left(\underset{\sim}{S}_n + 2\beta I_{1n}\,\underset{\sim}{\delta}\right)^{T}\left(\underset{\sim}{S}_n + 2\beta I_{1n}\,\underset{\sim}{\delta}\right) \tag{45}$$

In the above, the subscript n refers to time step t_n, $\Delta\varepsilon_n$ is the incremental strain vector, $\Delta\sigma_n$ is the incremental stress vector, $\Delta\underset{\sim}{\varepsilon}^{f}_n$ is the free expansion incremental strain vector, S is the deviatoric stress vector, $\underset{\sim}{\delta}$ is the Kronecker delta in vector form, $\underset{\approx}{D}$ is the elasticity matrix, $\underset{\approx}{I}$ is the identity matrix, $\Delta\underset{\approx}{C}$ is the change in $\underset{\approx}{C}$ due to a change in temperature and/or fluence; the other symbols have been defined. Equation 42 is the matrix equation that is modified for pellet fracture, as shown next, and is utilised in the derivations of finite element stiffness and force matrices.

10.3. Constitutive modelling of fuel pellet fracture

Cracking under normal operations is an inherent behaviour of solid fuel pellets, and as such it plays an important role in the fuel relocation and fuel-clad mechanical interaction mechanisms. It is modelled as a multi-axial direction-dependent distributed damage, (distributed over the element integration-point volume), in the constitutive relations and in the formulation of the equilibrium equations.

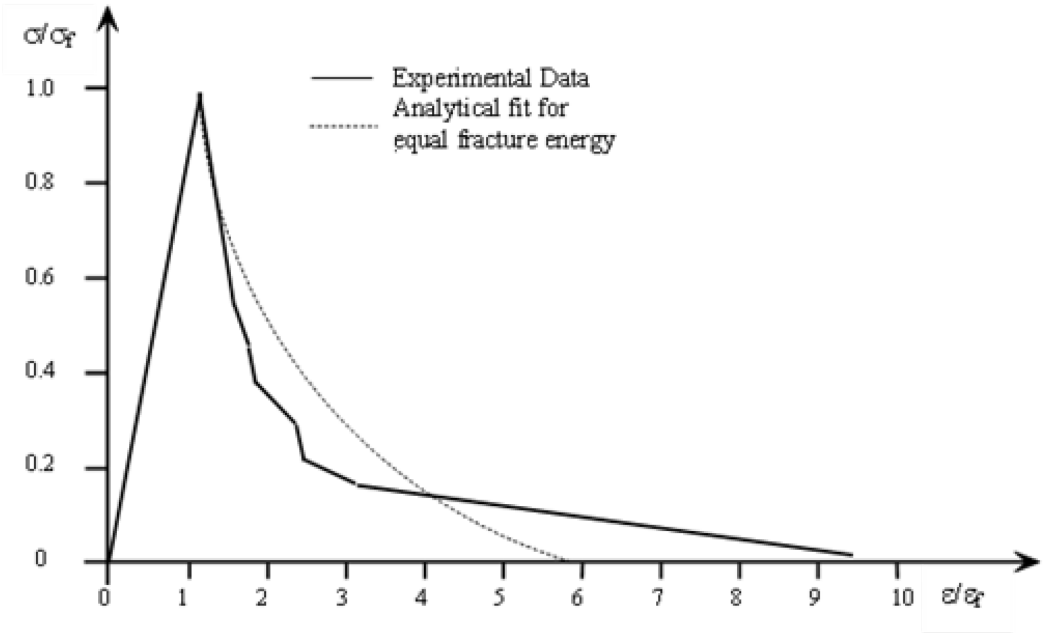
10.3.1. Uniaxial response

A typical representation of the tensile response of the material is displayed in Figure 10.3. which shows an initially linear behaviour up to a peak point followed by a strain softening regime. The peak point on this figure is the crack initiation point; however, the crack is not fully formed until the end of the strain softening regime, which can extend to a strain that is five to ten times larger than the fracture initiation strain. The modelling of this post-peak response of the material is very important to the behaviour of the fuel pellet, particularly under cyclic power histories where pellet expansion and contraction, which influences the gap conductance and consequently the thermal response, are greatly affected by that regime. Experimental stress-strain data in tension for fuel pellets is totally lacking. However, Figure 10.3. was constructed using the observed response of ceramic materials which generally exhibit the strain softening response shown in the figure. By knowing the fracture stress or strain and the elastic modulus and making an assumption on the extent of the strain softening regime, a normalised form of the stress-strain curve can be constructed as shown in Figure 10.3. It should be noted, however, that the model is not very sensitive to the shape of the strain softening curve. Using the analytical fit in Figure 10.3., the normalised stress is expressed as

$$\frac{\sigma}{\sigma_f} = A + B\, e^{-C\varepsilon/\varepsilon_f} \tag{46}$$

where, ε_f and σ_f are the fracture strain and stress respectively and A, B and C are functions of the ratio $\varepsilon_0/\varepsilon_f$, where ε_0 is the crack opening strain when the crack is assumed to be fully formed.

Figure 10.3. Normalised tension response curve for ceramic materials



Source: Rashid, 2019

For strains greater than ε_0, the material is not allowed to support any tensile stress normal to the crack plane. In the strain softening range between the peak point $\varepsilon = \varepsilon_f$ and $\varepsilon = \varepsilon_0$,

the material loses its stiffness and sheds the tensile stress gradually. This behaviour, however, is valid only for previously un-cracked material, but once the crack becomes fully formed, the material loses its tensile stiffness capacity in the normal direction to the crack regardless of the magnitude of the tensile strain. However, under strain reversals during power cycling, the crack is allowed to close and develop compressive stiffness even under residual tensile strain. This compressive stiffness recovery occurs gradually, from a very small residual value until the strain across the crack plane reaches ε_0, then gradually increasing to full value when the calculated crack strain becomes negative. This stiffness history is derived using a non-dimensional crack history parameter e(t), as will be discussed below. Equation (46) can be written as follows for $\varepsilon > \varepsilon_f$.

$$\frac{\sigma}{\sigma_f} = \frac{\sigma}{\varepsilon}\,\frac{\varepsilon}{\varepsilon_f}\,\frac{\varepsilon_f}{\sigma_f} = A + B\,e^{-C\varepsilon/\varepsilon_f} \tag{47}$$

recognising that $\sigma/\varepsilon = E$ and $\sigma_f/\varepsilon_f = E_0$, where E is the post-peak (cracking) modulus and E_0 is the initial (elastic) modulus, and introducing the definition of e(t), equation (47) becomes

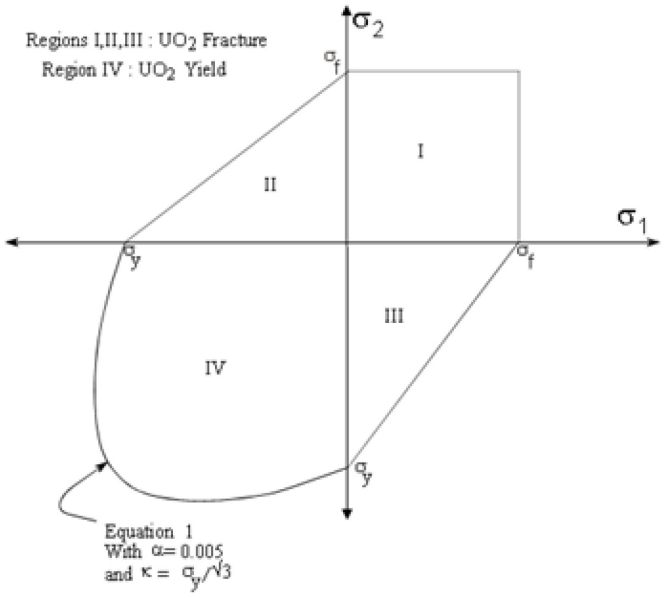
$$e(t) = E/E_0 = (\varepsilon_f/\varepsilon)\left(A + B\,e^{-C\varepsilon/\varepsilon_f}\right) \tag{48}$$

The function e(t) takes on the values of unity and zero respectively for the conditions $\varepsilon \leq \varepsilon_f$ and $\varepsilon \geq \varepsilon_0$. In the range $\varepsilon_f < \varepsilon < \varepsilon_0$, e(t) will have a value between unity and zero according to equation (48). In this sense, e(t) can be regarded as the memory function of the crack. This expression for e(t) will be used to modify the constitutive relations for the fuel pellet.

10.3.2. Multi-axial fracture

The preceding discussion dealt with the uniaxial response in tension. To generalise this one-dimensional (1D) response to two- and 3D stress states, the smeared-cracking model (Rashid, 1974) is utilised in which cracking is modelled as a direction-dependent damage distributed over the volume assigned to the element integration point. The crack is assumed to initiate at a point when the calculated principal strain reaches the material fracture strain as a temperature-dependent material property. The crack surface is the perpendicular plane to the principal direction. A failure/yield surface is illustrated in Figure 10.4., choosing, for convenience, a two-dimensional (2D) stress state. Stress states in the compression-compression quadrant of Figure 10.4. are characterised as elasto-plastic regimes, and for stress states in the positive-positive or positive-negative quadrants of the failure surface, the material responds elastically until the stress state reaches the failure surface where the material will fail in the cracking mode; the tangent to the failure surface determines the orientation of the crack surface.

Figure 10.4. 2D failure/yield surface for UO_2 fuel pellet represented in principal stress space



Source: Rashid, 2019

Consider the strain state $\bar{\underset{\sim}{\varepsilon}}_n$, expressed in the principal co-ordinate system, then the transformation of the co-ordinate strains $\underset{\sim}{\varepsilon}_n$ to $\bar{\underset{\sim}{\varepsilon}}_n$ is

$$\bar{\underset{\sim}{\varepsilon}}_n = \underset{\approx}{P}\, \underset{\sim}{\varepsilon}_n \tag{49}$$

where $\underset{\approx}{P}$ is function of the direction cosines of the principal direction unit vectors. Transforming the incremental quantities $\Delta\underset{\sim}{\varepsilon}_n$ and $\Delta\underset{\sim}{\sigma}_n$ to the principal axes gives:

$$\Delta\bar{\underset{\sim}{\varepsilon}}_n = \underset{\approx}{P}\, \Delta\underset{\sim}{\varepsilon}_n \tag{50}$$

$$\Delta\bar{\underset{\sim}{\sigma}}_n = \left(\underset{\approx}{P}^{-1}\right)^T \Delta\underset{\sim}{\sigma}_n \tag{51}$$

The transformed vectors $\Delta\bar{\underset{\sim}{\varepsilon}}_n$ and $\Delta\bar{\underset{\sim}{\sigma}}_n$ and the untransformed vectors $\Delta\underset{\sim}{\varepsilon}_n$ and $\Delta\underset{\sim}{\sigma}_n$ must, by conservation of energy, satisfy the inner product relation:

$$\Delta\underset{\sim}{\varepsilon}_n^T\, \Delta\underset{\sim}{\sigma}_n = \Delta\bar{\underset{\sim}{\varepsilon}}_n^T\, \Delta\bar{\underset{\sim}{\sigma}}_n \tag{52}$$

Re-writing equation (42) in short form

$$\Delta \underset{\sim}{\sigma}_n = \underset{\approx}{H}\, \Delta \underset{\sim}{\varepsilon}_n - \underset{\sim}{R}_n \tag{53}$$

where, $\underset{\sim}{R}_n$ is an internal stress vector which contains the rest of the terms in equation (42). Transforming equation (53) to the principal co-ordinate system yields

$$\Delta \underset{\sim}{\overline{\sigma}}_n = \underset{\approx}{\overline{H}}\, \Delta \underset{\sim}{\overline{\varepsilon}}_n - \left(\underset{\approx}{P}^{-1}\right)^T \underset{\sim}{R}_n \tag{54}$$

where,

$$\underset{\approx}{\overline{H}} = \left(\underset{\approx}{P}^{-1}\right)^T \underset{\approx}{H}\, \underset{\approx}{P}^{-1} \tag{55}$$

Now the memory function e(t) introduced in equation (48) can be generalised to describe the history of a crack in a principal plane, i.e. $e_i(t)$ describes the history of the crack whose normal is in the i^{th} principal direction, namely,

$$e_i(t) = \left(\varepsilon_f/\varepsilon_i\right)\left(A + B\, e^{-C\varepsilon/\varepsilon_f}\right) \quad , \quad i = 1, 2, 3 \tag{56}$$

Applying equation (56) to (53) leads to the following incremental stress-strain equation expressed in the principal co-ordinate system:

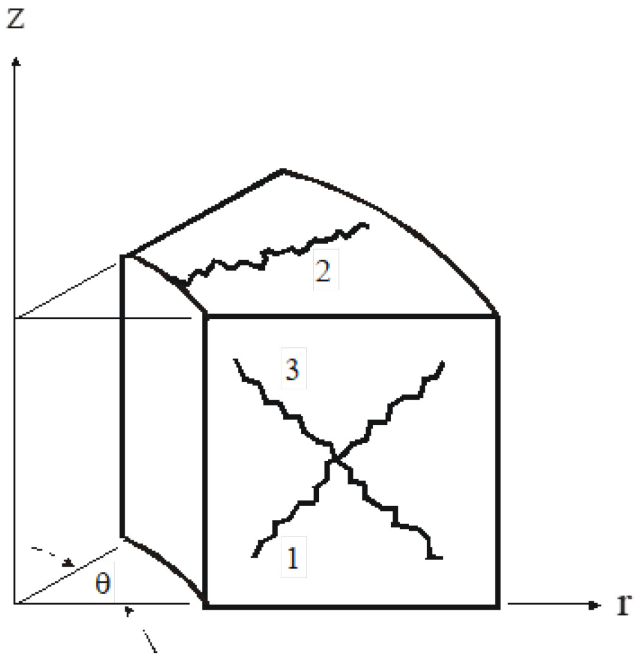
$$\Delta \underset{\sim}{\overline{\sigma}}_n = \underset{\approx}{E}^T \underset{\approx}{\overline{H}}\, \underset{\approx}{E}\, \Delta \underset{\sim}{\overline{\varepsilon}}_n - \left(\underset{\approx}{P}^{-1}\right)^T \underset{\sim}{R}_n + \left(\underset{\approx}{I} - \underset{\approx}{E}\right)\left[\left(\underset{\approx}{P}^{-1}\right)^T \underset{\sim}{R}_n - \underset{\sim}{\sigma}_{n-1}\right] \tag{57}$$

where, $\underset{\approx}{E}$ is a diagonal matrix whose diagonal elements are functions of the three possible crack histories $e_i(t)$, i = 1,2,3, defined by equation (56). Specifically, for the axisymmetric 2D case,

$$\underset{\approx}{E} = \begin{bmatrix} e_1 & 0 & 0 & 0 \\ 0 & e_2 & 0 & 0 \\ 0 & 0 & e_3 & 0 \\ 0 & 0 & 0 & e_1 e_3 \end{bmatrix} \tag{58}$$

In plane stress $\underset{\approx}{E}$ defines two possible orthogonal cracks. In an axisymmetric stress state, however, a third crack surface can develop along a radius, which implies the existence of an infinite number of such cracks. Figure 10.5. shows the three possible crack types that can develop in the fuel pellet in axisymmetric representation.

Figure 10.5. Fuel cracking pattern in axisymmetric fuel rod representation



Source: Rashid, 2019

In the present development the cracks can follow independent histories, including crack closing and possible healing if the fuel temperature rises above the sintering temperature. Finally, equation (57) is transformed back to the original co-ordinate system with the result

$$\Delta \underset{\sim}{\sigma}_n = \underset{\approx}{Z}^T \underset{\approx}{H} \underset{\approx}{Z} \Delta \underset{\sim}{\varepsilon}_n - \underset{\approx}{Z}^T \underset{\sim}{R}_n - \left(\underset{\approx}{I} - \underset{\approx}{Z}^T \right) \underset{\sim}{\sigma}_{n-1} \tag{59}$$

where,

$$\underset{\approx}{Z} = \underset{\approx}{P}^{-1} \underset{\approx}{E} \underset{\approx}{P} \tag{60}$$

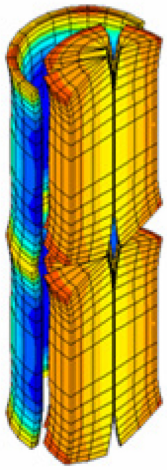
$$\underset{\sim}{R}_n = \underset{\approx}{H} \left[\Delta \underset{\sim}{\varepsilon}_n^f + \left(J_n \underset{\approx}{B} + \Delta \underset{\approx}{C} \right) \underset{\sim}{\sigma}_{n-1} \right] + 2\eta \frac{K \Delta K}{Q} \left(1 + J_n \underset{\approx}{M} \underset{\approx}{C} \right)^{-1} \left[S'_n + 2\beta + J_1 \underset{\sim}{\delta} \right] \tag{61}$$

10.4. Finite element implementation

Implementation of the above formulation in finite element analysis follows standard procedures, which are familiar to most numerical analysts, and need not be repeated here. However, the state of the art of finite element modelling of nuclear fuel is limited to 2D simulation, and the best known finite element based fuel performance code, and most functional as an industry tool, is EPRI's Falcon (Rashid, Montgomery, Dunham, 2004). Encouraging 3D finite element simulation of fuel response was developed by CEA using the 3D Toutatis code, Figure 10.6., (Brochard, Bentejac, Hourdequin, 1997). It should be

noted, however, that such 3D simulations still remain as special-effects research studies rather than a routine fuel performance evaluation tool.

Figure 10.6. 3D simulation of fuel pellets deformation in CEA's Toutatis Code



Source: Brochard, Bentejac, Hourdequin, 1997

With the exception of the Falcon code cited above, the nuclear industry's current practice in fuel performance modelling and analysis utilises finite-difference-based 1D axially stacked radial representation, referred to in the literature as 1½-D simulation. The movement towards higher dimensional simulation is discouraged by the large efforts needed for re-licensing fuel performance codes. However, efforts at developing 3D fuel behaviour codes have been initiated by the US Department of Energy at INL and ORNL within the NEAMS and CASL programmes (Larzelere, 2010), which are aimed at enabling science-based nuclear energy systems through advanced modelling and simulation, the outcome of which would be a new generation of high fidelity fuel behaviour codes.

10.5. Model limitations

The fuel's viscoplastic and fracture behaviour described above models the material as a continuum. With respect to pellet cracking, the constitutive model formulates pellet fracture in a framework that avoids computational instabilities which are inherent in a contiguous assembly of individual pellets due to rigid-body motion of the pellets and the cracking-induced pellet fragments. The above constitutive formulation of pellet fracture produces regular fracture patterns, which are clearly at variance with the observed random cracking. This limitation, however, is not particularly conducive to improvement through research since there is no known mathematical construct by which random processes become deterministic phenomena.

Another limitation of the model is that the evolution of the microstructure of the fuel material during irradiation is not represented in the model. This includes the high burnup structure and the evolution of the highly porous zone at the pellet periphery and the fission gas pressurised pores within the grains and at the grain boundaries. These porosities, unlike the uniformly distributed manufacturing porosities described above, are operationally induced randomly distributed pores that play a major role in fuel fragmentation during loss of coolant accidents. Characterising the associated fuel microstructure requires meso-scale

and even atomistic-scale modelling approaches for which atomistic theory and computational simulations can be of great value (Stan, 2009).

References

Brochard, J., F. Bentejac and N. Hourdequin (1997), "Nonlinear finite element studies of the pellet-cladding mechanical interaction in a PWR fuel," in *CW/4, Transactions of the 14th International Conference on Structural Mechanics in Reactor Technology (SMiRT 14)*, Lyon, France, August 17-22, 1997.

Hill, R. (1950), *The Mathematical Theory of Plasticity*, Clarendon Press, Oxford.

Kaufman, S.K. (1961), *The hot-pressing behavior of sintered Iow-density pellets of $U0_2$, ZrO_2-$U0_2$, ThO_2, and ThO_2-UO_2, WAPD-TM-751*, Westinghouse Electric Corporation, May, 1961.

Larzelere, A.R. (2010), *Nuclear Energy Advanced Modeling and Simulation (NEAMS)*, DOE Presentation, April 29, www.energy.gov/sites/prod/files/NEAC042910-NEAM.pdf.

Rashid, Y.R. (1974), "Mathematical Modeling and Analysis of Fuel Rods", *Nuclear Engineering and Design*, Vol. 29, pp. 22-32, November 1974.

Rashid, Y.R., H.T. Tang and E.B. Johansson (1974), "Mathematical Treatment of Hot Pressing of Reactor Fuel," *Nuclear Engineering and Design*, Vol. 29, No. 1, November 1974.

Rashid, Y.R., R.O. Montgomery and R.S. Dunham (2004), *Fuel Analysis and Licensing Code: FALCON MOD01, Volume 1: Theoretical and Numerical Bases*, EPRI Report, December 2004, -1011307.

Rashid, Y.R., S.K. Yagnik and R.O. Montgomery (2011), "LWR Fuel Performance Modeling and Multi-Dimensional Simulation – Current Status and Future Needs", *Journal of Materials (JOM)*, August, 2011.

Stan, M. (2009), "Multi-Scale Models and Simulations of Nuclear Fuels", *Nuclear Engineering and Technology*, Vol. 41, No.1, February 2009.

Conclusions

The most important deficiency in our understanding of irradiated material behaviour is microstructure evolution. The initial microstructure of fuels and structural materials is structurally and chemically complex. By necessity, approximations must be made to represent this complexity in a computational model. As the material undergoes irradiation, fission products are introduced and the chemistry changes. In addition, defects and fission products migrate under a non-uniform temperature and stress field leading to anisotropic swelling, gas release, segregation of species, nucleation and growth of new phases, and recrystallisation. At the macroscale, chemical, mechanical and thermal effects can combine to drive complex process, such as pellet-clad interaction, resulting in material failure. Modelling these interrelated processes in a reliable fashion remains an ongoing challenge.

This report describes computer simulation methods and models for studying nuclear fuels and structural materials across length and time scales. While some aspects of materials degradation under the combined effects of stress, elevated temperature and radiation can be isolated and studied at a given scale, it is essential to integrate models across scales to obtain a predictive understanding of the performance of materials.

At the most fundamental level, electronic structure calculations provide valuable information about point defect properties and interfacial energies. *Ab-initio* molecular dynamics (AIMD) enables the accurate simulation of a small collection of atoms over the nanosecond time scale. One can construct simple models of interfaces to understand segregation, interfacial transport and the early stages of phase formation. This method is especially relevant to the study of dissimilar material interfaces that are ubiquitous in nuclear materials but are challenging to model with existing interatomic potentials.

Classical molecular dynamics (MD) and accelerated dynamics methods offer the promise of extending the length and time scale of the simulation far beyond that possible with AIMD. However, the lack of reliable interatomic potentials remains a constraint. Electronic structure calculations can be used to determine properties of interest that are not available experimentally. A database of such properties can be used with machine learning to develop reliable potentials for complex materials systems. Most of the existing potentials are non-reactive and are unsuitable for modelling reactive chemical processes. At the same time, the available reactive force fields have not been rigorously tested and validated for nuclear materials. Existing potentials are parameterised for equilibrium properties but are often used to simulate systems driven far from equilibrium and outside the range of applicability of some potentials. This is especially a concern when fixed charge models are used to describe environments where charge transfer occurs. Even the widely studied Fe-Cr system does not have a completely reliable potential. It is therefore quite challenging at present to study alloys with more than two components by MD simulations. There is clearly a continuing need to develop, test and validate potentials for nuclear materials. It is also essential to continue the development of on-the-fly kinetic Monte Carlo methods that use machine learning to predict barriers instead of tabulating all possible events *a priori*. Such an approach is valuable when the local environment evolves rapidly and causes changes to the reaction rates and pathways.

There is considerable potential to connect density functional theory, MD and experimental data – e.g. transmission electron microscopy observations – with dislocation dynamics and crystal plasticity, to understand the mechanical behaviour of materials in extreme

environments. In addition, dislocation dynamics can be coupled with rate theory to provide insights into the failure of nuclear reactor materials. Information from the atomic level can be coarse-grained using the phase-field approach in order to extend the understanding to the continuum scale and explore nuclear materials phenomena across scales. The methodical integration of models will help us gradually replace empirical models in current computer codes with physics-based predictive models.

Finally, there is a pressing need for uncertainty quantification, verification and model validation in nuclear materials research. These must be performed for models at each scale. Validation requires experimental data. If such data is not available, validation experiments and lower length scale simulations must be designed to generate the needed data. In addition, the links and parameters passing between models must also be rigorously validated. Machine learning methods are being increasingly adopted in materials research. The nuclear materials community must combine the power of data analytics with physics-based models to gain new insights into materials degradation. The emergence of new computer architectures poses a challenge in relation to the reprogramming of existing codes, but it also offers the promise of reaching length and time scales well beyond what is possible at present.