Springer Series in Materials Science 226

Pierre A. Deymier
Keith Runge
Krishna Muralidharan *Editors*

Multiscale Paradigms in Integrated Computational Materials Science and Engineering

Materials Theory, Modeling, and Simulation for Predictive Design

Springer

Springer Series in Materials Science

Volume 226

The Springer Series in Materials Science covers the complete spectrum of materials physics, including fundamental principles, physical properties, materials theory and design. Recognizing the increasing importance of materials science in future device technologies, the book titles in this series reflect the state-of-the-art in understanding and controlling the structure and properties of all important classes of materials.

More information about this series at http://www.springer.com/series/856

Pierre A. Deymier · Keith Runge
Krishna Muralidharan
Editors

Multiscale Paradigms in Integrated Computational Materials Science and Engineering

Materials Theory, Modeling, and Simulation for Predictive Design

Springer

Editors
Pierre A. Deymier
Department of Materials Science and Engineering
University of Arizona
Tucson, AZ
USA

Krishna Muralidharan
Department of Materials Science and Engineering
University of Arizona
Tucson, AZ
USA

Keith Runge
Department of Materials Science and Engineering
University of Arizona
Tucson, AZ
USA

ISSN 0933-033X ISSN 2196-2812 (electronic)
Springer Series in Materials Science
ISBN 978-3-319-24527-0 ISBN 978-3-319-24529-4 (eBook)
DOI 10.1007/978-3-319-24529-4

Library of Congress Control Number: 2015950037

Springer Cham Heidelberg New York Dordrecht London

Printed on acid-free paper

Springer International Publishing AG Switzerland is part of Springer Science+Business Media (www.springer.com)

Preface

The exploration of materials and their properties has for some time evolved beyond the Edisonian paradigm, in which one sets about an experimental investigation of the synthesis and properties of materials. Advances have been experienced by a collaborative process between theory and experiment to create novel materials whose properties have been designed to serve many purposes, structural, chemical, and thermal and electrical transport, to name a few. With the advances of speed and complexity in computational resources, these collaborations have only increased in pace and scope.

Modern computing facilities allow for the modeling of large numbers of entities, whether they be the fundamental constituents of matter, electrons, and nuclei, or the collections of pieces of matter that might comprise a device, like a turbine or integrated circuit, for periods of time that could not have been imagined mere decades ago. These extraordinary capabilities have given rise to a new set of attainable goals for integrated computational materials science. For the realization of this potential, computational materials scientist must create new paradigms for the application of existing and new computational methodologies to problems that cross spatial and temporal scales.

The goal of this volume is to present the student and practicing computational materials scientist with a new framework for thinking about multiscale modeling. Integral to this framework are the concepts of theory, model, and simulation. We propose that a clear understanding of these concepts is vital for the appropriate combining of computational techniques across scales. The goal of any such framework is the faithful transmission of information across interfaces so that the modeling of physical phenomena translates the applicable theory into accurate simulations.

The chapters presented here are self-contained and intended to illustrate the state of the art in computational materials science for the modeling of the address various length and time scales. The reader can feel free to read each of these chapters according to the scales that appeal to their own interests. The final chapter explains how to create a predictive framework for communicating information across scale boundaries.

The work presented here represents not only the contributions of the authors, themselves, but also those from our mentors, collaborators, and students. Space is not sufficient here to appropriate, recognize, and thank them all for their contributions, but we hope that this short acknowledgment will begin to pay the debt we owe them.

Tucson, AZ, USA

Pierre A. Deymier
Keith Runge
Krishna Muralidharan

Contents

Contributors

Long-Qing Chen Department of Materials Science and Engineering, The Pennsylvania State University, State College, PA, USA

Pierre A. Deymier Department of Materials Science and Engineering, University of Arizona, Tucson, AZ, USA

Ibrahim Guven Virginia Commonwealth University, Richmond, VA, USA

G.E. Jabbour Department of Chemical and Materials Engineering, University of Nevada, Reno, NV, USA

Krishna Muralidharan Department of Materials Science and Engineering, University of Arizona, Tucson, AZ, USA

Ki-Dong Oh Department of Materials Science and Engineering, University of Arizona, Tucson, AZ, USA

Keith Runge Department of Materials Science and Engineering, University of Arizona, Tucson, AZ, USA

S.M. Valone Materials Science and Technology Division, Los Alamos National Laboratory, Los Alamos, NM, USA

Nan Wang Department of Materials Science and Engineering, The Pennsylvania State University, State College, PA, USA

Chapter 1
Introduction

Pierre A. Deymier and Keith Runge

Abstract In this chapter, we present the conceptual framework that motivates the development of a predictive multiscale approach to Integrated Computational Materials Science and Engineering. Clear distinction among theories, models, and simulations allow us to categorize the underlying physical principles used, the realization of those principles, and their implementation on a computational architecture. The physical theories that are explored in this volume can be combined into coherent multi-theory constructs that preserve the fidelity of each theory over its range of applicability. Similarly, models of these physical theories, which can be realized in computer simulations, must respect their range of validity over both spatial and temporal scales. The principles and procedures presented in this book promote a clear and concise methodological approach for the creation of predictive multi-theory, multi-model, multiscale simulations.

1.1 Conceptual Framework in Theory, Modeling and Simulation

Over the past few decades, a variety of computational modeling and simulation methods have provided powerful insight into the behavior of materials across the entire process-structure-properties-performance spectrum. These methods have most commonly been used to address materials phenomena at the level of a single scale belonging to the range of electronic, atomic, microstructural, mesostructural and macroscopic scales. However, many materials phenomena are inherently multiscale in both space and time and often it is making a drastic approximation to

P.A. Deymier (✉) · K. Runge
Department of Materials Science and Engineering, University of Arizona, Tucson, AZ 85721, USA
e-mail: deymier@u.arizona.edu

K. Runge
e-mail: krunge@u.arizona.edu

P.A. Deymier et al. (eds.), *Multiscale Paradigms in Integrated Computational Materials Science and Engineering*, Springer Series in Materials Science 226,
DOI 10.1007/978-3-319-24529-4_1

ignore the coupling between phenomena across multiple scales. While computational methods based on the most fundamental theoretical principles have the potential of modeling materials phenomena with a high degree of fidelity at all scales, the sheer computational cost of simulating materials phenomena at the largest spatio-temporal scales renders this approach impractical. Novel multiscale modeling and simulation paradigms are providing the means to bridge different modeling and simulation methods to achieve modeling and simulation capabilities of materials behavior from the atom all the way to the components of an engineered system. The combination of computational materials science and engineering methods with multiscale paradigms is therefore becoming of continually increasing importance for the emergence of a robust and predictive approach to materials design and analysis, namely the approach of Integrated Computational Materials Science and Engineering (ICMSE). The ICMSE approach can be understood by focusing on clear distinctions among Theory, Modeling, and Simulation, and their integration into multiscale methodologies. We address in detail below the Theory-Modeling-Simulation framework on which chapters in this book are based.

1.1.1 Theory

A *Theory* is a set of axioms and interpretative procedures that constitute a predictive mathematical description of the natural world. We will address ICMSE within the context of the theories of (a) non-relativistic quantum mechanics (QM), (b) classical mechanics, and continuum mechanics, (c) statistical mechanics and thermodynamics, (d) classical thermodynamics, etc.

Particles described within the framework of QM obey the time-independent Schrödinger equation and their behavior can be characterized through a physical interpretation of its many-particle wave function solutions. The reformulation of quantum mechanics theory in terms of electron density instead of wave functions forms the foundation of density functional theory (DFT). DFT is motivated by the Hohenberg-Kohn theorems [1] which state that the ground state of a many electron system is uniquely determined by the electron density and that this density minimizes an energy functional (a function of the spatially dependent density). For atom and molecules which include many electrons, and hence for the general many electron problem, DFT reduces the search for a three-dimensional wave function for each electron to that of a single three-dimensional function (electron density) for a given nuclear configuration or external potential.

Lagrange's formalism of classical mechanics finds the trajectory of some particle (position and velocity) in some potential as that which minimizes the action (a temporal integral of the Lagrangian, the difference between the kinetic and potential energy functions). Hamilton's approach replaces, via a Legendre transformation, position and velocities by positions and conjugated momenta as the canonical variables. The Hamilton's formalism constitutes a framework that can provide a deep physical insight into the behavior of dynamical systems.

The Hamiltonian expressing the energy of a system plays a central role in this formalism. The extension of classical mechanics from assemblies of particles to matter distributions and vector (displacement, force) and tensor (strain, stress) fields leads to the theory of continuum mechanics. For instance, linear elasticity theory addresses the deformation of solid continuum materials subjected to stress field that relate linearly to the strain field.

The formalism of the classical theory of continuum mechanics is based on partial differential equations (strain defined in terms of the gradient of displacement, mechanical equilibrium condition expressed in terms of the divergence of the stress field, etc.). The difficulty of defining partial derivatives at singularities such as discontinuities in materials (surfaces, interfaces, etc.) limits the direct application of this theory to materials exhibiting such discontinuous structural features. In contrast, the theory of peridynamics is a formulation of continuum mechanics based on force and displacement fields and integral equations of these fields. The integral equations of peridynamics are mathematically compatible with the structural discontinuities and singularities featured by most materials.

The theories of statistical mechanics and statistical thermodynamics aim at explaining the physical properties of macroscopic systems based on the knowledge of the behavior of their microscopic constituents (particles). QM and classical mechanics provide a framework for the description of systems of particles but have difficulties addressing large assemblies of interacting particles. Statistical theories enable the scale up from microscopic to macroscopic systems where only average values and fluctuations are observed. In the context of the fundamental concept of ergodicity, at equilibrium, the time average of some quantity that varies along the temporal trajectory of a dynamical system of particles is equivalent to a statistical average over a large number of macroscopically identical systems (i.e. a so-called ensemble). In that framework, one may consider the microcanonical ensemble (macroscopic systems with same energy, volume and composition), canonical ensemble (same temperature, volume and composition), and isothermal-isobaric ensemble (same temperature, pressure, composition) among others.

The theory of classical thermodynamics is based on the principle of maximization of the entropy of an isolated system. All macroscopic thermodynamics state variables are related through a fundamental equation. In the entropy representation of the fundamental equation, the equilibrium state of the system is one that maximizes the entropy with respect to the unconstrained state variables. Other representations of the fundamental equation are obtained through Legendre transformations of the entropy form, namely the Helmholtz free energy, Gibbs free energy, and enthalpy representations. The pertinent choice of these representations is made depending upon the convenience of the chosen state variables and the constraint they may be subjected to. The maximum entropy principle translates into a minimization variational principle for these other representations.

1.1.2 Model

A *Model* is a prescribed physical description of a real-world system or class of systems formulated within the concepts of a Theory.

Within the framework of DFT, Kohn and Sham [2] have reduced the problem of a many interacting electron system to the more tractable problem of a system composed of non-interacting electrons evolving in an effective potential. This effective potential includes any external potential (e.g. interactions with nuclei), the electron-electron Coulomb interactions and the exchange and correlation interactions. The latter two interactions are modeled via simplified exchange-correlation functionals such as the local density approximation (LDA) that is based on the exchange energy of the uniform electron gas. The kinetic energy is known for non-interacting particles as their wave function can be represented as a Slater determinant of orbitals. In ab initio molecular dynamics the forces acting on a set of nuclei are calculated from the electron densities determined from DFT combined with the nuclear (or ionic) repulsions.

In quantum statistical mechanics, all static properties and dynamic properties of a system of particles in thermal equilibrium are specified from the thermal density matrix. Using the closure conditions of quantum states, the partition function of this system may be expressed in terms of very high temperature thermal density matrices which can be approximated by Gaussian functions of position. This approach parallels the path integral approach to describing quantum systems. Approximate classical model Hamiltonian of systems of quantum particles results from this approach, providing an isomorphism between classical mechanics models and quantum systems. While the dynamics of the model classical system may not be realistic, statistical averages (equivalently time averages, provided the ergodic condition is satisfied) of quantities calculated from the classical Hamiltonian will be representative of the quantum system.

Classical mechanics models of interacting particles use interatomic potentials to describe the potential energy part of the classical Hamiltonian. The precision of atomistic models is dependent on the ability of the governing interatomic potential/force field to model the atomic interactions accurately, which can be achieved by parameterizing the potentials to reflect the underlying electronic effects. While equilibrium properties of materials can be successfully obtained from empirical potentials they cannot be expected to accurately model systems that are far away from equilibrium, such as when the systems are subjected to extreme conditions, where interatomic charge polarization and transfer may become significant. Thus potential formulations such as the ones based on chemical potential equalization and the recently developed charge-transfer embedded atom method (CT-EAM) and chemical environment dependent dynamic charge (EDD-Q) potentials have become extremely useful for modeling interatomic forces away from equilibrium [3]. Of particular interest are the EDD-Q formulations due to their ability to be parameterized based on quantum chemical data in a straight forward way.

A phase field model represents the non-uniformity of polycrystalline or polyphase materials via one or more ad hoc order parameters (phase fields) supplemented by other fields such as compositional, stress/strain fields, etc. An order parameter may model the grain orientation in a polycrystalline material or the different phases in a polyphase material. In this model, interfaces are not atomically sharp but modeled as diffuse continuous variations in an order parameter. The interfacial energy relates to the gradient in the order parameter. The formulation of the phase field model is then based on a free energy functional of the order parameter and other state variable fields (a fundamental equation) that may represent the thermodynamics of the system of interest [4].

The critical difference between the peridynamics approach and other approaches in classical continuum mechanics is that the former is formulated using integral equations as opposed to derivatives of the displacement components, allowing for a straight forward description of properties of materials without resorting to special boundary conditions to avoid singularities [5, 6]. Displacements in the peridynamics formalism are driven by a force which is expressed as a spatial integral of a force density. This force density is a function of the relative position between two points in the continuum and of the relative displacement between these points. Solid bodies in peridynamics are therefore modeled by the pairwise force density function for the "bond" connecting any two points. Discontinuities are easily handled by "cutting" the "bond" between two points across the discontinuity. The force density function requires a small number of materials parameters as input (e.g. elastic modulus, yield strength, critical interaction cutoff or stretch).

1.1.3 Simulation

A *Simulation* is the computational realization of a particular model. It consists of a set of algorithms and rules for reliable, refine-able (in the sense of precision) calculation of the properties of models. Improvements in the quality and reliability of computed results are separable in this framework.

Simulation methods enable the sampling of the phase space of Hamiltonian-based models, whether representing truly classical systems such as collections of particles interacting via interatomic potentials or a classical model of quantum systems (Path Integral Classical Isomorphism representation) or a quantum system. Both deterministic and stochastic simulation methods can be used to propagate the system through phase space. Molecular Dynamics (MD) is an example of a deterministic approach whereby one investigates the intrinsic dynamics of the system by integrating numerically in time the equations of motion. There exists a variety of algorithms for the states of the system to march in time. Stochastic methods such as Monte Carlo (MC) methods do not provide dynamical information but only evaluate the configurations of the system. Transitions from one configuration to another are driven by probabilistic evolution schemes such as a Markov process. This approach is therefore also applicable to model systems that may not have intrinsic dynamics.

The natural ensemble for the MD method is the microcanonical ensemble. Extension of the microcanonical MD method to simulating the dynamics of a model system in another ensemble (e.g. such as the canonical ensemble) requires the application of constraints on its degrees of freedom i.e. on its trajectories (e.g. thermostating). In that case, the dynamics of the system may be influenced by the constraint and one may consider that one only evaluates realistically the configurational part of the problem. Stochastic methods do not suffer from these constraints.

For equilibrium, variational methods are applied to the phase field free energy functional, which is minimized with respect to its degrees of freedom: order parameters and other state variable fields. For simulating time-dependent processes, one establishes dynamical equations relating the rate of change of the order parameters and state variables to the driving force (variational derivatives of the free energy functional) through mobility coefficients. Details of these equations may vary depending on the existence or not of conservation laws for the fields. Stochastic fluctuations such as thermal fluctuations may be added by introducing them in the dynamical equations stochastic terms.

The peridynamics model can be suitably discretized to model material properties, where the particle-to-particle (point-to-point in a continuum) interaction is represented via the parameterized force density function. Numerical methods to solve the classical equations of motion of each particle are used to drive the dynamics of the model solid.

1.1.4 Additional Remarks

The quality and reliability of results obtained within the *Theory-Modeling-Simulation* framework can be improved separately. For instance one can adopt a more refined theory that will include a more realistic description of the physical or chemical processes taking place in a natural system. This could be illustrated by employing a quantum mechanical description of a molecular system in place of a classical mechanics description of molecular forces. Another approach would be to adopt a more refined model within a theory. For instance, within the context of classical mechanics description of materials composed of collections of interacting atoms, the utilization of three-body interatomic potentials may enable a better prediction of mechanical properties compared to pair-wise additive potentials. Independently, within a given theory and for a fixed model, one may improve the simulation in two ways. The first one could involve the application of more accurate solvers of differential equations or the use of a smaller time step in solving numerically some equation of motion. The second may concern the adoption of a better thermostat in a canonical ensemble MD simulation.

In no way should the *Theory-Modeling-Simulation* framework allow the exploitation of offsetting errors by one or all of the three categories. What is done in the simulation must not offset limitations implicit in the choices upstream.

The *Model* must not "fix" the *Theory* by reintroducing features not in the theory. The algorithms and rules in the *Simulation* must not offset the intrinsic limits of the *Model* being treated.

Because separability of refinement is coupled with fidelity, the *Theory-Modeling-Simulation* framework explicitly supports systematic, quantitative exploration of the effects of empiricism. Convenient empirical assumptions can be replaced by more realistic ones and the consequences quantified. Moreover, the appraisal of computed data by visualization techniques is disciplined by explicit knowledge of the chain of relationships and assumptions contained in the *Model* and the *Simulation*.

1.2 Multiscale Modeling and Simulation

The Integrated Computational Materials Science and Engineering (ICMSE) paradigm captured in the *Theory*, *Modeling*, and *Simulation* triad is integrated into multiscale approaches.

1.2.1 Multiscale Approaches

To this point we have not mentioned the multiscale aspects of the *Theory*, *Modeling* and *Simulation* paradigm explicitly. For materials science, this is an issue of the desired level of fundamental understanding. Consider the simulation of a material from a quantum mechanical perspective. The limitations of computing power imposed by current computer platforms enable the simulation of systems containing at most a few hundred atoms. This means that this type of simulation cannot account for phenomena with long-range characteristics. Newtonian atomistic models simulated by MD can handle the simulation of systems that may contain billions of atoms thus permitting the investigation of phenomena at scales larger than those possible by a quantum description. However, the reliance on classical interatomic potentials makes the details of the very shortest scale of the atomic bond and related electronic phenomena inaccessible. Even with billions of atoms, very long range phenomena are inaccessible to MD simulations. Simulations of continuum models open a window on very long range scales but distance themselves from phenomena resulting from the discrete nature of matter. While, our argument above is based on the inability of some type of simulation or model to address the complete range of spatial scales that may be needed to truthfully describe a natural system, similar observations can be made concerning temporal scales. For instance, current computing power limits the time over which the dynamics of an atomistic system can be simulated. This makes long-time scale phenomena inaccessible. Again, simulations of coarser models may permit the investigation of a system over

much longer periods of time without however providing access to the detailed processes taking place at very short time scales.

The inability to handle large systems over large periods of time can be viewed as a *Simulation* limitation, not a *Model* limitation. However, as we saw, the adoption of a different *Model,* for example a MD simulation using some interatomic potential versus a quantum level simulation, might be helpful in circumventing the *Simulation* limitations. This choice will have a cost though in the inability, for instance, to handle the electronic phenomena taking place at the level of the bonds. This is a *Model* limitation. Development of multiscale models and the corresponding simulations is a direct consequence of these *Model* limitations.

In the *Theory-Modeling-Simulation* framework, a multiscale model can be defined as one in which more than one theory is incorporated, each corresponding to one of several physically important scales involved in the model.

1.2.2 Serial and Concurrent Multiscale Modeling and Simulation

Most of the modern simulation methods for coupling length scales can be characterized as either serial or concurrent. Within serial methods, a set of calculations at a fundamental level (small length scale) is used to evaluate parameters as input for a more phenomenological model that describes a system at longer length scales. For example the Quasi-continuum (QC) method is a zero temperature technique with a formulation based on standard continuum mechanics (e.g. Finite Element (FE) method) with the additional feature that the constitutive equations are drawn from calculations at the atomic scale [7–10]. In contrast, concurrent methods build around the idea of describing the physics of different regions of a material with different models and linking them via a set of boundary conditions. The archetype of concurrent methods divides the space into atomistic regions coupled with a continuum modeled via FE [11, 12]. Coarse graining has been proposed as a mean to couple seamlessly a molecular dynamics (MD) region to a FE mesh [13]. Coarse-grained MD produces equations of motion for a mean displacement field at the nodes of a coarse-grained mesh partitioning the atomistic system. Other algorithms that allow the coupling between atomistic and continuum regions have been proposed [14–17]. The fundamental issue in multiscale modeling and simulation relates to the nature of the "interface" between the models that are used to bridge their individual ranges of scales. All models must be connected as seamlessly as possible. This is the challenge of "scale-parity" simulation. Successful multiscale modeling and simulation requires treatment of all the relevant scales even-handedly. Scale-parity requires that no one scale be favored. This has led to the development of an approach to multiscale modeling called "consistent embedding" [18].

1.2.3 Consistent Embedding

The principle of consistent embedding dictates that the information that comprises a model at a larger spatial or longer time scale be compatible with the model used at a smaller or shorter scale. This principle is essential to the development of predictive theory and modeling, as the materials that exist on either side of a multiscale interface must be physically consistent. For example, if we wish to look at the phenomenon of fracture, then the stress-strain relationship for the model at the shorter length scale should, at the very least, display the same small strain behavior, Young's modulus, as the model used at the longer length scale. Enforcing this kind of constraint, that the Young's modulus of the quantum chemical and atomistic models be equivalent, up to some controllable error, is an example of consistent embedding. Other criteria can be developed, based on the physical properties being modeled, which improve the likelihood that emergent behavior of larger systems is grounded in the theoretical description of the smaller system.

The consistent embedding requirement has led to the development of techniques (e.g. based on wavelet analysis) that enable the mathematically rigorous construction of a multiscale model which range of scales is the union of the scales of the individual models that are integrated [19–25]. These techniques are applicable to serial and current methods and to bridging not only spatial but also temporal scales.

1.2.4 Multi-theory and Multi-model Approaches

We have addressed the concept of mulitscale approaches in computational simulations in the previous sections. We wish to clarify this concept in the context of our *Theory-Modeling-Simulation* framework by introducing the additional concepts of multi-theory and multi-model approaches.

The three differing concepts of multi-theory, multi-model, and multiscale approaches can occur singly or in combination. An example of a multi-theory approach would be the usual description of quantum chemical theory where the electrons are taken to behave according to quantum mechanics, while the nuclei, or ions, are taken to behave according to classical mechanics.

Another example of multi-theory approach results from the dynamical treatment of some classical mechanics system composed of particles modeled with some interatomic potential model. Using an ergodic sampling of configuration space, the system is then being studied under the rubric of statistical mechanics or thermodynamics, at the level of theory, in addition to the classical mechanics theory that was at the foundation of the model.

Following these examples a little further, we can consider a multi-theory, multi-model description of a material. If considered in the context of dynamics, the system can be comprised of (a) quantum mechanical electrons described within the

framework of DFT and modeled with some specific choice of an approximate density functional, and classical mechanical ions modeled using some interatomic potential. This system falls into the combined theories of QM, Classical Mechanics and Statistical Thermodynamics as well as two distinct models.

A single theory but multi-model example would consist of the description of a system of different particles (say electrons and ions) both in the framework of the theory of QM but the electrons being modeled with DFT and the ions modeled with an approximate classical Hamiltonian of systems of quantum particles derived from the path integral approach.

It is also possible to consider multiscale simulation versions of multi-theory, multi-model approaches. An example might be the use of atomistic and continuum mechanical theories to describe the linear mechanical response of a material (i.e. elastic behavior) in the context of a dynamical simulation. Some portion of the material may be modeled as a combination of classical ions and quantum electrons as illustrated in previous examples while another portion of the same system might be modeled as a set of mesh points using continuum mechanics. The theories used here include quantum mechanics, classical atomic and continuum classical mechanics as well as statistical mechanics. This simulation would then be based on multiple theories captured by multiple models working at distinctive spatio-temporal scales such as a density functional model of the electrons, a classical Hamiltonian model of the ions and linear elastic model of the continuum. The quantum model may be solved statically within the Born-Oppenheimer approximation (electrons always in their ground state relative to atomic motions), while the dynamics of the atomistic region is simulated with MD schemes, the elastic medium is dynamically simulated with a finite difference scheme in the time domain. This type of multi-theory, multi-model and multiscale simulations will be addressed in details in Chap. 6.

Acknowledgments The authors would like to thank Dr. Samuel B. Trickey for many helpful and informative conversations that form the basis for this chapter.

References

1. P. Hohenberg, W. Kohn, Phys. Rev. **136**, B864 (1964)
2. W. Kohn, L.J. Sham, Phys. Rev. **140**, A1133 (1965)
3. K. Muralidharan, C. Cao, Y. Wan, K. Runge, H.-P. Cheng, Chem. Phys. Lett. **437**, 92 (2007)
4. W.J. Boettinger, J.A. Warren, C. Beckermann, A. Karma, Annu. Rev. Mater. Res. **32**, 163 (2002)
5. S.A. Silling, J. Mech. Phys. Solids **48**, 175 (2000)
6. S.A. Silling, M. Epton, O. Weckner, J. Xu, A. Askari, J. Elast. **88**, 151 (2007)
7. E.B. Tadmor, M. Ortiz, R. Phillips, Phil. Mag. A **73**, 1529 (1996)
8. E.B. Tadmor, R. Phillips, M. Ortiz, Langmuir **12**, 4529 (1996)
9. V.B. Shenoy, R. Miller, E.B. Tadmor, R. Phillips, M. Ortiz, Phys. Rev. Lett. **80**, 742 (1998)
10. G.S. Smith, E.B. Tadmor, E. Kaxiras, Phys. Rev. Lett. **84**, 1260 (2000)
11. M. Mullins, M.A. Dokainish, Phil. Mag. A **46**, 771 (1982)

12. H. Kitagawa, A. Nakatami, Y. Sibutani, Mat. Sci. Eng. A **176**, 263 (1994)
13. R.E. Rudd, J.Q. Broughton, Phys. Rev. B **58**, R5893 (1998)
14. J.Q. Broughton, F.F. Abraham, N. Bernstein, E. Kaxiras, Phys. Rev. B **60**, 2391 (1999)
15. S. Kohlhoff, P. Gumbsch, H.F. Fischmeister, Phil. Mag. A **64**, 851 (1991)
16. F.F. Abraham, J.Q. Broughton, N. Bernstein, E. Kaxiras, Comp. Phys. **12**, 538 (1998)
17. S. Ogata, E. Lidorikis, F. Shimojo, A. Nakano, P. Vashista, R.K. Kalia, Comp. Phys. Commun. **138**, 143 (2001)
18. K. Runge, in *Optical Measurements, Modeling, and Metrology,* vol. 5, ed. by T. Proulx. Conference Proceedings of the Society for Experimental Mechanics Series, p. 83 (2011)
19. G. Frantziskonis, P.A. Deymier, Model. Simul. Mater. Sci. Eng. **8**, 649 (2000)
20. G. Frantziskonis, P. Deymier, Phys. Rev. B **68**, 024105 (2003)
21. G. Frantziskonis, S.K. Mishra, S. Pannala, S. Simunovic, C.S. Daw, P. Nukala, R.O. Fox, P.A. Deymier, Int. J. Mult. Comp. Eng. **4**, 755 (2006)
22. S.K. Mishra, K. Muralidharan, S. Pannala, S. Simunovic, C.S. Daw, P. Nukala, R.O. Fox, P.A. Deymier, G. Frantziskonis, Intl. J. Chem. React. Eng. **6**, A28 (2008)
23. K. Muralidharan, S.K. Misra, G. Frantziskonis, P.A. Deymier, P. Nukala, S. Pannala, S. Simunovic, Phys. Rev. E **77**, 026714 (2008)
24. G. Frantziskonis, K. Muralidharan, P. Deymier, S. Simunovic, P. Nukala, S. Pannala, J. Comp. Phys. **228**, 8085 (2009)
25. P.A. Deymier, Ki-Dong Oh, Krishna Muralidharan, G. Frantzikonis, K. Runge, J. Comput. Aided Mater. Des. **13**, 17 (2006)

Chapter 2
Path Integral Molecular Dynamics Methods

Pierre A. Deymier, Keith Runge, Ki-Dong Oh and G.E. Jabbour

Abstract We present the aspects of the path integral molecular dynamics (PIMD) method relative to (a) its theoretical fundamental principles, (b) its applicability to model quantum systems and (c) its implementation as a simulation tool. The PIMD method is based on the discretized path integral representation of quantum mechanics. In this representation, a quantum particle is isomorphic to a closed polymer chain. The problem of the indistinguishability of quantum particles is tackled with a non-local exchange potential. When the exact density matrix of the quantum particles is used, the exchange potential is exact. We use a high temperature approximation to the density matrix leading to an approximate form of the exchange potential. This quantum molecular dynamics method allows the simulation of collections of quantum particles at finite temperature. Our algorithm can be made to scale linearly with the number of quantum states on which the density matrix is projected. Therefore, it can be optimized to run efficiently on parallel computers. We apply the PIMD method to the electron plasma in 3-dimension. The kinetic and potential energies are calculated and compared with results for similar systems simulated with a variational Monte Carlo method. Both results show good agreements with each other at all the densities studied. The method is then use to model the thermodynamic behavior of a simple alkali metal. In these simulations, ions and valence electrons are treated as classical and quantum particles, respectively. The simple metal undergoes a phase transformation upon heating.

P.A. Deymier (✉) · K. Runge · K.-D. Oh
Department of Materials Science and Engineering, University of Arizona,
Tucson, AZ 85721, USA
e-mail: deymier@u.arizona.edu

K. Runge
e-mail: krunge@u.arizona.edu

K.-D. Oh
e-mail: kidong@u.arizona.edu

G.E. Jabbour
Department of Chemical and Materials Engineering, University of Nevada,
Reno, NV 89557, USA
e-mail: ghassan.jabbour@unr.edu

P.A. Deymier et al. (eds.), *Multiscale Paradigms in Integrated Computational Materials Science and Engineering*, Springer Series in Materials Science 226,
DOI 10.1007/978-3-319-24529-4_2

Furthermore, to demonstrate the richness of behaviors that can be studied with the PIMD method, we also report on the metal to insulator transition in a hydrogenoid lattice. Finally, in previous studies of alkali metals, electrons interacted with ions via local pseudo-potentials, an extension of the method to modeling electrons in non-local pseudo-potentials is also presented with applications.

2.1 Introduction

Modeling and simulation have become a vital part of materials research. Modeling and simulation techniques are maturing to the point where they offer hope for a practical and reliable approach for the study of real materials. The development of materials models has evolved from the infancy of specific empirical descriptions, to highly accurate and sophisticated representations based on first principle calculations. In the field of ab initio molecular dynamics method, the method of Car and Parrinello [1, 2], based on the Density Functional Theory (DFT) has enjoyed a great popularity over recent years. DFT molecular dynamics has been employed to investigate a very large number of problems from condensed matter to chemistry to biology [3]. In contrast, early applications of molecular dynamics simulations using the discretized path-integral [4] representation of quantum particles have been limited mostly to the simulation of systems containing a small number of quantum degree of freedom (such as in the solvation of a single quantum particle in a classical fluid [5]) or to problems where quantum exchange is not dominant [6]. We should also mention the path-integral based method of Alavi and Frenkel that allows for the calculation of the grand canonical partition function of fermion systems [7]. With this method the fermion sign problem in the evaluation of the partition function is solved exactly in the case of non-interacting fermions. When combined with DFT, this method provides a means of doing ab initio molecular dynamics of systems with interacting high temperature electrons [8].

The modeling and simulation of fermion systems by path-integral Monte Carlo method [9–14] have opened the way toward the implementation of a path integral based finite temperature ab initio molecular dynamics method (PIMD). In this chapter, we describe such a molecular dynamics method applicable to the simulation of many-fermion systems at finite temperatures. The method is based on (a) the discretized path integral representation of quantum particles as closed polymeric chains of classical particles (or beads) coupled through harmonic springs [4], (b) the treatment of quantum exchange as crosslinking of the chains [15], (c) the non-locality of crosslinking (exchange) along the imaginary time chains [9–11], and (d) the restricted path integral [12, 16, 17] to resolve the problem of negative weights to the partition function resulting from the crosslinking of even numbers of quantum particles.

The PIMD presented here is applied to the description of one-component plasma, which consists of the electron gas with a uniform neutralizing background, at the border of the degenerate and semi-degenerate regimes where the ratio of the

temperature to the Fermi temperature (T_F) $\sim$ 0.1. The electron plasma is naturally the first focus of the application of the PIMD method because it is the simplest many-body fermion system. It has been extensively studied via path-integral, variational, and diffusion Monte Carlo methods since the calculation of the equation of states of a Fermi one-component plasma. The behavior of the interacting electron gas is a problem of fundamental practical importance as one uses its properties in the density functional theory. The one-component plasma is also a good prototype system as there exists a large amount of theoretical and numerical data concerning its equation of state. The zero-temperature perturbative expansion of the energy of a three-dimensional uniform electron plasma in the high density limit, where $r_s \ll 1$ ($r_s = r/a_o$ where r is the electron sphere radius and a_o is the Bohr radius; the electron sphere radius is defined by $r = \left(\frac{4}{3}\pi\frac{N_{el}}{V}\right)^{-1/3}$ where N_{el}/V is the electron density), was calculated theoretically quite some time ago [18]. Accurate variational Monte Carlo calculations have extended the zero-temperature equation of states of the degenerate Fermi one-component plasma to a wide range of lower densities from $r_s = 1$ to 500 [19, 20]. The exchange-correlation free energy has been subsequently calculated to encompass the full range of thermal degeneracy [21–23]. The first PIMD calculations of the properties of the electron plasma at finite temperature and metallic density have been reported in [24, 25].

The application of the PIMD method to alkali metals is a natural extension of the work on the electron plasma. The PIMD method was applied to the study of a simple alkali metal, namely potassium K [26–28]. An alkali metal was chosen because (1) it is a prototype free-electron metal, (2) there exist experimental data for the pair correlation function of a liquid potassium [29] and the power spectrum of crystalline potassium at 9 K is known [30], (3) DFT molecular dynamics has had problems with metals when electrons leave the Born Oppenheimer surface and therefore violate one of the basic assumption of the method. Note that this problem had been approached technically in an ad hoc manner with the introduction of appropriate thermostats for the electronic and ionic degrees of freedom [31]. In the simple metal case, the discretized restricted path integral representation of electron is the same as that of the electron plasma. Classical ionic degrees of freedom representing potassium ions are added to the model. The electron degrees of freedom interact with the ionic degrees of freedom via empty core pseudo-potentials. It was shown that the PIMD successfully modeled the body centered crystal structure of the solid state of potassium at low temperature. The PIMD method had also the ability of simulating the melting transition of potassium. In particular, the PIMD method enables the study of the response of the electronic structure to the loss of long range order in the atomic structure in the form of localization.

Electronic transitions such as the Metal to Insulator (MI) transition are challenging phenomena for computational methods. The MI transition (MIT) is important for several materials such as transition metal oxides [32, 33]. The correlation-induced MI transition, known as Mott transition, is of fundamental importance in condensed matter theory [34]. The transition results from a competition between the electrons' potential energy that tends toward localization and

kinetic energy that favors delocalization. This competition is captured in a canonical lattice model for correlated electrons, namely the half-filled Hubbard model [35]. The PIMD method has been utilized to characterize the transition of the correlated electronic structure of a half-filled expanded three-dimensional hydrogenoid body-centered cubic lattice at finite temperatures [36]. Starting from a paramagnetic metallic state with electron gas character, it was found that bound electrons formed upon expansion of the lattice. The bound electrons were spatially localized with their center for the motion of gyration located on ionic positions. The region of coexistence of bound and unbound states in the temperature-density plane was reminiscent of that associated with a first-order transition. At constant temperature, the number of bound electrons increased monotonously with decreasing density and the width of the region of coexistence narrowed with increasing temperature.

To complete this introduction concerning the development and applications of the PIMD method, one should also mention the extension of the path integral formalism to account for the non-locality of pseudo-potentials [37].

2.2 Theoretical Framework and Model Development

In this section, we establish the theoretical framework for the PIMD method. This theoretical framework leads to the development of classical-like models of quantum particles.

2.2.1 Feynman Path Integral

2.2.1.1 Partition Function for a Single Particle

Since Feynman [38] introduced the path integral of a quantum system, it has been well developed [39–41] and applied to many-body systems [5, 12, 15, 24, 25, 42]. In quantum statistical mechanics, all static properties and dynamic properties of a system in thermal equilibrium are specified from the thermal density matrix. If we work in the canonical ensemble, which is a system of fixed number of particles in a fixed volume in equilibrium with a thermal reservoir, the probability of observing a state with energy E is proportional to e^{-E/k_BT} where k_B is Boltzmann's constant and T is the temperature. Let us consider a single particle system governed by the Hamiltonian operator $\widehat{H}$. The partition function of this system may be written

$$\begin{aligned} Z &= Tre^{-\beta\widehat{H}} \\ &= \int dr \left\langle r|e^{-\beta\widehat{H}}|r \right\rangle \end{aligned} \tag{2.1}$$

where β is $1/k_BT$ and $|r\rangle$ is the Eigenstate of $\widehat{H}$. In the path integral formalism for many-body systems, we normally represent the matrix element in the partition function as a density matrix, $\rho(r_i, r_j; \beta)$. The density matrix is defined as

$$\rho(r_i, r_j; \beta) = \langle r_i | e^{-\beta \widehat{H}} | r_j \rangle \tag{2.2}$$

Then the partition function can be rewritten as the trace of the density matrix

$$Z = \int dr \rho(r, r; \beta) \tag{2.3}$$

Before proceeding with the development of approximate forms for the density matrix, we will first consider the matrix element in real time, t, for physical clarity. By substituting $i(t_f - t_i)/\hbar$ for β, the density matrix $\rho(r_i, r_f; \beta)$ for a particle governed by the Hamiltonian, $\widehat{H}$, becomes

$$K(r_f, t_f; r_i, t_i) = \left\langle r_f | e^{-\frac{i}{\hbar}\widehat{H}(t_f - t_i)} | r_i \right\rangle \tag{2.4}$$

where t_i and t_f are an initial time and a final time, respectively. The matrix element in real time, $K(r_f,\ t_f; r_i, ; t_i)$, which is the so called Kernel [4], is a solution of the real time dependent Schrödinger Equation,

$$\frac{\hbar}{i}\frac{\partial K(f; i)}{\partial t_f} = \widehat{H}_f K(f; i) \quad t_f > t_i \tag{2.5}$$

wherein the Hamiltonian $\widehat{H}_f$ operates on the variables r_f and t_f only. By analogy, the density matrix in imaginary time is a solution of an equation of the form

$$-\frac{\partial \rho}{\partial \beta} = \widehat{H}\rho \tag{2.6}$$

Equation (2.6) is a diffusion like equation. This fact will become important when we introduce the restricted path integral and in particular, when we consider boundary conditions on the density matrix. The kernel or propagator $K(r_f, t_f; r_i, t_i)$ obeys the superposition principle, since it is an exact solution of the Schrödinger equation in real time. By the superposition principle, we mean that

$$K(r_f, t_f; r_i, t_i) = \int dr K(r_f, t_f; r, t) K(r, t; r_i, t_i) \tag{2.7}$$

at any time t, where $t_i < t < t_f$. The basic idea is to break a finite time interval into infinitesimal time steps and then evaluate the matrix element of the propagation operator for each step. Equation (2.7) indicates that one may calculate the matrix

element to any desired degree of accuracy for infinitesimal time interval, although the matrix element, or the kernel $K(r_f, t_f; r_i, t_i)$, cannot be calculated exactly for a finite time interval, $t_f - t_t$ In other words, we calculate a matrix element only for each infinitesimal time interval after breaking a finite time into infinitesimal intervals. The value for the finite time interval can be obtained from the results of the evaluations for all infinitesimal time intervals. Similarly, we may evaluate the thermal density matrix $\rho(r_i, r_j; \beta)$ with appropriate accuracy if we divide a finite temperature term ß in P infinitesimal intervals, where $P \to \infty$.

To evaluate the thermal density matrix, we will consider a set of P different configurational states, $\{|r_i>; i = 1, P\}$, where each state is an Eigenstate of $\widehat{H}$.

With the relation

$$e^{-\beta\widehat{H}} = \left(e^{-\frac{\beta\widehat{H}}{P}}\right)^P,$$

or

$$e^{-\beta\widehat{H}} = e^{-\epsilon\widehat{H}} e^{-\epsilon\widehat{H}} \ldots,$$

where $\epsilon = \beta/P$, the partition function becomes

$$Z = \int dr \left\langle r|e^{-\epsilon\widehat{H}} e^{-\epsilon\widehat{H}} \cdots e^{-\epsilon\widehat{H}}|r\right\rangle.$$

Using the closure relation of the Eigenstates $|r_i\rangle$ of $\widehat{H}$,

$$\int dr_i |r_i\rangle\langle r_i| = 1$$

and projecting the particle on (P-1) intermediate states, the partition function can be written as

$$Z = \int dr_1 dr_2 \ldots dr_{\mathrm{P}} \left\langle r_1|e^{-\epsilon\widehat{H}}|r_2\right\rangle \left\langle r_2|e^{-\epsilon\widehat{H}}|r_3\right\rangle \ldots \left\langle r_{\mathrm{P}}|e^{-\epsilon\widehat{H}}|r_{\mathrm{P}+1}\right\rangle \tag{2.8}$$

or

$$Z = \int dr_1 dr_2 \ldots dr_{\mathrm{P}}\, \rho(r_1, r_2; \epsilon)\rho(r_2, r_3; \epsilon) \ldots \rho(r_{\mathrm{P}}, r_{\mathrm{P}+1}; \epsilon) \tag{2.9}$$

where $r_1 = r_{P+1}$ indicates complete closure. Figure 2.1 illustrates a complete necklace of a single particle for P = 4.

Each matrix element or each density matrix in the above relation represents a propagator from one state to another state for infinitesimal imaginary time. In other

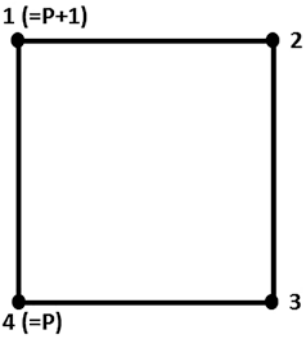

Fig. 2.1 Necklace representation of a single particle for P = 4. The *number labels* indicates the different imaginary times

words, the density matrix of a single particle is connected to look like a polymeric necklace consisting of P beads. To evaluate the density matrix of a single particle for infinitesimal imaginary time, we assume PT is very large. Then we can adopt Trotter's first order approximation [43]. According to the Trotter formula [43–45], we have

$$e^{-\beta\widehat{H}} = e^{-\beta\left(\widehat{T}+\widehat{V}\right)} = \lim_{\mathrm{P}\to\infty}\left(e^{-\frac{\beta\widehat{T}}{P}}e^{-\frac{\beta\widehat{V}}{P}}\right)^{P} \tag{2.10}$$

and

$$e^{-\epsilon\left(\widehat{T}+\widehat{V}\right)} \approx e^{-\epsilon\widehat{T}}e^{-\epsilon\widehat{V}}, \quad \textit{for small}\,\epsilon \tag{2.11}$$

where the Hamiltonian $\widehat{H} = \widehat{T} + \widehat{V}$. $\widehat{T}$ and $\widehat{V}$ are a kinetic energy operator and a potential energy operator, respectively. With (2.11), we can approximate the exact density matrix by the product of the density matrix for $\widehat{T}$ and the density matrix for $\widehat{V}$. The error of this approximation is of order of ϵ^2.

Now we are going to evaluate the density matrix for infinitesimal imaginary time step $\epsilon = \beta/P$. Let us assume that the Hamiltonian $\widehat{H} = \widehat{T} + \widehat{V}$, where $\widehat{T} = -\widehat{p}^2/2m$ and $\widehat{V}$ is a local potential energy operator. Introducing a complete set of momentum states, $|p_n\rangle$ and using (2.11), the density matrix becomes

$$\begin{aligned}\rho(r_n, r_{n+1};\epsilon) &= \left\langle r_n|e^{-\epsilon\widehat{H}}|r_{n+1}\right\rangle \\ &= \int dp_n\langle r_n|p_n\rangle\left\langle p_n|e^{-\epsilon\widehat{H}}|r_{n+1}\right\rangle \\ &= \int dp_n\langle r_n|p_n\rangle\left\langle p_n|e^{-\epsilon\widehat{p^2}/2m}e^{-\epsilon\widehat{V}}|r_{n+1}\right\rangle + \mathcal{O}(\epsilon^{\in})\end{aligned} \tag{2.12}$$

After applying the momentum operator to $|p_n\rangle$ and the local potential operator to $|r_n + 1\rangle$, we then obtain

$$\rho(r_n, r_{n+1}; \epsilon) \approx \rho_0(r_n, r_{n+1}; \epsilon) e^{-\epsilon V(r_{n+1})} \tag{2.13}$$

In (2.13) we define $\rho_0(r_n, r_{n+1}; \epsilon)$ as the density matrix of a free particle (also called the free particle propagator), as

$$\rho_0(r_n, r_{n+1}; \epsilon) = \left\langle r_n | e^{-\epsilon \widehat{T}} | r_{n+1} \right\rangle \tag{2.14}$$

or

$$\rho_0(r_n, r_{n+1}; \epsilon) = \int dp_n \langle r_n | p_n \rangle \langle p_n | r_{n+1} \rangle e^{-\epsilon p_n^2 / 2m} \tag{2.15}$$

Using the Gaussian integral, the free particle propagator becomes

$$\rho_0(r_n, r_{n+1}; \epsilon) = \left(\frac{m}{2\pi\epsilon\hbar^2} \right)^{\frac{3}{2}} \exp\left(-\frac{m(r_n - r_{n+1})^2}{2\epsilon\hbar^2} \right) \tag{2.16}$$

We give more details of the derivation of (2.16) in the Appendix 1. From (2.9) with $\epsilon = \beta/P$, the partition function of a single particle can be written as

$$\begin{aligned} Z &= \int \prod_{n=1}^{\mathrm{P}} * dr_n \rho\left(r_n, r_{n+1}; \frac{\beta}{P} \right) \\ &\approx \int \prod_{n=1}^{P} * dr_n \rho_0\left(r_n, r_{n+1}; \frac{\beta}{P} \right) e^{-\epsilon V(r_{n+1})} \end{aligned} \tag{2.17}$$

or

$$Z \approx \left(\frac{m}{2\pi\epsilon\hbar^2} \right)^{3P/2} \int \prod\nolimits_{n=1}^{P} * dr_n e^{-\beta V_{eff}} \tag{2.18}$$

$$V_{eff}(r_1, r_2, \ldots, r_P) = \sum\nolimits_{n=1}^{P} * \frac{Pm}{2\hbar^2\beta^2} (r_n - r_{n+1})^2 + \frac{1}{P} V(r_{n+1}) \tag{2.19}$$

The (*) on the product and the summation indicate that $r_{P+1} = r_1$. The partition function of (2.18) is similar to a classical partition function. The first term of the effective potential, V_{eff}, originates from the kinetic energy of a particle. We may interpret this term as a harmonic type interaction between the nearest neighbor beads in a closed necklace. The coupling constant is $C_1 = Pm/\hbar^2\beta^2$. Equation (2.19), therefore, established a classical model of a quantum particle.

In the high temperature limit, the necklace of P beads collapses to a single point so that a quantum particle becomes a classical particle. The classical partition function is valid in the limit $\epsilon = \beta/P \to 0$. The classical isomorphism is therefore

more accurate at high temperature, T, and for a large number of P. At low temperature, the quantum particle possesses some spatial extent associated with its de Broglie wavelength.

2.2.1.2 Systems of Interacting Particles Obeying Maxwell-Boltzmann Statistics

In the previous section, we discussed the thermal density matrix and the partition function of a single particle in a canonical ensemble. Here, we will generalize them to systems containing many quantum particles. We will attach particular attention to the contrast between the discretized path integral form of the partition function of particles obeying Maxwell-Boltzmann, Bose-Einstein, and Fermi-Dirac statistics.

The partition function for a N-body system may be written as

$$Z = \int dr_1 \ldots dr_N \, \rho(r_1, \ldots, r_N; \beta) \tag{2.20}$$

where the thermal density matrix of N distinguishable particles, that are obeying Maxwell Boltzmann statistics, is defined by

$$\rho(r_1, \ldots, r_N; \beta) = \left\langle r_1 \ldots r_N | e^{-\beta \widehat{H}} | r_1 \ldots r_N \right\rangle \tag{2.21}$$

with $|r_1 \cdots r_N\rangle$ being an Eigenstate of $\widehat{H}$. We will assume that the Hamiltonian of the N-particle system takes the form

$$\widehat{H} = \sum\nolimits_{i=1}^{N} \frac{\widehat{p}_i^2}{2m} \sum\nolimits_{i=1}^{N} \widehat{\emptyset}(r_i) + \frac{1}{2} \sum\nolimits_{i>j}^{N} \widehat{v}_{ij} \tag{2.22}$$

with $\widehat{v}_{ij} = \widehat{v}\left(r_i - r_j\right)$. $\widehat{\emptyset}(r_i)$ is an external potential acting on the ith particle and $\widehat{v}_{ij}$ is a pair potential describing the interactions between particles i and j. In order to evaluate the partition function, let us discretize the density matrix by inserting (P-1) intermediate states for each particle.

With the property of completeness of the Eigenstates $|r_1 \cdots r_N\rangle$ of $\widehat{H}$, the density matrix becomes

$$\begin{aligned} \left\langle r_1 \ldots r_N | e^{-\beta \widehat{H}} | r_1 \ldots r_N \right\rangle &= \left\langle r_1 \ldots r_N | e^{-\epsilon \widehat{H}} e^{-\epsilon \widehat{H}} \ldots e^{-\epsilon \widehat{H}} | r_1 \ldots r_N \right\rangle \\ &= \int \prod_{v=1}^{P} \prod_{n=1}^{N} dr_n^{(v)} \prod_{k=1}^{P} * \left\langle r_1^{(k)} \ldots r_N^{(k)} | e^{-\epsilon \widehat{H}} | r_1^{(k+1)} \ldots r_N^{(k+1)} \right\rangle \end{aligned} \tag{2.23}$$

In (2.23), the subscript and superscript of $r_i^{(k)}$ denote the kth element (or so-called *bead*) of the necklace of the ith electron. For the sake of convenience, we will use a new notation, $|R\rangle = |r_1 \cdots r_N\rangle$ and $|R^{(k)}\rangle = |r_1^{(k)} r_2^{(k)} \cdots r_N^{(k)}\rangle$. Using the Trotter approximation and applying the potential operator, we may write the infinitesimal density matrix as

$$\begin{aligned}
&\left\langle R^{(k)} \middle| e^{-\epsilon \widehat{\mathrm{H}}} \middle| R^{(k+1)} \right\rangle \approx \left\langle R^{(k)} \middle| e^{-\epsilon \widehat{T}} e^{-\epsilon \widehat{V}} \middle| R^{(k+1)} \right\rangle \\
&= \left\langle R^{(k)} \middle| e^{-\epsilon \sum_{i=1}^{N} \frac{\hat{p}_i^2}{2m}} e^{-\epsilon \left(\sum_{i=1}^{N} \widehat{\emptyset}(r_i) + \sum_{i=1}^{N} \widehat{v}(r_i - r_j) \right)} \middle| R^{(k+1)} \right\rangle \\
&= \left\langle R^{(k)} \middle| e^{-\epsilon \sum_{i=1}^{N} \frac{\hat{p}_i^2}{2m}} \middle| R^{(k+1)} \right\rangle \exp\left\{ -\epsilon \left(\sum_{i=1}^{N} \emptyset(r_i^{(k)}) + \sum_{i>j}^{N} v\left(r_i^{(k)} - r_i^{(k)} \right) \right) \right\}
\end{aligned} \tag{2.24}$$

The term $\left\langle R^{(k)} | e^{-\epsilon \sum \widehat{p_i^2}/2m} | R^{(k+1)} \right\rangle$ in (2.24) is a free particle propagator of N distinguishable particles. If we use the expression for the free particle propagator for a single particle, (2.16), we have

$$\begin{aligned}
\rho_0\left(R^{(k)}, R^{(k+1)}; \epsilon \right) &\equiv \left\langle R^{(k)} | e^{-\epsilon \sum_{i=1}^{N} \frac{\widehat{p_i^2}}{2m}} | R^{(k+1)} \right\rangle \\
&= \int \prod_{n=1}^{N} dp_n^{(k)} \left\langle R^{(k)} | p_1^{(k)} \ldots p_N^{(k)} \right\rangle \left\langle p_1^{(k)} \ldots p_N^{(k)} | e^{-\epsilon \sum_{i=1}^{N} \frac{\widehat{p_i^2}}{2m}} | R^{(k+1)} \right\rangle \\
&= \int \prod_{n=1}^{N} dp_n^{(k)} \left\langle R^{(k)} | p_1^{(k)} \ldots p_N^{(K)} \right\rangle \left\langle p_1^{(k)} \ldots p_N^{(k)} | R^{(k+1)} \right\rangle e^{-\epsilon \sum_{i=1}^{N} \frac{p_i^{(k)2}}{2m}} \\
&= \left(\frac{m}{2\pi\epsilon\hbar^2} \right)^{3N/2} \exp\left(- \sum_{i=1}^{N} \frac{m \left(r_i^{(K)} - r_i^{(k+1)} \right)^2}{2\epsilon\hbar^2} \right)
\end{aligned} \tag{2.25}$$

If we repeat the evaluation of an infinitesimal density matrix over all intermediate states, the partition function of the N-body system can be written as

$$Z \approx \left(\frac{C_1}{2} \right)^{3NP/2} \int \prod_{i=1}^{P} \prod_{k=1}^{N} dr_i^{(k)} e^{-\beta(V_1 + V_2)} \tag{2.26}$$

where

$$V_1 = \frac{C_1}{2}\sum_{k=1}^{P} * \sum_{i=1}^{N}\left(r_i^{(k)} - r_i^{(k+1)}\right)^2 \tag{2.27}$$

and

$$V_2 = \frac{1}{P}\sum_{k=1}^{P}\sum_{i=1}^{N} \emptyset\left(r_i^{(k)}\right) + \frac{1}{P}\sum_{k=1}^{P}\sum_{i>j}^{N} v\left(r_i^{(k)} - r_j^{(k)}\right) \tag{2.28}$$

The effective potential, $V_1 + V_2$, of the N-body system in absence of quantum exchange is similar to the effective potential of one-quantum particle system. V_1 represents the harmonic potential which corresponds to interactions between the first neighbors in the closed necklaces. V_2 is nothing but the potential energy resulting from the external field and the particle/particle interactions. Equations (2.27) and (2.28) establish an isomorphism between a system of interacting quantum particles evolving in some external potential field and a classical system of separate necklaces composed of bead connected via harmonic springs. Beads with the same label in the different necklaces interact with each other and evolve in the same external potential.

2.2.1.3 Two-Electron System

In the previous section, we have established an isomorphism between a classical partition function and a path integral of *N* distinguishable particles. To extend this to many-fermion (and many-boson) systems, we first investigate a system of two electrons.

Subsequently, we will generalize the system to *N* indistinguishable fermions (or bosons).

Since identical particles cannot occupy the same state by the Pauli exclusion principle, the total wave function of a two-electron system should be antisymmetric upon exchange between electrons. Using this fact, we will introduce a new space in order to represent a state of two indistinguishable fermions. The new space is defined as

$$|r_1 r_2\} = \frac{1}{\sqrt{2}}(|r_1 r_2\rangle - |r_2 r_1\rangle \tag{2.29}$$

with one particle in state r_1 and one particle in state r_2. The closure relation of this space is

$$\frac{1}{2}\int dr_1 dr_2 |r_1 r_2\}\{r_1 r_2| = 1 \tag{2.30}$$

The density matrix of the 2-electron system can be written as

$$\rho\left(r_1, r_2; r_1', r_2'; \beta\right) = \left\{r_1, r_2 | e^{-\beta \widehat{H}} \middle| r_1', r_2'\right\} \tag{2.31}$$

If we consider an intermediate state $\left|r_1'' r_2''\right\}$ using the closure relation, we have a convolution relation for two identical particles such that

$$\rho\left(r_1, r_2; r_1', r_2'; \beta\right) = \int dr_1'' dr_2''\, \rho\left(r_1, r_2; r_1'', r_2'', \beta/2\right) \rho\left(r_1'', r_2''; r_1', r_2'; \beta/2\right) \tag{2.32}$$

We recall that the partition function is the trace of the density matrix and

$$Z = \int dr_1 dr_2 \rho(r_1, r_2; r_1, r_2; \beta) \tag{2.33}$$

Using the convolution relation, (2.32), for (P-1) intermediate states in imaginary time, the partition function now becomes

$$Z = \int \prod_{v=1}^{P} dr_1^{(v)} dr_2^{(2)} \prod_{k=1}^{P} \rho\left(r_1^{(k)}, r_2^{(k)}; r_1^{(k+1)}, r_2^{(k+1)}; \epsilon\right) \tag{2.34}$$

where $r_i = r_i^{(1)} = r_i^{(P+1)}$ with $i = 1, 2,$ that is, each electron forms a closed necklace with P nodes. Let us evaluate an infinitesimal element of the density matrix with (2.29).

$$\begin{aligned}
&\rho\left(r_1^{(k)}, r_2^{(k)}; r_1^{(k+1)}, r_2^{(k+1)}; \epsilon\right) \\
&\quad = \frac{1}{2}\left\{\left(\left\langle r_1^{(k)} r_2^{(k)}\right| - \left\langle r_2^{(k)} r_1^{(k)}\right|\right) e^{-\epsilon \widehat{H}} \left(\left|r_1^{(k+1)} r_2^{(k+1)}\right\rangle - \left|r_2^{(k+1)} r_1^{(k+1)}\right\rangle\right)\right\} \\
&\quad = \frac{1}{2}\left(\left\langle r_1^{(k)} r_2^{(k)}\middle| e^{-\epsilon \widehat{H}} \middle| r_1^{(k+1)} r_2^{(k+1)}\right\rangle + \left\langle r_2^{(k)} r_1^{(k)}\middle| e^{-\epsilon \widehat{H}} \middle| r_2^{(k+1)} r_1^{(k+1)}\right\rangle\right. \\
&\qquad \left. - \left\langle r_1^{(k)} r_2^{(k)}\middle| e^{-\epsilon \widehat{H}} \middle| r_2^{(k+1)} r_1^{(k+1)}\right\rangle - \left\langle r_2^{(k)} r_1^{(k)}\middle| e^{-\epsilon \widehat{H}} \middle| r_1^{(k+1)} r_2^{(k+1)}\right\rangle\right)
\end{aligned} \tag{2.35}$$

To evaluate the last terms of (2.35) including the cross terms between particles, we will consider a general case.

$$\begin{aligned}
\left\langle r_a^{(k)} r_b^{(k)}\middle| e^{-\epsilon \widehat{H}} \middle| r_c^{(k+1)} r_d^{(k+1)}\right\rangle &\approx \left\langle r_a^{(k)} r_b^{(k)}\middle| e^{-\epsilon \widehat{T}} e^{-\epsilon \widehat{V}} \middle| r_c^{(k+1)} r_d^{(k+1)}\right\rangle \\
&= r_a^{(k)} r_b^{(k)}\left| e^{-\epsilon \sum \widehat{p_i^2}/2m} \right| r_c^{(k+1)} r_d^{(k+1)}\right\rangle e^{-\epsilon V\left(r_c^{(k+1)} r_d^{(k+1)}\right)}
\end{aligned} \tag{2.36}$$

The term in (2.36), $\left\langle r_a^{(k)} r_b^{(k)} \left| e^{-\epsilon \sum \widehat{p_i^2}/2m} \right| r_c^{(k+1)} r_d^{(k+1)} \right\rangle$, is a density matrix of a free particle propagators, which can be evaluated exactly in the same way as for the classical particles obeying Maxwell-Boltzmann statistics. With the results of Appendix 1, we have

$$\left\langle r_a^{(k)} r_b^{(k)} \left| e^{-\epsilon \sum \widehat{p_i^2}/2m} \right| r_c^{(k+1)} r_d^{(k+1)} \right\rangle = \left(\frac{C_1}{2}\right)^{3N/2} e^{-C_2\left\{\left(r_a^{(k)} - r_c^{(k+1)}\right)^2 + \left(r_b^{(k)} - r_d^{(k+1)}\right)^2\right)\}} \tag{2.37}$$

Equation (2.37) shows that

$$\left\langle r_a^{(k)} r_b^{(k)} \left| e^{-\epsilon \sum \widehat{p_i^2}/2m} \right| r_c^{(k+1)} r_d^{(k+1)} \right\rangle = \left(\frac{C_1}{2}\right)^{3N/2} e^{-C_2\left\{\left(r_a^{(k)} - r_c^{(k+1)}\right)^2 + \left(r_b^{(k)} - r_d^{(k+1)}\right)^2\right)\}} \tag{2.38}$$

If we consider a symmetric potential, such as a pair-wise additive central potential, (2.35) can be simplified as

$$\begin{aligned} \rho\left(r_1^{(k)}, r_2^{(k)}; r_1^{(k+1)}, r_2^{(k+1)}; \epsilon\right) &= \left\langle r_1^{(k)} r_2^{(k)} | e^{-\epsilon \widehat{H}} | r_1^{(k+1)} r_2^{(k+1)} \right\rangle \\ &\quad - \left\langle r_1^{(k)} r_2^{(k)} | e^{-\epsilon \widehat{H}} | r_2^{(k+1)} r_1^{(k+1)} \right\rangle \end{aligned} \tag{2.39}$$

From (2.34–2.39), the partition function of the two electrons finally becomes

$$\begin{aligned} Z = &\left(\frac{C_1}{2}\right)^{3NP/2} \int \prod_{v=1}^{P} dr_1^{(v)} dr_2^{(v)} \prod_{k=1}^{P} \left\{ e^{-\beta C_0 \left(\left(r_1^{(k)} - r_1^{(k+1)}\right)^2 + \left(r_2^{(k)} - r_2^{(k+1)}\right)^2\right)} \right. \\ &\left. - e^{-\beta C_0 \left(\left(r_1^{(k)} - r_2^{(k+1)}\right)^2 + \left(r_2^{(K)} - r_1^{(k+1)}\right)^2\right)} \right\} e^{-\beta \frac{1}{P} V_2^{(k)}} \end{aligned} \tag{2.40}$$

The term in braces in (2.40) is representative of an exchange process between the two electrons as shown in the Fig. 2.2.

By factoring the first term out of { } in (2.40), we have

$$\begin{aligned} \{\} = &e^{-\beta C_0 \left(\left(r_1^{(k)} - r_1^{(k+1)}\right)^2 + \left(r_2^{(k)} - r_2^{(k+1)}\right)^2\right)} \\ &\times \left\{ 1 - e^{-\beta C_0 \left(\left(r_1^{(k)} - r_2^{(k+1)}\right)^2 - \left(r_1^{(k)} - r_1^{(k+1)}\right)^2\right) + \left(\left(r_2^{(k)} - r_1^{(k+1)}\right)^2 - \left(r_2^{(k)} - r_2^{(k+1)}\right)^2\right)} \right\}. \end{aligned}$$

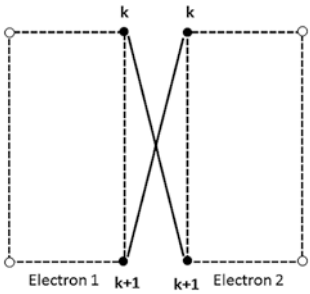

Fig. 2.2 Exchange process between two electrons

For convenience, we may rewrite this term by considering the product terms as

$$\begin{aligned}\prod_{k=1}^{P}\{\} &= \prod_{k=1}^{P} e^{-\beta C_0\left(\left(r_1^{(k)}-r_1^{(k+1)}\right)^2+\left(r_2^{(k)}-r_2^{(k+1)}\right)^2\right)}\cdot\prod_{k=1}^{P}\det\left(E^{(k,k+1)}\right)\\ &= \exp\left(-\beta C_0\sum_{i=1}^{2}\sum_{k=1}^{P}\left(r_i^{(k)}-r_i^{(k+1)}\right)^2\right)\cdot\prod_{k=1}^{P}\det\left(E^{k,k+1}\right)\end{aligned} \tag{2.41}$$

with

$$\det\left(E^{(k,k+1)}\right) = \begin{vmatrix} 1 & e^{-\beta C_0\left(\left(r_2^{(k)}-r_1^{(k+1)}\right)^2+\left(r_2^{(k)}-r_2^{(k+1)}\right)^2\right)} \\ e^{-\beta C_0\left(\left(r_2^{(k)}-r_1^{(k+1)}\right)^2+\left(r_2^{(k)}-r_2^{(k+1)}\right)^2\right)} & 1 \end{vmatrix}$$

From (2.40) and (2.41), the partition function of the two-electron system becomes

$$\begin{aligned}Z = \left(\frac{C_1}{2}\right)^{\frac{3NP}{2}}\int\prod_{v=1}^{P}dr_1^{(v)}dr_2^{(v)}\exp\left(-\beta C_0\sum_{i=1}^{2}\sum_{k=1}^{P}\left(r_i^{(k)}-r_i^{(k+1)}\right)^2\right)\\ \times\prod_{k=1}^{P}\det\left(E^{(k,k+1)}\right)e^{-\beta V_2^{(k)}/P}\end{aligned} \tag{2.42}$$

We note that the density matrix of two electrons estimated between two infinitesimally close imaginary times k and $k + 1$ is proportional to $\det\left(E^{(k,k+1)}\right)$ so that the sign of the infinitesimal density matrix can be either a positive number or a negative number. This is the origin of the so-called sign problem in the simulation of many-fermion systems. In particular, the partition function (2.42) is an integration of terms which can be either positive or negative so one can think of the partition function as the difference between a large positive number and a large

negative number to give an overall positive value. Therefore the noise level of the partition function is large and this is the difficulty to overcome during the modeling and simulation of a many-fermion system. Furthermore since $\det\left(E^{(k,k+1)}\right)$ can be negative, it is not possible to rewrite Z in the form of a classical partition function. Thus one cannot establish yet an isomorphism between the fermion system and a classical one. In order to solve the sign problem, we will adopt the fixed-node path integral method in our model. The details of the fixed-node path integral method and its usage in our model of many-electron system will be discussed in Sect. 2.3.

2.2.1.4 Many-Electron System

We may extend the partition function of the two-electron system, (2.42), to that of an indistinguishable N-body quantum system. Because of the indistinguishability between particles, we will introduce an orthonormal basis of the indistinguishable N-body fermion system;

$$|r_1 r_2 \cdots r_N\} = \frac{1}{\sqrt{N!}} \sum\nolimits_{\wp} (-1)^{\wp} \left| r_{\wp 1} r_{\wp 2} \cdots r_{\wp N} \right\rangle \tag{2.43}$$

where $|r_1 r_2 \cdots r_N$ is an orthonormal basis of an distinguishable N-body system, and $\wp$ is the parity of the permutation and $(-1)^{\wp}$ become +1 and −1 for even and odd number of the permutations between fermions, respectively. The closure relation of the new basis becomes

$$\frac{1}{N!} \sum\nolimits_{i=1}^{N} |r_1 r_2 \cdots r_N\}\{r_1 r_2 \cdots r_N| = 1 \tag{2.44}$$

With the new basis, we may define the density matrix of the indistinguishable N–body fermion system as

$$\rho(R, R'; \beta) = \left\{ R | e^{\beta \widehat{H}} | R' \right\} \tag{2.45}$$

or

$$\rho(R, R'; \beta) = \frac{1}{N!} \sum\nolimits_{\wp} (-1)^{\wp} \left\langle R_{\wp} \left| e^{\beta \widehat{H}} \right| R' \right\rangle \tag{2.46}$$

where $|R\} = |r_1 r_2 \cdots r_N\}$ and $|R\rangle = |\, r_1 r_2 \cdots r_N\rangle$. If we consider P intermediate states and the closure relation, (2.44), we may have the following convolution relation for the density matrix:

$$\rho(R,R';\beta) = \int \prod_{i=1}^{P} d\mathcal{R}^{(i)}\, \rho^{(0)}\, \rho^{(1)} \cdots \rho^{(P)} \tag{2.47}$$

where $\rho^{(k)} = \rho\left(R^{(k)}, R^{(k+1)}; \epsilon\right), R^{(0)} = R, R^{(P)} = R'$, and $\epsilon = \beta/\mathrm{P}$, which is the infinitesimal imaginary time interval.

The partition function, which is the trace of the density matrix, for an indistinguishable N-body quantum system may be written in the form

$$\begin{aligned} Z &= \int dR \rho(R,R;\beta) \\ &= \int \prod_{i=0}^{P} dR^{(i)}\, \rho^{(0)}\, \rho^{(1)} \cdots \rho^{(P)} \end{aligned} \tag{2.48}$$

In order to calculate the partition function, we have to evaluate the infinitesimal density matrix, $\rho\left(R^{(k)}, R^{(k+1)}; \epsilon\right)$. For further development, we will assume that the Hamiltonian of the system is: $\widehat{H} = \widehat{T} + \widehat{V}_2$, where again

$$\begin{aligned} \widehat{T} &= \sum_{i=1}^{N} \frac{\widehat{p}^2}{2m} \\ \widehat{V}_2 &= \sum_{i=1}^{N} \widehat{\phi}(r_i) + \sum_{i,j}^{N} \widehat{v}\left(\left|r_i - r_j\right|\right), \end{aligned}$$

$\phi(r_i)$ is an external potential at r_i and $v\left(\left|r_i - r_j\right|\right)$ is a pair potential between particles i and j. With the Trotter's approximation, the infinitesimal density matrix becomes

$$\begin{aligned} &\rho\left(R^{(k)}, R^{(k+1)}; \epsilon\right) \\ &\quad = \frac{1}{N!} \sum_{\wp} (-1)^{\wp} \left\langle r_{\wp 1}^{(k)} \cdots r_{\wp N}^{(k)} \middle| e^{-\epsilon \widehat{H}} \middle| r_1^{(k+1)} \cdots r_N^{(k+1)} \right\rangle \\ &\quad \approx \frac{1}{N!} \sum_{\wp} (-1)^{\wp} \left\langle r_{\wp 1}^{(k)} \cdots r_{\wp N}^{(k)} \middle| e^{-\epsilon \widehat{H}} e^{-\epsilon \widehat{V}} \middle| r_1^{(k+1)} \cdots r_N^{(k+1)} \right\rangle \end{aligned}$$

If we apply the closure relation, $\int dR^{(k)} \left|R^{(k)}\right\rangle \left\langle R^{(k)}\right| = 1$ and the orthonormal property of the basis $\left|R^{(k)}\right\rangle$, the infinitesimal density matrix can be written as

$$\rho\left(R^{(k)}, R^{(k+1)}; \epsilon\right) = \rho_0\left(R^{(k)}, R^{(k+1)}; \epsilon\right) e^{-\epsilon V_2\left(R^{(k)}\right)} \tag{2.49}$$

where the infinitesimal density matrix of free particles is defined as

$$\rho_0\left(R^{(k)}, R^{(k+1)}; \epsilon\right) = \frac{1}{N!}\sum_{\wp}(-1)^{\wp}\left\langle r_{\wp 1}^{(k)} \cdots r_{\wp N}^{(k)}\middle| e^{-\epsilon \widehat{T}} \middle| r_1^{(k+1)} \cdots r_N^{(k+1)}\right\rangle$$

The infinitesimal density matrix of free particles contains the exchange processes between identical particles and it can be written in a determinant form (see Fig. 2.3) From (2.25), the infinitesimal density matrix of free particles becomes

$$\rho_0\left(R^{(k)}, R^{(k+1)}; \epsilon\right) = \frac{1}{N!}\sum_{\wp}(-1)^{\wp}\left(\frac{C_1}{2}\right)^{3N/2}\exp\left(-C_2\sum_{i=1}^{N}\left(r_{\wp i}^{(k)} - r_i^{(k+1)}\right)^2\right),$$

with $C_2 = m/2\epsilon\hbar^2$. If we perform the summation with the permutation operator $\wp$, we can simplify the density matrix by using a determinant:

$$\rho_0\left(R^{(k)}, R^{(k+1)}; \epsilon\right) = \left(\frac{C_1}{2N!}\right)^{3NP/2}\det\left(A^{(k,k+1)}\right) \tag{2.50}$$

where the matrix $A^{(k,k+1)}$ is defined as

$$\left(A^{(k,k+1)}\right)_{ij} = e^{-\beta C_0\left(r_i^{(k)} - r_j^{(k+1)}\right)^2}$$

Furthermore, the exchange processes between particles can be illustrated by writing the infinitesimal density matrix of free particle (3-particle case) as

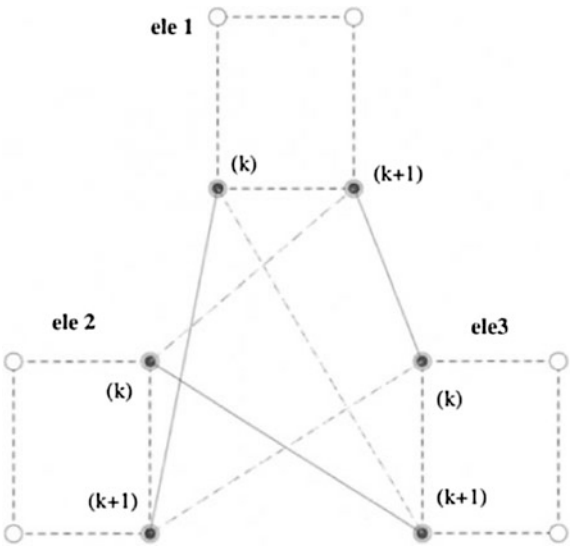

Fig. 2.3 Exchange process among three identical particles. *Solid lines* and *dashed lines* represent two different exchange processes. Electrons are labeled "ele" 1, 2, and 3. The time slices are labeled, k and $k + 1$

$$\rho_0\left(R^{(k)},R^{(k+1)};\epsilon\right)=\left(\frac{C_1}{2N!}\right)^{3NP/2}A_0\cdot(1-f_{12}-f_{23}-f_{31}+g_{231}+g_{312}) \quad (2.51)$$

where

$$\begin{aligned}A_0&=e^{-C_2\sum_{i=1}^{N}\left(r_i^{(k)}-r_i^{(k+1)}\right)^2}\\ f_{ii}&=e^{-C_2\left(\left(r_i^{(k)}-r_j^{(k+1)}\right)^2-\left(r_i^{(k)}-r_i^{(k+1)}\right)^2\right)}\cdot e^{-C_2\left(\left(r_j^{(k)}-r_i^{(k+1)}\right)^2-\left(r_j^{(k)}-r_j^{(k+1)}\right)^2\right)}\\ g_{ijl}&=\frac{1}{A_0}\cdot e^{-C_2\left(\left(r_i^{(k)}-r_j^{(k+1)}\right)^2}e^{-C_2\left(\left(r_j^{(k)}-r_l^{(k+1)}\right)^2}e^{-C_2\left(\left(r_l^{(k)}-r_i^{(k+1)}\right)^2}\end{aligned} \quad (2.52)$$

The quantities f_{ij} and g_{ijl} correspond to exchange processes between two particles (i, j) and exchange processes among three particles (i, j, l), respectively (again see Fig. 2.3). The determinant of the density matrix of the free particles in absence of quantum exchange is factored out of (2.50);

$$\det\left(A^{(k,k+1)}\right)_{ij}=\prod_{i=1}^{N}\left(A^{(k,k+1)}\right)_{ii}\cdot\det\left(E^{(k,k+1)}\right) \quad (2.53)$$

where all the exchange effects (including the sign of the density matrix) are included in $\det\left(E^{(k,k+1)}\right)$ which elements are defined as $\left(E^{(k,k+1)}\right)_{ij}=\left(A^{((k,k+1)}\right)_{ij}\Big/\left(A^{((k,k+1)}\right)_{ii}$.

Specifically, the matrix element of $\left(E^{(k,k+1)}\right)$ for the N-particle system is given by

$$\left(E^{(k,k+1)}\right)_{ij}=\exp\left[-\beta C_1\left\{\left(r_i^{(K)}-r_j^{(k+1)}\right)^2-\left(r_i^{(k)}-r_i^{(k+1)}\right)^2\right\}\right] \quad (2.54)$$

From (2.48, 2.50, 2.53, 2.54), we finally write the partition function of the N-fermion system as

$$\begin{aligned}Z=&\left(\frac{C_1}{2N!}\right)^{3NP/2}\int\prod_{v=1}^{P}\prod_{j=1}^{N}dr_j^{(v)}\exp\left(-\beta C_0\sum_{i=1}^{N}\sum_{k=1}^{P}\left(r_i^{(k)}-r_i^{(k+1)}\right)^2\right)\\ &\times\prod_{k=1}^{P}\det\left(E^{(k,k+1)}\right)e^{-\beta V_2/P}\end{aligned} \quad (2.55)$$

where the classical potential energy V_2 is the same as in (2.28), i.e.

$$V_2=\frac{1}{P}\sum_{k=1}^{P}\sum_{i=1}^{N}\emptyset\left(r_i^{(k)}\right)+\frac{1}{P}\sum_{k=1}^{P}\sum_{i>j}^{N}v\left(r_i^{(k)}-r_j^{(k)}\right) \quad (2.56)$$

2.2.2 Path Integral with Non-local Exchange Using a Mean Field Approximation

In the preceding sections, we have developed a local non-interacting density matrix which does not describe electron correlation, since the free particle density matrix has been obtained by using a complete set of states represented by Slater determinants of plane waves and Slater determinant of plane waves are solutions to the Hartree-Fock equation for free electrons. Although the local non-interacting density matrix does not include electron correlation, in the limit of high temperature, its nodes approximate reasonably well those of the exact density matrix [12]. We now construct an approximate form for the density matrix that includes electron correlation. In order to treat the correlation between like-spin electrons, Hall has proposed a non-local exchange pseudo-potential [9, 10]. In the local form of the density matrix (2.51), $\det\left(E^{(k,k+1)}\right)$ includes all the exchange effects, although exchange occurs only between consecutive beads in imaginary time. Furthermore, in the limit of $\epsilon \to 0$, the matrix $E^{(k,k+1)}$ converges to the identity matrix and the system collapses into a boson state. To avoid this undesirable behavior and inspired by the consideration of quantum chemistry simulation, where exchange is treated as a non-local interaction in space, Hall has suggested a non-local form of a density matrix of a two-electron system as follows;

$$\det\left(E^{(k,k+1)}\right) \to \prod_{l=1}^{P} \det\left(E_{\alpha}^{(k,l)}\right) \tag{2.57}$$

where the matrix element $\left(E_{\alpha}^{(k,l)}\right)_{ij}$ is defined as

$$\left(E_{\alpha}^{(k,l)}\right)_{ij} = \exp\left(-\frac{m\,\alpha}{2\epsilon\hbar^2}\left|\left(r_i^{(k)} - r_j^{(l)}\right)^2 - \left(r_i^{(k)} - r_i^{(l)}\right)^2\right|\right) \tag{2.58}$$

In the preceding relations, the superscripts and the subscripts label the beads and electrons, respectively. We note that α in (2.57 and 2.58) is a system dependent parameter and the absolute value of the argument of the exponential prevents negative weights. Because it is not easy to find a proper parameter α for a system and it is clear that one underestimates the contribution from negative values by choosing the absolute value, we generalize the non-local density matrix of an N-fermion system by choosing the following:

$$\det\left(A^{(k,k+1)}\right) \to \prod_{i=1}^{N}\left(A^{(k,k+1)}\right)_{ii} \cdot \prod_{l=1}^{P} \det\left(E^{(k,l}\right)^{1/P} \tag{2.59}$$

The non-local form of the infinitesimal density matrix of an N-fermion system and the corresponding partition function now becomes

$$\rho_0\left(R^{(k)}, R^{(k+1)}; \epsilon\right) = \left(\frac{C_1}{2N!}\right)^{3NP/2} \prod_{i=1}^{N} \left(A^{(k,k+1)}\right)_{ii} \cdot \prod_{l=1}^{P} \det\left(E^{(k,l)}\right)^{1/P}. \tag{2.60}$$

and

$$Z = \left(\frac{C_1}{2N!}\right)^{3NP/2} \int \prod_{v=1}^{P} \prod_{j=1}^{N} dr_j^{(v)} \exp\left(-\beta C_0 \sum_{i=1}^{N} \sum_{k=1}^{P} \left(r_i^{(k)} - r_i^{(k+1)}\right)^2\right) \times \prod_{l=1}^{P} \prod_{k=1}^{P} \det\left(E^{(k,l)}\right) e^{-\beta V_2/P} \tag{2.61}$$

The non-local form for the density matrix cannot be obtained from simple Slater determinants of plane waves. Equation (2.60) should therefore represent electrons beyond the Hartree-Fock approximation. A non-local density matrix would account for some electron correlation. In Fig. 2.4, we illustrate both the local exchange process in (a) and the non-local exchange process in (b). In the non-local exchange model, a bead of an electron interacts with all beads of any other electron.

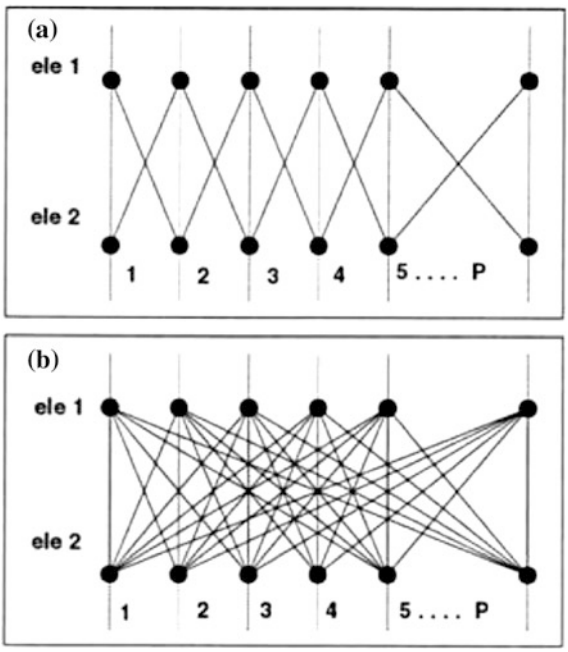

Fig. 2.4 **a** Local and **b** non-local exchange processes between two electrons. The *number labels* indicate the imaginary times. The electrons are denoted "ele" 1 and 2

Although the non-local form of the density matrix does not collapse to a boson state when $P \rightarrow \infty$, it still has the sign problem, because the determinant values are either positive and negative. In the next section, we will discuss and present a solution to the fermionic sign problem by introducing the restricted path-integral Monte Carlo method.

2.2.3 Restricted Path Integral Method

There is a fundamental difficulty in the simulation of many-body fermionic systems, called the sign problem. The sign problem arises from permutations between identical particles. The contribution to the partition function of even permutations is almost the same as the contribution from odd permutations. In practical calculations of thermodynamic properties, one cannot expect accurate results, because of the large signal to noise ratio. The study of the sign problem and the search for better conditioned models and simulation algorithm are widely discussed subjects in the simulation of many-electron systems [8, 12, 16, 46–48]. In more recent studies, a restricted fixed-node path integral approximation has been suggested to solve the sign problem of a many-fermion system with path integral Monte Carlo simulation. In this approximation, the paths of all fermions in imaginary time are restricted to remain within the region of phase space where the density matrix is positive. This approximation has been applied to liquid ^{3}He above 1 K [12] and the hydrogen plasma at high temperature and reasonable agreement to the existing theories was found.

The main idea of the restricted fixed-node approach was initially proposed by Metropolis and Ulam [49, 50]. They suggested the extension of the random-walk process typically used to simulate the diffusion equation for solving Schrodinger equation. Anderson [46] applied the restricted fixed-node scheme to obtain the ground state of simple quantum molecular systems. Anderson also solved the Schrodinger equation using the random-walk methods. We summarize below the restricted fixed-node path integral idea suggested by Ceperley [12, 16, 17] and then apply it to the development of models of fermion systems in the framework of the path- Integral formalism.

The one-dimensional diffusion equation can be written as

$$\frac{\partial C}{\partial t} = D\frac{\partial^2 C}{\partial x^2} \tag{2.62}$$

where $C = C(x, t)$ is a concentration, D is a diffusion coefficient and t is a real time.

The solution of the diffusion equation for an initial delta concentration at x_o (i.e. we assume that the initial condition is $C(x, t = 0) = \delta(x - x_0)$), is given by the Gaussian function:

$$C(x,t) = \frac{1}{\sqrt{4\pi Dt}} e^{-\frac{(x-x_0)^2}{4Dt}} \tag{2.63}$$

As we mentioned in Sect. 2.1, the diffusion equation is isomorphic to the imaginary time-dependent Schrodinger equation of a free particle density matrix;

$$\frac{\partial \rho(x,t)}{\partial \beta} = \frac{\hbar^2}{2m} \frac{\partial^2 \rho(x,t)}{\partial x^2} \tag{2.64}$$

by replacing β by $it/\hbar$.

If we assume a boundary condition of the diffusion equation in the form:

$$C(x - x') = 0, \tag{2.65}$$

the solution of (2.62) becomes

$$C(x,t) = \frac{1}{\sqrt{4\pi Dt}} e^{-\frac{(x-x_0-x')^2}{4Dt}} - \frac{1}{\sqrt{4\pi Dt}} e^{-\frac{(x+x_0-x')^2}{4Dt}} \tag{2.66}$$

This solution is completely determined by the boundary condition. The solution given by (2.66) has an antisymmetric form in space about x' and we can impose an infinite potential barrier at the boundary x' without loss of generality. The locus of points x' will be called a *fixed node*. By virtue of the isomorphism between the diffusion equation and the imaginary-time Schrodinger equation, we may apply an infinite potential barrier at the nodes of the density matrix, with the fact that the trace of the density matrix should be a positive real physical quantity and be spatially antisymmetric if exchange occurs between like-spin electrons. In the path integral, the trace of a density matrix $\rho(x, x; \beta)$ is always positive, but the infinitesimal density matrix, $\rho(x_i, x_{i+1}; \tau = \beta/P)$, where $i = 1, \ldots, P$, can be either positive or negative. We can choose a configuration of beads of a necklace such that x_i and x_{i+1} are on the same side of the boundary node so that $\rho(x_i, x_{i+1}; \tau)$ for all i are positive. In the fixed-node Path Integral Monte Carlo method, one begins with a trial density matrix, which nodes are known. So if the trial density matrix is exact, the method then becomes exact.

To clarify the meaning of the fixed-node and of the sign problem, we consider the special case of a two-body system (electrons 1 and 2) in 3-dimensions with an exchange process between a pair of beads (i) and (j). An element of the infinitesimal density matrix is

$$\left(E^{(i,j)}\right)_{12} = e^{-C_1\left\{\left(x^{(i)}-y^{(j)}\right)^2-\left(x^{(i)}-x^{(j)}\right)^2\right\}}$$

where $x^{(i)}$ and $y^{(j)}$ are the coordinates of the electrons 1 and 2, respectively. Then the determinant of the matrix is

$$\det\left(E^{(i,j)}\right)_{12} = \begin{vmatrix} 1 & e^{-C_1\left\{\left(x^{(i)}-y^{(j)}\right)^2-\left(x^{(i)}-x^{(j)}\right)^2\right\}} \\ e^{-C_1\left\{\left(y^{(i)}-x^{(j)}\right)^2-\left(y^{(i)}-y^{(j)}\right)^2\right\}} & 1 \end{vmatrix}$$
$$= 1 - e^{-2C_1\left(x^{(i)}-y^{(i)}\right)\cdot\left(x^{(j)}-y^{(j)}\right)}$$

With relative coordinates $x = \left(x^{(i)} - y^{(i)}\right)/\sqrt{2c_1}$ and $y = \left(x^{(j)} - y^{(j)}\right)/\sqrt{2c_1}$, the determinant can be rewritten as

$$\det\left(E^{(i,j)}\right) = 1 - e^{-x\cdot y} \tag{2.67}$$

In the relative coordinates, the sign of the determinant now has the same sign as the dot product $(x \cdot y)$. The boundary between regions with positive or negative sign is $x \cdot y = 0$. This is illustrated in Fig. 2.5. Sampling the region of the (x, y) plane

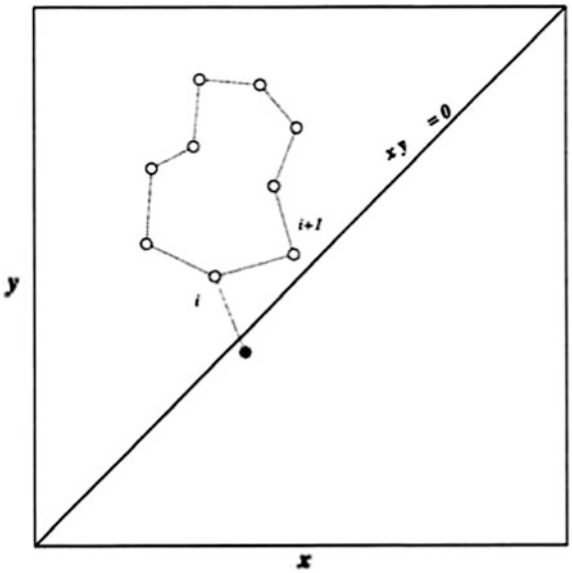

Fig. 2.5 Diagrammatic representation of configuration space for one exchange process between two electrons (see text for details). The *diagonal line* is $x \cdot y = 0$ and is the boundary delimiting the transition from a positive to a negative sign of the density matrix. All *unfilled dots* correspond to configurations with positive density matrices and the *filled dot* has a density matrix with a negative value

with only positive determinants to calculate the corresponding partition function is equivalent to calculating the partition function by sampling the entire (x, y) plane. The fixed node restriction based on an assumed form of the density matrix imposes the necessary boundary condition.

2.2.4 Classical Isomorphism for Many-Body Fermion System

In establishing a model for a many-body fermion system, we are not evaluating the partition function directly. Instead we are trying to establish an isomorphism between the quantum partition function and a classical partition function. With the restricted path-integral method, we can always constrain the configurations of the particles in the system to regions of configuration space with positive density matrices. We therefore can rewrite the partition function of the many-body fermion system, (2.61), in a form isomorphic to a classical one:

$$Z \approx \frac{1}{N!}\left(\frac{mP}{2\pi\beta\hbar^2}\right)^{3NP/2} \int \prod_{i=1}^{N} \prod_{k=1}^{P} dr_i^{(k)} e^{-\beta V_{eff}} \cdot e^{-\frac{\beta}{P} V_2} \tag{2.68}$$

where the integration is limited to configurations with $\det\left(E^{(k,l)}\right) > 0$. In (2.68), V_2 is a classical potential energy function of position. If the potential V_2 represents the Coulomb interactions between electrons, it can be written as

$$V_2 = V_2\left(r_{ij}^{(k)}\right) \tag{2.69}$$

where $r_{ij}^{(k)} = \left|r_i^{(k)} - r_j^{(k)}\right|$ denotes the distance between the kth beads of the electrons i and j. If the potential V_2 arises from interactions between electrons and ions, the distance between some ion I and the kth beads of the electrons j is $r_{Ij}^{(k)} = \left|r_I - r_j^{(k)}\right|$. V_{eff} in (2.68) is a quantum effective potential energy. The classical potential function includes electron/electron Coulomb interactions or electron/ion interactions. In contrast, V_{eff} includes quantum exchange energy between electrons. From (2.68, 2.61), we can define the effective potential, V_{eff}, as

$$V_{eff} = V_{eff}^{harm} + V_{eff}^{exch} \tag{2.70}$$

where

$$V_{eff}^{harm} = \frac{mP}{2\beta^2\hbar^2} \sum\nolimits_{i=1}^{N} \sum\nolimits_{k=1}^{P} * \left(r_j^{(k)} - r_j^{(k+1)}\right)^2 \tag{2.71}$$

and

$$V_{eff}^{exch} = -\frac{1}{\beta}\sum_{k=1}^{P}\sum_{l=1}^{P}\frac{1}{Q}\ln\left(\det\left(E^{(k,l)}\right)\right) \tag{2.72}$$

where $\det\left(E^{(k,l)}\right) > 0$ and Q is the effective number of paths with $\det\left(E^{(k,l)}\right) > 0$.

V_{eff}^{harm} is a non-exchange harmonic potential and V_{eff}^{exch} is a non-local quantum exchange potential. In the non-local form of path-integral dynamics, a quantum particle is still represented by a necklace of P beads such that a point in the necklace interacts with its next consecutive neighbor along the chained necklace through a harmonic potential with a strength $mP/2\beta^2\hbar^2$. In contrast to the harmonic potential, the exchanges between different particles are not limited to the nearest neighbors along the necklaces, but act over all beads (imaginary time slices) of the different necklaces.

In order to represent V_{eff}^{exch} with a more practical form for implementation of a restricted path-integral molecular dynamics, we introduce a step function θ^{+}. The function θ_{kl}^{+} ensures the path restriction by taking on the values 1 and 0 for paths with positive and negative $\det\left(E^{(k,l)}\right)$, respectively. In addition, for a system containing electrons with two different types of spins (i.e. spin-up and spin-down electrons), we may use the fact that the density matrix is approximated as the product of two determinants taking the form of (2.72); one determinant for the electrons with one type of spin and another determinant for the electrons with the other type of spin [51]. We now rewrite V_{eff}^{exch} as

$$V_{eff}^{exch} = -\frac{1}{\beta}\sum_{s=up}^{down}\sum_{k=1}^{P}\sum_{l=1}^{P}\frac{1}{Q_s}\ln\left(\det\left(E^{(k,l)}\right)\right)\theta_{kls}^{+} \tag{2.73}$$

where

$$Q_s = \sum_{k=1}^{P}\sum_{l=1}^{P}\theta_{kls}^{+}$$

In the previous two equations, s in the summation denotes the spin of the electrons.

Up to now, we have written the partition function of a quantum fermion system in the form of a classical partition function. The potential V_2 was assumed to local and independent of angular momentum. In the next section, we derive expressions for the path integral of a quantum particle in a non-local pseudo-potential.

2.2.5 Path Integral with Non-local Pseudo-potential

The spherical symmetry of an atomic potential that an electron may experience makes it easier to derive the path integral in spherical coordinates. Actually, an expression for the path integral in these coordinates has already been found [52–54]. The potential that was used in these derivations was of local nature. Also, the derivation of the path integral using non-local potentials has already been done [55]. In this section we present a new way of incorporating a non-local pseudo-potential in the derivation of the path integral in spherical coordinates.

2.2.5.1 Local and Nonlocal Pseudo-potential

A local potential is a function of the spatial coordinates. In other words, a local operator, C_{OP}, operates on a given function of coordinates, $\Psi(\vec{r})$, in the following manner:

$$C_{op}\Psi(\vec{r}) = C(\vec{r}).\Psi(\vec{r}) \tag{2.74}$$

where $C(\vec{r})$ is the Eigenvalue of the operator corresponding to the physical observable C (for example, C could be the linear momentum). To clarify this point further, if the local operator is a potential then the value of the wave function at any point in space is independent of its value at other points. In general, the coordinates of the local operator do not have to be only spatial for an operator to be local. An operator can be a function of both spatial and spin coordinates and still be local [56].

On the other hand, a non-local potential is not only a function of spatial and/or spin coordinates, but involves other types of coordinates. In other words, the value of the wave function at any point in space will depend not only on the potential at that point, but also on the potential at points other than the one in question. For example, a non-local potential is equivalent under certain conditions, to a velocity dependent potential [57]. When nonlocal pseudo-potentials are used, the Schrödinger equation becomes an integro-differential equation that is not solvable exactly in a closed form.

In operator language, a non-local operator, C_{OP}, operates on a given function, $\Psi(\vec{r})$, in the following manner:

$$C_{op}\,\Psi(\vec{r}) = \int C\left(\vec{r}, \overrightarrow{r'}\right)\Psi(\vec{r}')\,d\vec{r}' \tag{2.75}$$

With the non-locality one is able to account for the difference in potential felt by the different angular momentum components of the valence wave function. After all, the pseudo potential, as originally derived by Phillips and Kleinman [58], is a non-local operator.

For a particle evolving in a central potential the representation of the propagator based on Lagrangian formulation of the path integral is given by [52–54]:

$$K(\vec{r}_2, \vec{r}_1, \tau) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} \mathrm{K}_\ell(r_2, r_1, \tau)_\ell Y^*_{\ell m}(\theta_2, \phi_2) Y_{\ell m}(\theta_1, \phi_1) \tag{2.76}$$

where τ is the time it takes the particle to go from the initial position $\vec{r}_1$ to the final one $\vec{r}_2$. The amplitude in the radial direction is given by [35]:

$$\begin{aligned} \mathrm{K}_\ell(r_2, r_1, \tau) = & \lim_{\mathrm{N}\to\infty} (4\pi)^{\mathrm{N}} \left(\frac{m}{2\pi i \in_0}\right)^{\frac{3N}{2}} \\ & \times \int \prod_{i=1}^{N} [R_\ell(r_i, r_{i-1})] \prod_{i-1}^{N-1} r_i^2 dr_i \end{aligned} \tag{2.77}$$

where $\mathrm{R}_\ell(r_\mathrm{i}, r_{\mathrm{i}-1})$ is given by:

$$\begin{aligned} R_\ell(r_i, r_{i-1}) = & \left[\frac{i\pi \in_0}{2mr_i r_{i-1}}\right]^{\frac{1}{2}} \exp\left[\frac{im}{2 \in_0} \left(r_i^2 + r_{i-1}^2\right) - i \in_0 V(r_i)\right] \\ & \times I_{\ell+\frac{1}{2}}\left(\frac{m}{i \in_0} r_i r_{i-1}\right) \end{aligned} \tag{2.78}$$

where V is the central force potential, ϵ_0 is τ/N, *and* $I_{l+1/2}\,(mr_\mathrm{i}r_{\mathrm{i}-1}/\mathrm{i}\epsilon_0)$ is the modified Bessel function. The path integral given in (2.76) applies to local potentials or local pseudo-potentials only. Due to the nature of non-local operators, one cannot simply add the non-local pseudo-potential to the exponent in the exponential function of (2.2). Certain measures of caution must be taken in deriving the expression for the path integral in the case of non-local pseudo potentials.

2.2.5.2 One Electron System

In this section we deal with a one electron system. That is, an atom with one valence electron. If we decompose the effective potential into a local parts and in non-local one, then the Hamiltonian operator of our system could be written as follows:

$$H_{op} = T_{op} + V_{loc} + V_{NL} \tag{2.79}$$

where T_{op} is the kinetic energy operator, V_{loc} is the local potential operator, and V_{NL} is the non-local pseudopotential operator.

We substitute (2.79) into the density matrix element expression:

$$\rho = \left\langle \vec{r}_n \middle| e^{-\frac{\beta H_{op}}{P}} \middle| \vec{r}_{n+1} \right\rangle \tag{2.80}$$

to obtain the form:

$$\rho = \left\langle \vec{r}_n \middle| e^{-\frac{\beta}{P}\left[T_{op+} V_{loc} + V_{NL}\right]} \middle| \vec{r}_{n+1} \right\rangle \tag{2.81}$$

One can now use the Trotter formula, to write:

$$\left\langle \vec{r}_n \middle| e^{-\epsilon H_{op}} \middle| \vec{r}_{n+1} \right\rangle = \left\langle \vec{r}_n \middle| e^{-\epsilon T_{op}} e^{-\epsilon V_{loc}} e^{-\epsilon V_{NL}} \middle| \vec{r}_{n+1} \right\rangle \tag{2.82}$$

where $\epsilon = \beta/P$.

The non-local pseudo potential operator that we use in our derivation is of the form [59, 60]:

$$V_{NL} = \sum\nolimits_{\ell=0}^{\infty} \sum\nolimits_{m-\ell}^{\ell} |\ell m\rangle V_\ell(r) \langle \ell m| \tag{2.83}$$

This expression carries the fact that each angular momentum pseudo-valence-orbit feels a different potential. Notice that a different local potential $V_l(r)$ Potential is needed for different l's. From (2.83) we can see that the non-local operator could be decomposed into a sum of different operators that have particular symmetry [61]:

$$V_{NL} = \sum\nolimits_{\ell=0}^{\infty} \sum\nolimits_{m-\ell}^{\ell} |\ell m\rangle V_\ell(r) \langle \ell m| \tag{2.84}$$

$$V_{NL} = V_s + V_p + V_d + \ldots \tag{2.85}$$

where, for example, the operator versus acts only on orbitals processing an s symmetry.

Now if we introduced the closure relation in (2.82) we get

$$\rho = \int d\vec{r} \left\langle \vec{r}_n \middle| e^{-\in V_{loc}} \middle| \vec{r} \right\rangle \int d\vec{r}' \left\langle \vec{r} \middle| e^{-\in V_{NL}} \middle| \vec{r}' \right\rangle \times \left\langle \vec{r}' \middle| e^{-\in T_{op}} \middle| \vec{r}_{n+1} \right\rangle \tag{2.86}$$

where the ρ stands for the left-hand side of the (2.82). Using power series expansion of the exponential it is easy to see that:

$$\rho = \int d\vec{r} \delta(\vec{r} - \vec{r}_n) e^{-\epsilon V_{loc}(\vec{r})} \times \int d\vec{r}' \langle \vec{r} | e^{-\epsilon V_{NL}} | \vec{r}' \rangle \langle \vec{r}' | e^{-\epsilon T_{op}} | \vec{r}_{n+1} \rangle \tag{2.87}$$

where we have used the orthonormality relation. Note that in this case $V_{loc}(\vec{r})$ is a function of the position vector and not an operator anymore. The free particle density matrix can be evaluated, as shown in Appendix 1 and ρ becomes:

$$\rho = \left(\frac{mP}{2\pi\beta\hbar^2}\right)^{\frac{3}{2}} e^{-\frac{\beta}{P}V_{loc}(\vec{r}_n)} \times \int d\vec{r}' \left\langle \vec{r}_n \middle| e^{-\frac{\beta}{P}V_{NL}} \middle| \vec{r}' \right\rangle e^{-\left(\frac{mp}{2\beta\hbar^2}\right)(\Delta\vec{r})^2} \tag{2.88}$$

where the sampling property of the delta function has been used, i.e.:

$$\int \delta(\vec{r} - \vec{r}_n) f(\vec{r}) d\vec{r} = f(\vec{r}_n) \tag{2.89}$$

and the difference $\Delta\vec{r}$ between the two vectors is given by:

$$\Delta(\vec{r}) = \vec{r}' - \vec{r}_{n+1} \tag{2.90}$$

2.2.5.3 Evaluation of the Non-local Density Matrix Elements

The non-local density matrix elements are those that involve V_{NL} in the exponential as follows:

$$G_{NL} = \left\langle \vec{r}_n \middle| e^{-\frac{\beta}{P}V_{NL}} \middle| \vec{r}' \right\rangle \tag{2.91}$$

Again, using the power series expansion of the exponential function we will have

$$G_{NL} = \langle \vec{r}_n | 1 | \vec{r}' \rangle + \left(-\frac{\beta}{P}\right) \langle \vec{r}_n | V_{NL} | \vec{r}' \rangle + \frac{1}{2!}\left(-\frac{\beta}{P}\right)^2 \langle \vec{r}_n | V_{NL} \cdot V_{NL} | \vec{r}' \rangle + O\left(V_{NL}^3\right) \tag{2.92}$$

where the higher orders are not neglected here. The zero order term (the first term on the RHS of (2.92) is:

$$\langle \vec{r}_n | 1 | \vec{r}' \rangle = \delta(\vec{r}' - \vec{r}_n) = \frac{\delta(r_n - r')}{r'^2} \delta(\cos\theta_n - \cos\theta') \delta(\phi_n - \phi') \tag{2.93}$$

where the set of spherical harmonics is a complete orthonormal basis set that could be expressed in the following closure relation:

$$\delta(\cos\theta_n - \theta')\delta(\phi_n - \phi') = \sum_{l=0}^{\infty} \sum_{m=-l}^{m=l} Y_{lm}^*(\theta_n, \phi_n) Y_{lm}(\theta', \phi') \tag{2.94}$$

The above relation will allow the zero order term (2.93) to take the form

$$\langle \vec{r}_n | \vec{r}' \rangle = \frac{\delta(r_n - r')}{r'^2} \sum_{l,m} Y_{lm}^*(\theta_n, \phi_n) Y_{lm}(\theta', \phi') \tag{2.95}$$

where the indices l, m stand for a double sum as appears in (2.94).

The first order term is given by [40]:

$$\langle \vec{r}_n | V_{NL} | \vec{r}' \rangle = \sum_l \sum_m Y^*_{lm}(\theta_n, \phi_n) V_l(r_n, r') Y_{lm}(\theta', \phi') \tag{2.96}$$

The evaluation of the second order term involves more bookkeeping. We first introduce the closure relation:

$$\langle \vec{r}_n | V_{NL} \cdot V_{NL} | \vec{r}' \rangle = \left\langle \vec{r}_n | V_{NL} \int d\vec{r} | \vec{r} \right\rangle \langle \vec{r} | V_{NL} | \vec{r}' \rangle \tag{2.97}$$

Taking the integral outside and using (2.96), the second term (2nd) will become:

$$\text{2nd} = \int d\vec{r} \sum_{l,m} Y^*_{lm}(\theta_n, \phi_n) V_l(r_n, r) Y_{lm}(\theta, \phi) \times \sum_{L,M} Y^*_{LM}(\theta, \phi) V_L(r, r') Y_{LM}(\theta', \phi') \tag{2.98}$$

where upon using the orthonormality relation:

$$\int_0^{2\pi} d\phi \int_0^{\pi} \sin\theta d\theta Y^*_{lm}(\theta, \phi) Y_{LM}(\theta, \phi) = \delta_{lL} \delta_{mM} \tag{2.99}$$

Equation (2.98) takes the form:

$$\text{2nd} = \int r^2 dr \sum_{l,m} Y^*_{lm}(\theta_n, \phi_n) V_l(r_n, r') Y_{lm}(\theta', \phi') V_l(r, r') \tag{2.100}$$

Now if we choose the non-local potential in the form:

$$V_l(r_n, r) = \frac{V_l(r_n)}{r_n^2} \delta(r_n - r) \tag{2.101}$$

where there is a different $V_{l(rn)}$ for each value of l. Using this non-local potential in (2.100) we have:

$$\text{2nd} = \int r^2 d\vec{r} \sum_{l,m} Y^*_{lm}(\theta_n, \phi_n) Y_{lm}(\theta', \phi') \frac{V_l(r_n)}{r_n^2} \delta(r_n - r) \times \frac{V_l(r)}{r^2} \delta(r - r') \tag{2.102}$$

Using the sampling property of the delta function, as introduced in the (2.89), we can reduce (2.102) to look like:

$$2\text{nd} = \sum\nolimits_{l,m} Y^*_{lm}(\theta_n, \phi_n)\, V_l(r_n, r') Y_{lm}(\theta', \phi') V_l(r') \tag{2.103}$$

The higher order terms could be evaluated in a similar fashion (as done above). For example, the third order term will be:

$$\left\langle \vec{r}_n \middle| V^3_{LN} \middle| \vec{r}' \right\rangle = \sum\nolimits_{l,m} Y^*_{lm}(\theta_n, \phi_n) V_l(r_n, r') Y_{lm}(\theta', \phi') V^2_l(r') \tag{2.104}$$

The expressions for the various terms are substituted back in expansion of G_{NL} to get:

$$\begin{aligned} G_{NL} = & \sum_{l,m} Y^*_{lm}(\theta_n, \phi_n) Y_{lm}(\theta', \phi') \frac{\delta(r_n - r')}{r'^2} + \left(-\frac{\beta}{P}\right) \sum_{l,m} Y^*_{lm}(\theta_n, \phi_n) Y_{lm}(\theta', \phi') V_l(r_n, r') \\ & + \frac{\left(-\frac{\beta}{P}\right)^2}{2!} \sum_{l,m} Y^*_{lm}(\theta_n, \phi_n) Y_{lm}(\theta', \phi') V_l(r') \\ & + \text{higher order terms} \end{aligned} \tag{2.105}$$

To simplify the above equation, we use the form of the non-local pseudo-potential as defined in (2.102) to obtain:

$$G_{NL} = \sum\nolimits_{l,m} Y^*_{lm}(\theta_n, \phi_n) Y_{lm}(\theta', \phi') \frac{\delta(r_n - r')}{r'^2} e^{-\frac{\beta}{P} V_l(r')} \tag{2.106}$$

Let us expand the square of the differences between the two vectors (2.90) as follows:

$$(\vec{r}' - \vec{r}_{n+1})^2 = r'^2 + r^2_{n+1} - 2\vec{r}' \cdot \vec{r}_{n+1} \cos\alpha \tag{2.107}$$

where the cosine of the angle between the two vectors is given by:

$$\cos\alpha = \cos\theta' \cos\theta_{n+1} + \sin\theta' \sin\theta_{n+1} \cos\left(\phi' - \phi_{n+1}\right) \tag{2.108}$$

Using (2.107) we can write the following:

$$e^{-\beta C(\Delta\vec{r})^2} = e^{-\beta C\left(r'^2 + r^2_{n+1}\right)} e^{2\beta C r' r_{n+1} \cos\alpha} \tag{2.109}$$

where

$$C = \frac{Pm}{2\beta^2\hbar^2} \tag{2.110}$$

In order to get the expression for the propagator ρ we will need to use the identity [62]:

$$e^{\xi\cos\alpha} = \left(\frac{\pi}{2\xi}\right)^{\frac{1}{2}} \sum_{l=0}^{\infty} (2l+1) P_l(\cos\alpha) I_{l+\frac{1}{2}}(\xi) \tag{2.111}$$

where $P_l(\cos a)$ is the familiar Legendre polynomial, and a is the angle between the two position vectors. We also need to write the Legendre polynomials in a form that is helpful in simplifying the expression for G_{NL}. Fortunately, the addition theorem of spherical harmonics would help in our case. We can write:

$$P_l(\cos\alpha) = \frac{4\pi}{2l+1} \sum_{m=-l}^{l} Y_{lm}^*(\theta_1, \phi_1) Y_{lm}(\theta_2, \phi_2) \tag{2.112}$$

which when substituted in (2.111) will yield:

$$e^{\xi\cos\alpha} = 4\pi \left(\frac{\pi}{2\xi}\right)^{\frac{1}{2}} \sum_{l,m} Y_{lm}^*(\theta_1, \phi_1) Y_{lm}(\theta_2, \phi_2) I_{l+\frac{1}{2}}(\xi) \tag{2.113}$$

Now using these relations and (2.105) we can write the short time (if P is large) propagator in the following manner:

$$\begin{aligned} \rho = {} & 4\pi \left(\frac{\beta C}{\pi}\right)^{\frac{3}{2}} e^{-\frac{\beta}{P} V_{loc}(\vec{r}_n)} \int r'^2 dr' \int \sin\theta' d\theta' \int d\phi' \\ & \times \sum_{l,m} Y_{lm}^*(\theta_n, \phi_n) Y_{lm}(\theta', \phi') \frac{\delta(r_n - r')}{r'^2} e^{-\frac{\beta}{P} V_l(r')} \\ & \times \left[\frac{\pi}{4\beta C r' r_{n+1}}\right]^{\frac{1}{2}} \sum_{L,M} Y_{LM}\left(\theta_{n+1}, \phi_{n+1}\right) Y_{LM}^*(\theta', \phi') \\ & \times I_{l+\frac{1}{2}}(2\beta C r' r_{n+1}) e^{-\beta C\left(r'^2 + r_{n+1}^2\right)} \end{aligned} \tag{2.114}$$

From the equation above it is easy then to get the final expression for ρ which takes the form:

$$\begin{aligned} \rho = {} & 4\pi \left(\frac{\beta C}{\pi}\right)^{\frac{3}{2}} \left[\frac{\pi}{4\beta C r_n r_{n+1}}\right]^{\frac{1}{2}} \sum_{l,m} Y_{lm}^*(\theta_n, \phi_n) Y_{lm}\left(\theta_{n+1}, \phi_{n+1}\right) \\ & \times e^{-\frac{\beta}{P}[V_{loc}(r_n) + V_l(r_n)]} I_{l+\frac{1}{2}}(2\beta C r_n r_{n+1}) e^{-\beta C\left(r_n^2 + r_{n+1}^2\right)} \end{aligned} \tag{2.115}$$

where V_{loc} acts locally in radial coordinates.

We are interested in developing a classical model of an electron in a non-local pseudo-potential. To that effect, we determine the path integral representation of the partition function.

2.2.5.4 One Electron Partition Function

The partition function is given by:

$$Z = \int d\vec{r}_1 \left\langle \vec{r}_1 \middle| e^{-\beta H_{op}} \middle| \vec{r}_1 \right\rangle \approx \lim_{P\to\infty} \int d\vec{r}_1 \left\langle \vec{r}_1 \middle| \left(e^{-\frac{\beta H_{op}}{P}} \right)^P \middle| \vec{r}_1 \right\rangle \tag{2.116}$$

Upon introducing (P-1) intermediate states in Z we will get as seen in previous sections:

$$Z \approx \int d\vec{r}_1\, d\vec{r}_2 \ldots d\vec{r}_P \left\langle \vec{r}_1 \middle| e^{-\epsilon H_{op}} \middle| \vec{r} \right\rangle_2 \times \left\langle \vec{r}_2 \middle| e^{-\epsilon H_{op}} \middle| \vec{r}_3 \right\rangle \ldots \left\langle \vec{r}_P \middle| e^{-\epsilon H_{op}} \middle| \vec{r} \right\rangle_1 \tag{2.117}$$

where we have used the periodic boundary condition (cyclic condition): $\vec{r}_{P+1} = \vec{r}_1$. Our one electron system is now transformed to look like a polymeric necklace consisting of P beads.

We, write the partition function in more compact form:

$$Z \approx \int \prod_{n=1}^{P} * d\vec{r}_n \left\langle \vec{r}_n \middle| e^{-\frac{\beta}{P} H_{op}} \middle| \vec{r}_{n+1} \right\rangle \tag{2.118}$$

and using (2.115) we can write the partition function as follows:

$$\begin{aligned} Z \approx & \left(\frac{\beta C}{\pi} \right)^{\frac{3P}{2}} \int d\vec{r}_1 d\vec{r}_2 \ldots d\vec{r}_p \prod_{n=1}^{P} \left[e^{-\frac{\beta}{P} V_{loc}(r_n)} \right. \\ & \left. \times \left\{ \sum_{l=0}^{\infty} (2l+1)\, P_l(\cos\alpha)\, .\mathfrak{F}_{l+\frac{1}{2}}(\chi)\, .e^{-\frac{\beta}{P} V_l(r_n)} \right\} .e^{-\beta C (r_n - r_n + 1)^2} \right] \end{aligned} \tag{2.119}$$

In the preceding equation, we have defined a modified Bessel function:

$$\mathfrak{F}_{l+\frac{1}{2}}(\chi) = \left(\frac{\pi}{2\chi} \right)^{\frac{1}{2}} \cdot e^{-\chi} \cdot I_{l+\frac{1}{2}}(\chi)$$

where $\chi = 2\beta C r_n r_{n+1}$.

Upon inspecting (2.119) we see that the summation over the angular momenta runs up to infinity. This is inconvenient in a computational setting. Therefore, there must be some way that will help to truncate the sum, in order to reduce the

computational effort, without the loss of accuracy. The truncation is achieved by observing [63] that the local potential $V_l(r)$ is almost the same for all values of l that do not exist in the core region. This could be written as:

$$V_l = V_{l_{max}+1} \quad for \quad l > l_{max}$$

where l_{max} is the maximum value for l in the core region.

Then, (2.115) can be rewritten in the following fashion:

$$\rho = 4\pi\left(\frac{\beta C}{\pi}\right)^{\frac{3}{2}} e^{-\frac{\beta}{P}V_{loc}} \left\{ \sum_{m=-l}^{m=l} \sum_{l=0}^{l_{max}-1} Y^*_{lm}(\theta_n, \phi_n) Y_{lm}(\theta_{n+1}, \phi_{n+1}) \times \mathfrak{F}_{l+\frac{1}{2}}(2\beta C r_n r_{n+1}) e^{-\frac{\beta}{P}V_l(r_n)} \right.$$
$$\left. + \sum_{m=-l}^{m+l} \sum_{l=l_{max}}^{\infty} Y^*_{lm}(\theta_n, \phi_n) Y_{lm}(\theta_{n+1}, \phi_{n+1}) \times \mathfrak{F}_{l+\frac{1}{2}}(2\beta C r_n r_{n+1})] e^{-\frac{\beta}{P}V_l(r_n)} \right\} e^{-\beta C(r_n - r_{n+1})^2} \tag{2.120}$$

Upon further manipulation of the above equation we arrive at:

$$\rho = 4\pi\left(\frac{\beta C}{\pi}\right)^{\frac{3}{2}} e^{-\frac{\beta}{P}V_{loc}} e^{-\beta C(r_n - r_{n+1})^2} \left\{ \sum_{m=-1}^{m=1} \sum_{l=0}^{l_{max}-1} Y^*_{lm}(\theta_n, \phi_n) \right.$$
$$\left. \times Y_{lm}(\theta_{n+1}, \phi_{n+1}) \mathfrak{F}_{l+\frac{1}{2}}(2\beta C r_n r_{n+1}) \left[e^{-\frac{\beta}{P}V_l(r_n)} - e^{-\frac{\beta}{P}V_{l_{max}}} \right] \right\} \tag{2.121}$$
$$+ \left(\frac{\beta C}{\pi}\right)^{\frac{3}{2}} e^{-\frac{\beta}{P}[V_{loc} + V_{l_{max}}(r_n)]} e^{-\beta C(\vec{r}_n - \vec{r}_{n+1})^2}$$

Notice that in the case where the non-local potential is not significant, or not considered, this equation will lead to the density matrix of a particle in the local potential V_{loc}.

2.2.5.5 Model of a Valence Electron in a Sodium Atom

The sodium (Na) ion has the following electronic configuration:

$$Na^+ : 1s^2 2s^2 2p^6$$

which shows a single valence in an angular momentum state of $l = 0$.

The maximum l that exists in the core is one (the p state). Therefore, we have $l_{max} = 1$.

Using (2.121) we can find ρ for the sodium ion:

$$\rho_{Na^+} = \left(\frac{\beta C}{\pi}\right)^{\frac{3}{2}} e^{-\frac{\beta}{P}V_{loc}} e^{-\beta C(r_n - r_{n+1})^2} \left\{ P_0(\cos\alpha) \times \mathfrak{F}_{\frac{1}{2}}(2\beta C r_n r_{n+1}) \left[e^{-\frac{\beta}{P}V_0(r_n)} - e^{-\frac{\beta}{P}V_1(r_n)} \right] \right\} + \left(\frac{\beta C}{\pi}\right)^{\frac{3}{2}} e^{-\frac{\beta}{P}[V_{loc}(r_n)+V_1(r_n)]} e^{-\beta C(\vec{r}_n - \vec{r}_{n+1})^2} \tag{2.122}$$

where the zero order Legendre polynomial $P_0(\cos a) = 1$.

In order to proceed with the molecular dynamics simulation for our quantum system, we need an effective potential to describe the interactions of the nodes with each other, and with the sodium ion. As we did before, this effective potential could be achieved by writing the partition function as follows:

$$Z \approx \left(\frac{\beta C}{\pi}\right)^{\frac{3P}{2}} \int \left[\prod_{n=1}^{P} d\vec{r}_n\right] . e^{-\beta V_{eff}(\vec{r}_1, \cdots, \vec{r}_P)} \tag{2.123}$$

We proceed from here to find the partition function for Na^+, which is:

$$Z_{Na^+} \approx \left(\frac{\beta C}{\pi}\right)^{\frac{3P}{2}} \int d\vec{r}_1 \ldots d\vec{r}_P \exp\left\{ \sum_{n=1}^{P} \mathrm{Ln}\left[e^{-\frac{\beta}{P}V_{loc}} \times e^{-\beta C(r_n - r_{n+1})^2} \mathfrak{F}_{\frac{1}{2}}(2\beta C r_n r_{n+1}) \left\langle e^{-\frac{\beta}{P}V_0(r_n)} - e^{-\frac{\beta}{P}V_1(r_n)} \right\rangle + e^{-\frac{\beta}{P}[V_{loc}+V_1(r_n)]} e^{-\beta C(\vec{r}_n - \vec{r}_{n+1})^2} \right] \right\} \tag{2.124}$$

However, such an approximation is valid only if the terms in the <> of (2.124) is a positive number. We implement here an approach similar to the restricted path integral method described previously, whereby, we will limit the states to those with a positive <>. Notice that the above expression contains both an "s" and a "p" pseudopotentials. This will be enough to study the 3s and the 3p electronic behavior.

Upon inspecting the partition function we can see that the classical effective potential is readily obtained as:

$$V_{eff} = -\frac{1}{\beta}\sum_{n=1}^{P} \mathrm{Ln}\left\{ e^{-\frac{\beta}{P}V_{loc}} e^{-\beta C(r_n - r_{n+1})^2} \times \mathfrak{F}_{\frac{1}{2}}(2\beta C r_n r_{n+1}) \left[e^{-\frac{\beta}{P}V_0(r_n)} - e^{-\frac{\beta}{P}V_1(r_n)} \right] + e^{-\frac{\beta}{P}[V_{loc}+V_1(r_n)]} e^{-\beta C(\vec{r}_n - \vec{r}_{n+1})^2} \right\} \tag{2.125}$$

2.3 Path Integral Molecular Dynamics Simulation Method

In Sect. 2.2, we developed several classical models of quantum systems using the Path Integral formalism of quantum mechanics. Here, we integrate these classical models within the framework of the simulation of Molecular Dynamics (MD).

2.3.1 Molecular Dynamics Method

The simulation method of MD computes phase space trajectories of a system of interacting particles which individually obey the classical Newtonian equations of motion. Specifically, if a N-particle system is described by a classical Hamiltonian $H = \sum_i \frac{mv_i^2}{2} + \sum_{i>j} \phi(r_{ij})$, where v_i is a velocity of a particle i and $\phi(r_{ij})$ is a pairwise central potential between particles i and j separated by the distance $r_{ij} = |r_i - r_j|$, one can analytically specify the phase space trajectories, $(r_i(t), v_i(t))$ by solving the Newtonian equations of motion with certain initial conditions. In other words, we can find the time evolution trajectories by solving the following equations;

$$F_i(t) = m\frac{d^2 r_i}{dt^2}$$
$$v_i(t) = \frac{dr_i}{dt}$$

where F_i is the total force on the particle i, which is the sum of the forces due to all the other particles, j, in the system. The force on the particle i due to interactions with a particle j, $f_{ij}(t)$, can be obtained from $f_{ij} = -\nabla_j \phi_{ij}$. To solve the equations of motion in a numerical MD simulation, one discretizes in time the differential equations of motion. A variety of numerical schemes can then be used to integrate the equations of motion and march the system in time. For instance, the simplest method is based on straightforward finite differences. Since Alder and Wainwright [44] used MD for simulating an N -body system, this method has been applied to a large variety of systems [64–67]. In its most basic form the trajectories generated by the MD algorithm conserve energy (Hamiltonian), volume and number of particles, i.e. they are representative of the microcanonical ensemble in statistical mechanics.

The MD method developed by Alder and Wainwright is conceptually the simplest to illustrate the concept of MD simulation. To solve the equations of motion above, for each particle "i", we apply a finite difference scheme for the second-order differential equation:

$$\frac{d^2 r_i(t)}{dt^2} = \frac{1}{(\Delta t)^2}\{r_i(t+\Delta t) - 2r_i(t) + r_i(t-\Delta t)\}$$

where Δt is the simulation time step. The discretized equations of motion become:

$$r_i(t+\Delta t) = 2r_i(t) - r_i(t-\Delta t) + \frac{F_i(r_i)}{m}(\Delta t)^2, \quad for\ i = 1, 2, \ldots, N$$

With this equation one can obtain a new position at time $t + \Delta t$ from the positions at two consecutively preceding time steps and the force acting at time t. Provided a set of initial conditions $r_i\,(t = 0)$ and $r_i(\Delta t)$ we can obtain the time evolution of the position of every particle in the system. We can calculate the velocity of the ith particle with the Euler backward scheme according to $v_i(t) = \frac{1}{\Delta t}\{r_i(t) - r_i(t-\Delta t)\}$, $for\ i = 1, 2, \ldots, N$, thus obtained the trajectory in phase space of the collection of N particles.

This simplest of the MD simulation algorithm is also called the "leap-frog" integration algorithm used by Verlet [68]. Note that the trajectories produced by this method conserve the Hamiltonian and therefore the energy of the system. In this form the MD method is simulating only a small system of N particles. The number of particles is limited by the availability of computational resources. It is common to use Periodic Boundary Conditions (PBC) to mimic the behavior of a small part of a material embedded into an infinite system. Let us consider a N-particle system within a physical volume, the basic cell or the MD cell. We assume for the sake of simplicity that the cell is a rectangular box-like cell with edge lengths, $L = \left(L_x, L_y, L_z\right)$, and a volume $\Omega = L_x L_y L_z$. To retain a constant number density and reduce surface effects associated with the small size of the simulation cell we impose periodic boundary conditions (PBC) [69]. With PBC, the MD simulation cell is repeated periodically in all directions of space. The repeated cells are called image cells. There are 26 image cells in three dimensions. Due to the periodicity imposed by these boundary conditions, we may write a physical quantity $A(r)$ in the MD-simulation cell in the form of a periodic function:

$$A(r) = A(r + n \cdot L)$$

where $n = \left(n_x, n_y, n_z\right)$ whose components are integers, r is confined within the cell, i.e. $|r| < L$. PBC means that if a particle crosses a boundary of the cell as it evolves along its trajectory, it re-enters through the opposite side of the cell at the same instant. PBC conserve number density. Furthermore, with PBC the potential energy is represented by

$$V = \sum_{i<j} \phi\left(r_{ij}\right) + \sum_{n} \sum_{i<j} \phi\left(\left|r_i - r_j + n \cdot L\right|\right)$$

where $\phi\left(r_{ij}\right)$ is the pair potential, $r_{ij} = \left|r_i - r_j\right|$, and r_i and r_j are restricted with in the MD-cell. In order to avoid the infinite summation in the last term one typically

introduces a cut-off range (r_c) for the potential [64, 70]. A particle in the basic cell interacts only with each of the N - 1 other particles in the MD-cell or its image cells. Effectively we may cut off the potential at a range

$$r_c < L/2$$

With this cut-off range r_c, we rewrite theotential energy as

$$V = \sum_{i(<j)} \sum_{j \epsilon MDcell} \phi\left(r_{ij}(<r_c)\right)$$

Although we can effectively simulate a system with MD method with PBC and a truncated potential, such a truncation may have a detrimental effect on systems with the long range interactions such as systems of particles interacting via the Coulomb potential. This issue will be discussed specifically in subsequent subsections as our quantum systems composed of charged particles (electrons).

The simulation of the model quantum systems developed in Sect. 2.2 requires that we define some classical Hamiltonian from which we will derive dynamical equations of motion to sample the states of the quantum system.

2.3.2 *Molecular Dynamics Classical Hamiltonian for N-Electron Plasma*

In this section, we develop a classical Hamiltonian compatible with MD simulations based on the restricted path-integral representation of quantum particles. Once we set up the Hamiltonian of a system, we can calculate the phase space trajectories of all particles with the MD method.

Let the system consist of N_{el} unpolarized electrons i.e. an electron plasma. The number of spin-up electrons is the same as the number of spin-down electrons. If there is no external force, the total potential for the system is the sum of the electron-electron Coulomb potential and the effective potentials given by (2.71, 2.73). By considering a kinetic energy term for the N_{el} electrons, a classical Hamiltonian can be defined as:

$$H = \sum_{i=1}^{N_{el}} \sum_{k=1}^{P} \frac{1}{2} m^* \left(\dot{r}_i^{(k)}\right)^2 + \sum_{k=1}^{P} \sum_{i(>j)}^{N_{el}} \sum_{j=1}^{N_{el}} \frac{e^2/P}{\left|r_i^{(k)} - r_j^{(k)}\right|} + \sum_{i=1}^{N_{el}} \sum_{k=1}^{P} {}' \frac{m_e P}{2\beta\hbar^2} \left(r_i^{(k)} - r_i^{(k+1)}\right)^2$$
$$- \frac{1}{\beta} \sum_{s=up}^{down} \sum_{k=1}^{P} \sum_{l=1}^{P} \frac{1}{P_s^*} \ln\left(\det\left(E^{(k,l)}\right)\right) \theta_{kls}^{+} \tag{2.126}$$

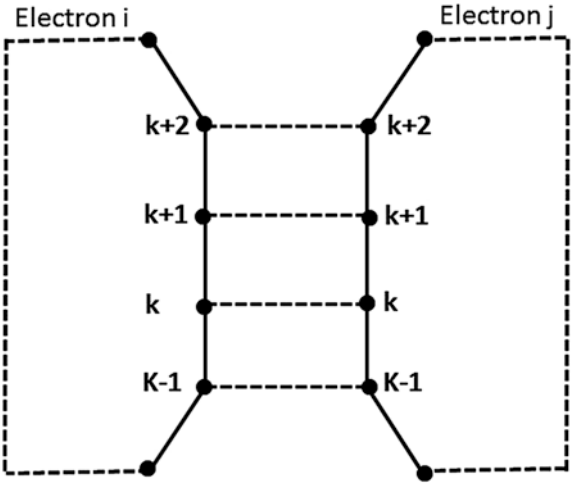

Fig. 2.6 Coulomb interaction between two electrons i and j at different imaginary times. Note that the electron-electron Coulomb interactions occur between all the pairs of beads in the discretized necklaces i.e. are not retarded in imaginary time

where

$$P_s^* = \sum_{k=1}^{P} \sum_{l=1}^{P} \theta_{kls}^{+}$$

In (2.126), m^* is an arbitrary mass of a bead (we chose $m^* = 1$ amu) [5] used to define an artificial kinetic energy for the quantum states in order to explore the effective potential surface, V_{eff}, and m_e is an electron mass. We recall that the subscript and the superscript on the position are for labeling an electron and a bead in this electron, respectively. The second term in the above equation accounts for the electron-electron Coulomb potential energy. The forces derived from the non-local exchange potential, the last term of (2.126), are calculated as means over the restricted paths with positive determinants. Therefore, an effective force calculation requires a satisfactory sample of such paths. Since the exchange potential offers a barrier to paths with negative determinants, it biases the sampling of phase space toward configurations with positive determinants. Although many configurations with negative determinants may exist and evolve, they do not contribute to the exchange forces.

In the previous Hamiltonian, the Coulomb interactions between electrons are long range and we use the Ewald summation method (see subsequent sections) to calculate their contributions to the energy and forces. To satisfy charge neutrality, the electron plasma system needs to evolve in a uniform positively charged

background. As we have seen, in the path-integral MD, each electron is considered as a necklace of P beads. The Coulomb potential in (2.126) can be written as

$$V_2 = \sum_{k=1}^{P} \frac{1}{P} \sum_{i(>j)}^{N} \sum_{j=1}^{N-1} \frac{e^2}{r_{ij}^{(k)}}$$

where $r_{ij}^{(k)} = \left| r_i^{(k)} - r_j^{(k)} \right|$ is the distance between the bead (k) of the ith electron necklace and the bead (k) of the jth electron necklace. We illustrate the Coulomb interactions between beads in Fig. 2.6.

2.3.3 Molecular Dynamics Hamiltonian for N-Alkali Atom Metal

To extend the restricted PIMD method to a simple metal such as an alkali metal system (system composed of N monovalent alkali ions and N electrons), we simply add the potential terms related to the classical ionic degrees of freedom to the Hamiltonian of the electron plasma system. For a system containing N_{el} un-polarized electrons and N_{ion} ions, the Hamiltonian becomes:

$$\begin{aligned} H = H_{ele} + \sum_{I=1}^{N_{ion}} \frac{1}{2} M_I \dot{R}_i^2 + \sum_{I>(J)}^{N_{ion}} \sum_{J}^{N_{ion}-1} \phi(R_{IJ}) \\ + \sum_{i=1}^{N_{el}} \sum_{I=1}^{N_{ion}} \frac{1}{P} V_{pseudo}\left(R_I - r_i^{(k)}\right) \end{aligned} \tag{2.127}$$

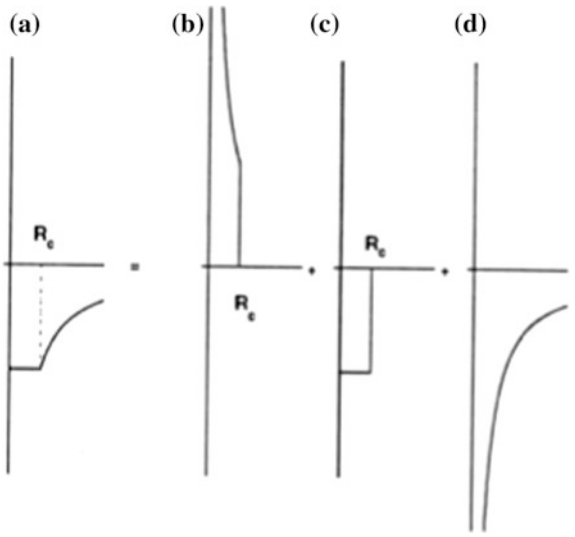

Fig. 2.7 Empty core pseudo-potential model for the alkali ion-electron interaction. The sum of the potentials **b**, **c**, and **d** is equal to (**a**)

where M_I is the ion mass and $\boldsymbol{R}$ is the position of an ion. In (2.127), H_{ele} is the Hamiltonian of the electron plasma described in (2.126) and the second term is the kinetic energy of the ions. $\phi(R_{IJ})$ and V_{pseudo} are the ion-ion potential and the ion-electron pseudo-potential, respectively. This Hamiltonian assumes that the electron-ion pseudo-potential is local. To that effect, the electron-ion interaction is modeled by a local empty core pseudopotential model. In this model, we assume that a positive ion is a conducting sphere with radius R_c and its total charge is $+e$. The local pseudo-potential describing interactions between the Ith ion and the kth bead of an electron j is defined as

$$V_{pseudo}\left(R_I - r_i^{(k)}\right) = \varphi\left(r_{Ij}^{(k)}\right) = \begin{cases} -e/R_c, & if\ r_{Ij}^{(k)} \le R_c \\ -e/r_{Ij}^{(k)}, & if\ r_{Ij}^{(k)} > R_c \end{cases} \tag{2.128}$$

where $r_{IJ}^{(k)} = \left|R_I - r_j^{(k)}\right|$ is the distance of the kth bead of the electron j measured from the center of the Ith ion. Figure (2.7) illustrate this pseudo-potential. In the figure the pseudo-potential (a) is represented by the sum of a usual coulombic potential (d), a repulsive $\frac{1}{r}$ potential inside the core (b) and a constant potential inside the core (c). The total potential on the Ith ion then becomes

$$\varphi(r_I) = \sum_{k=1}^{P} \frac{(-e)}{P} \left\{ \sum_{j=1}^{N_e} \left(\frac{1}{R_c} - \frac{1}{r_{Ij}^{(k)}} \right) \theta_{Ij} + \sum_{j=1}^{N_e} \frac{1}{r_{Ij}^{(k)}} \right\}$$

where θ_{Ij} is equal to 1 for $r_{Ij} \le R_c$, otherwise 0.

The electron-ion potential energy in our Hamiltonian can be calculated by $V_2^{e-ion} = \sum_I e\varphi(r_I)$.

A variety of inter-ionic potentials can be used for ϕ. In this chapter, the cation-cation pair potential we use is of the Bohn-Mayer (BM) type and was proposed by Fumi and Tosi [71] for alkali halides. This potential includes a long range Coulomb potential term and a short range repulsive term:

$$V_2^{ion-ion} = \sum\nolimits_{I>J}^{N} \sum\nolimits_{J=1}^{N-1} \frac{e^2}{r_{IJ}} + A_{IJ} e^{-r_{IJ}/\rho_{IJ}} \tag{2.129}$$

where $r_{IJ} = |r_I - r_J|$ is the distance between an ion pair I and J, and A_{IJ} and ρ_{IJ} are parameters of the potential.

2.3.4 Molecular Dynamics Hamiltonian for a Single Alkali Atom with Non-local Pseudo-potential

The single electron classical MD Hamiltonian is:

$$H = \sum_{k=1}^{P} \frac{1}{2} m^* \left(\dot{r}^{(k)} \right)^2 + V_{eff} \tag{2.130}$$

where the effective potential is that given by (2.125) which we repeat below for completeness:

$$V_{eff} = -\frac{1}{\beta} \sum_{n=1}^{P} Ln \left\{ e^{-\frac{\beta}{P} V_{loc}} e^{-\beta C (r_n - r_{n+1})^2} \times \Im_{\frac{1}{2}} (2\beta C r_n r_{n+1}) \left[e^{-\frac{\beta}{P} V_0(r_n)} - e^{-\frac{\beta}{P} V_1(r_n)} \right] + e^{-\frac{\beta}{P}[V_{loc} + V_1(r_n)]} e^{-\beta C (\vec{r}_n - \vec{r}_{n+1})^2} \right\}$$

Again, m^* is the arbitrary mass of a bead "k" in the single electron necklace. We chose $m^* = 1$ amu. It is convenient to rewrite this effective potential in a more revealing form:

$$V_{eff} = \sum_{n=1}^{P} C (\vec{r}_n - \vec{r}_{n+1})^2 + \frac{V_1(r_n)}{P} - \frac{1}{\beta} Ln \left\{ 1 + \Im_{\frac{1}{2}} (2\beta C r_n r_{n+1}) e^{2\beta C r_n r_{n+!} (1 - \cos \alpha)} \times \left[e^{-\frac{\beta}{P}(V_0 - V_1)} - 1 \right] \right\}$$

where we have used the identity:

$$(r_n - r_{n+1})^2 = r_n^2 + r_{n+1}^2 - 2 r_n r_{n+1} + 2 r_n r_{n+1} \cos \alpha - 2 r_n r_{n+1} \cos \alpha$$

2.3.5 *Periodic Boundary Conditions*

In this section, we address effects associated with periodic boundary conditions (PBC) that are specific the simulation of the quantum systems described by the Hamiltonians given by (2.126) (electron plasma) and (2.127) (Alkali metal). These effects are associated with the electric charge of the particles and their Coulomb interactions, the representation of quantum particles by a closed necklace, and the non-local exchange potential associated with the indistinguishability of the quantum particles.

2.3.5.1 Ewald Summation Method

As we have seen MD simulations are often conducted with PBC. That is the system of interest is composed of a finite size simulation cell of specified shape and volume containing the particles of interest (here our electron plasma composed of N-electrons). This simulation cell is repeated periodically in all directions in space to mimic the behavior of a small element of material embedded into an infinite

system. If L is the length of the edge of a cubic simulation cell, then, the range of an interaction between two particles (here, electron-electron Coulomb interaction) is restricted to no more than half the size of the cell (i.e. $r_c < 1/2L$). This restriction is imposed to avoid unphysical interactions such as one electron interacting with one of its images or with another electron and simultaneously with an image of that same electron. Since we may be limited computationally to simulating a small number of electrons in a small simulation cell, we may neglect a critical amount of long range interactions. We use the Ewald summation method [72], to accurately account for the contribution of long range interactions to the Coulomb energy in the Hamiltonian of (2.126) and (2.127). Following the method, we consider a lattice made up of charged particles with positive and/or negative charges and we superpose onto these point charges spherically symmetrical Gaussian distributions of charges (while maintaining charge neutrality with another set of Gaussian distributions with opposite charge). The Gaussian charge distributions are proportional to $e^{-\eta r^2}$ where r is the radial distance from the center of the Gaussian. The parameter η is called the Ewald parameter. The calculation procedure has two distinct but related parts. One is computing the potential from a structure with Gaussian charge distributions located at the positions of our charged particles. The other one is the potential of our lattice of point charges with an additional Gaussian charge distribution of opposite sign superposed upon the point charges. η is chosen such that both potentials at some reference points converge rapidly.

The original Ewald summation method has been further developed by Nijboer et al. [73]. They generalized the summation method to interactions having form of $1/r^n$. This generalized Ewald summation method has been used in the Monte Carlo simulation of the classical one component plasma by Brush et al. [74] and of the fermion one component plasma by Ceperley [19]. Brush et al. showed that the Coulomb potential at some position within a cubic cell with PBC containing an electron plasma and a charge neutralizing positive background can be written as

$$\begin{aligned}\varphi(r_i) = \frac{4\pi}{\Omega}\sum_{k(\neq 0)}^{\infty}\frac{e^{-k^2/4\eta}}{k^2}\sum_{j(\neq i)}^{N} q_j e^{-ik\cdot r_{ij}} - 2q_i\sqrt{\frac{\eta}{\pi}} \\ + \sum_{j(\neq i)}^{N}\frac{q_j}{r_{ij}}erfc\left(\eta r_{ij}\right) - \frac{q_i\pi(N-1)}{\Omega\eta^2}\end{aligned} \tag{2.131}$$

where $q_i = -e$ for all $i, \Omega = L^3$ is the volume of a cell, k is a reciprocal space vector, and $r_{ij} = \left|r_i - r_j\right|$. The wave vector can be written in terms of an integer vector $n = \left(n_x n_y n_z\right)$ as $k = \frac{2\pi}{L}n$. The usual complementary error function, *erfc*, is defined as

$$\mathrm{erfc}(x) = 1 - \mathrm{erf}(x) = 1 - \frac{2}{\sqrt{\pi}} \int_0^x e^{-y^2} dy$$

The last term in (2.131) is the contribution of the positive background. In the path- Integral MD Hamiltonian, we may write the potential energy corresponding to electron-electron Coulomb interactions as

$$V_2 = \sum\nolimits_{k=1}^{P} \frac{1}{P} \frac{1}{2} \sum\nolimits_{i=1}^{N} q_i \varphi\left(r_i^{(k)}\right) \tag{2.132}$$

where the summations are performed over all beads of the electrons in the simulation cell. Specifically, the electron-electron Coulomb potential energy becomes:

$$\begin{aligned} V_2 = \sum_{k=1}^{P} \frac{1}{P} \Bigg\{ & \frac{2}{\pi L} \sum_{i(\neq j)}^{N} \sum_{j=1}^{N} \sum_{n>0}^{\infty} \frac{q_i q_j}{|n|^2} \exp\left(-\frac{\pi^2 |n|^2}{\eta^2 L^2}\right) \cos\left(\frac{2\pi}{L} n \cdot r_{ij}^{(k)}\right) \\ & - \sum_{i=1}^{N} q_i^2 \frac{\eta}{\sqrt{\pi}} + \frac{1}{2} \sum_{i(\neq j)}^{N} \sum_{j=1}^{N} \frac{q_i q_j}{r_{ij}^{(k)}} erfc\left(\eta r_{ij}^{(k)}\right) - \sum_{i=1}^{N} q_i^2 \frac{\pi (N-1)}{\Omega \eta^2} \end{aligned} \tag{2.133}$$

If we use the relation $r_{ij}^{(k)} = r_i^{(k)} - r_j^{(k)}$, the first term in the right-hand side of (2.133) can be made to take a considerably more efficient form [75]

$$\begin{aligned} V_2^{1st-term} = \sum_{i=1}^{P} \frac{1}{P} \frac{1}{\pi L} \sum_{n>0}^{\infty} \frac{1}{|n|^2} \exp\left(-\frac{\pi^2 |n|^2}{\eta^2 L^2}\right) \Bigg\{ & \left[\sum_{i=1}^{N} q_i \cos\left(\frac{2\pi}{L} n \cdot r_i^{(k)}\right)\right]^2 \\ & + \left[\sum_{i=1}^{N} q_i \sin\left(\frac{2\pi}{L} n \cdot r_i^{(k)}\right)\right]^2 \Bigg\} \end{aligned} \tag{2.134}$$

The double sum over i and j in (2.133) converts to two single sums. That is, $N^2 P$ calculations are reduced to $\sim NP$ calculations.

In many cases reported in this chapter, the long-range Coulomb potentials will be replaced by a short-range screened potential of the form $(1/r)erfc(\eta r)$ based on the real space part of the potential given by (2.133). With a well-chosen value for the parameter, η, the reciprocal-space part of the Ewald sum is negligible compared to the real-space part [76].

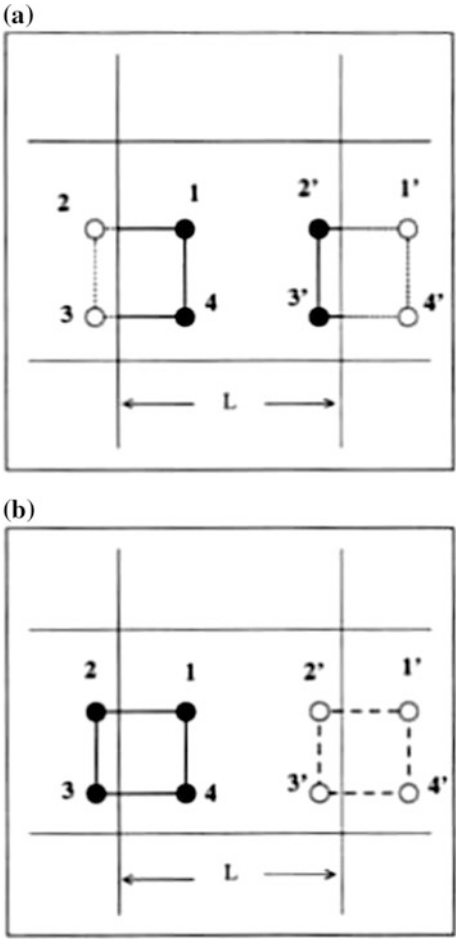

Fig. 2.8 Beads of an electron necklace in the simulation cell and the image cells. Beads (*1* and *4*) are in the simulation cell and beads (*2* and *3*) which belong to the same necklace are in the image cell. In standard MD methods, we use the *filled circles* in (**a**) to calculate distances between beads. In **b**, the necklace is reconstructed after translating beads (*2* and *3*) by *L* from the *left*

2.3.5.2 Maintaining the Continuity of the Necklace Representation of Quantum Particles

The electron necklaces does not always remain confined in the basic simulation cell but may spreads over its neighboring image cells. When some bead of an electron necklace expands beyond the simulation cell and into an image cell, the continuity of the necklace is broken by the traditional way of imposing PBC. Let us assume that two beads belonging to the same necklace of an electron i are at $r_i^{(k)}$ and $r_i^{(k+1)}$ within the simulation cell. Let us suppose that bead $(k + 1)$ moves outside the simulation cell in the direction (x, 0, 0), then its coordinates within the simulation cell becomes $r_i^{(k+1)} - (L_1, 0, 0)$, where L_x is the length of the cell in the direction (x). PBC has broken the continuity of the necklace. We cannot calculate the distance between beads k and $k + 1$ by standard MD methods. In Fig. 2.8 we show the

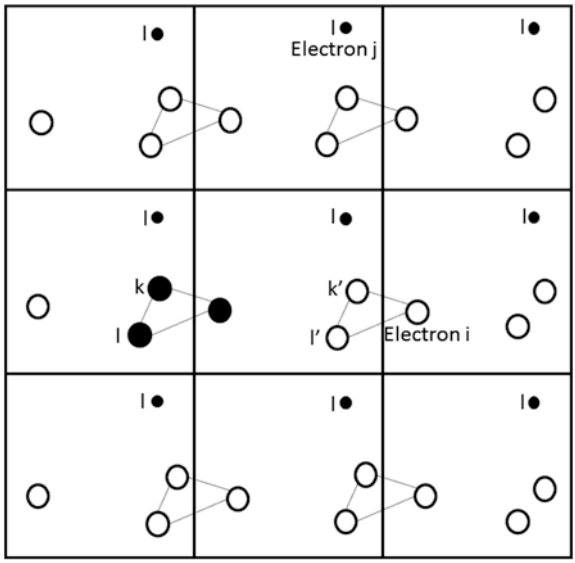

Fig. 2.9 Simulation cell with PBC containing two electrons. The *smaller filled dots* represent the *l*th beads of the electron *j* and the *larger circles* represent the beads of the electron *i*

effect of PBC on the continuity of the necklace for $P = 4$. We assumed that beads (1, 2, 3, and 4) make a complete necklace.

If we calculate the distance between the beads within the simulation cell, for instance between the filled circles in (a) of the figure, the calculated distance between 1 and 2′ is not equal to the true physical distance between beads 1 and 2. A MD algorithm for necklace representations of quantum particles, requires bookkeeping of the position of a bead within the simulation cell but also of the address of the image cell where the bead is actually located in order to allow for the reconstructing of the complete necklace. In Fig. 2.8b, we reconstruct the necklace by translating beads 2 and 3 from 2′ and 3′ by $-L$, respectively.

Practically, in the MD simulation with electron necklaces, the position of the *k*th bead of the *i*th electron is denoted by $r_i^{(k)}$ and $X_i^{(k)}$, where $r_i^{(k)}$ is always defined inside the simulation cell and $X_i^{(k)}$ contains the information necessary for reconstructing the integral necklace of the electron. In other words, a set $\left\{r_i^{(k)} + X_i^{(k)}; k = 1, 2, .., P\right\}$ represents a complete closed necklace of the electron *i*. When we measure the distance between two beads of the same electron *i*, we calculate

$$\left|\left(r_i^{(k)} + X_i^{(k)}\right) - \left(r_i^{(l)} + X_i^{(l)}\right)\right|$$

However if one asks about the distance between two beads in different electrons with PBC, it does not become an obvious question. The difficulty arises because there is no reference point for the distance measurement. In Fig. 2.9, we show the

beads of two electrons i and j, and their images. The larger circles represent the beads of the electron i and the smaller filled dots the bead l for the electron j. If we define a reference point as $r_i^{(k)} + X_i^{(k)} - X_i^{(l)}$, then we can easily calculate the distance from this point to the ith bead of any other electrons which can be in any of 27 cells (simulation cell plus image cells). For example, if we want to measure the minimum distance between the kth bead of the ith electron and the lth bead of the jth electron in any cell, we have to calculate all the distances

$$\left(r_i^{(k)} + X_i^{(k)} - X_i^{(l)}\right) - \left(r_j^{(l)} + n_{cell}L\right)$$

where n_{cell} is an integer vector running over the simulation cell (0, 0, 0) and 26 image cells {(1, 0, 0), (0, 1, 0), etc.}.

2.3.5.3 Evaluation of $\det\left(E^{(p,q)}\right)$ of the Effective Exchange Potential with PBC

When we evaluate the exchange potential and eventually the exchange forces, we encounter some practical and fundamental difficulties. One of the difficulties is associated with the number of operations to calculate the quantum non-local exchange effective potential. Since PBC are imposed on the system, the size of the computational problem increases dramatically. For a simulation in 3-dimensional space, there are actually 26 image cells which are the exact copies of the simulation cell. In that case the effective potential for N'-*iso*-spin electron:

$$V_{eff} = V_{eff}^{harm} + V_{eff}^{exch} \tag{2.135}$$

where

$$V_{eff}^{harm} = \sum\nolimits_{i=1}^{P} \sum\nolimits_{k=1}^{N\prime} \frac{Pm}{2\hbar^2\beta^2} \left(r_k^{(i)} - r_k^{(i+1)}\right)^2 \tag{2.136}$$

and

$$V_{eff}^{exch} = -\frac{1}{\beta} \sum_{i=1}^{P} \sum_{j=1}^{P} \frac{1}{P^*} \ln\det\left(E^{(i,j)}\right) \theta_{ij}^{+} \tag{2.137}$$

V_{eff}^{harm} and V_{eff}^{exch} are the harmonic potential and the exchange potential functions, respectively. A matrix element $\left(E^{(p,q)}\right)_{ij}$ is given by

$$\left(E^{(p,q)}\right)_{ij} = \exp\left[-\frac{Pm}{2\beta\hbar^2}\left\{\left(r_i^{(p)} - r_j^{(q)}\right)^2 - \left(r_i^{(p)} - r_i^{(q)}\right)^2\right\}\right] \tag{2.138}$$

In the above equations, the subscripts and the superscripts are indexes for the Electrons necklaces and nodes of a given electron necklace, respectively. Because of PBC, exchange processes must include not only electrons in the simulation cell but electrons in all the image cells. We, therefore, have to expand the size of the matrix $E^{(p,q)}$to a size of$(27N' \times 27N')$instead of$(N' \times N')$. This make the efficient evaluation of the exchange potential and subsequently of the corresponding forces practically impossible. To find a good approximation, we expand the determinant in the following manner:

$$\begin{aligned}\det\left(E^{(p,q)}\right) = 1 - \sum_{cells}\sum_{i=1}^{N'}\sum_{j=1}^{N'} f_{ij}^{(p,q)} + \sum_{cells}\sum_{i=1}^{N'}\sum_{j=1}^{N'}\sum_{k=1}^{N'} g_{ijk}^{(p,q)} \\ - \sum_{cells}\sum_{i=1}^{N'}\sum_{j=1}^{N'}\sum_{k=1}^{N'}\sum_{l=1}^{N'} h_{ijkl}^{(p,q)} + \ldots\end{aligned} \tag{2.139}$$

where the summation $\sum_{cells}$ means that every index of particles (i.e. *i, j, k* and *l*, …) is running over all 27 cells. f_{ij}, g_{ijk}, and h_{ijkl} represent exchange processes involving two electrons, three electrons, four electrons, respectively. Since we are interested in calculating the exchange force acting on the electrons within the simulation cell, we can restrict the *i*th electron to be within the simulation cell. The size of the matrix $E^{(p,q)}$ is then reduced to $(27N' \times N')$. T reduce the computational load further, we note that the second term of (2.139) which corresponds to exchange between pairs of electrons may now be written as

$$\begin{aligned}\sum_{cells}\sum_{i,j=1}^{N'} f_{ij}^{(p,q)} &= \sum_{cells}\sum_{i,j=1}^{N'} \exp\left[-\frac{Pm}{2\hbar^2\beta}\left\{\left(r_i^{(p)} - r_j^{(q)}\right)^2 - \left(r_i^{(p)} - r_i^{(q)}\right)^2 + \left(r_j^{(p)} - r_i^{(q)}\right)^2 - \left(r_j^{(p)} - r_j^{(q)}\right)^2\right\}\right] \\ &= \sum_{cells}\sum_{i=1}^{N'}\sum_{j=1}^{N'} e^{-\left(d_{ij}^2 + d_{ji}^2 - d_{ii}^2 - d_{jj}^2\right)} \\ &= \sum\nolimits_{i=1}^{N\prime}\left\{\sum_{j=1}^{N'} e^{-L_{ij}^{cell1}} + \sum_{j=1}^{N'} e^{-L_{ij}^{cell2}} + \ldots + \sum_{j=1}^{N'} e^{-L_{ij}^{cell27}}\right\}\end{aligned} \tag{2.140}$$

where

$$d_{ij}^2 = \frac{Pm}{2\hbar^2\beta}\left(r_i^{(p)} - r_j^{(q)}\right)^2$$

$$d_{ij}^2 = \frac{Pm}{2\hbar^2\beta}\left(r_j^{(p)} - r_i^{(q)}\right)^2$$

and

$$L_{ij}^{cellK} = d_{ij}^2 + d_{ji}^2 - d_{ii}^2 - d_{jj}^2 \quad K = 1, 2, \ldots, 27$$

The index i in the above equations is restricted to electrons within the simulation cell, but the index j can point to any electron over all neighbor cells including the simulation cell itself. By choosing a minimum value of L_{ij}^{cellK} over all K, we have

$$\sum_{cells} \sum_{i=1}^{N'} \sum_{j=1}^{N'} f_{ij}^{(p,q)} = \sum_{i=1}^{N'} \sum_{j=1}^{N'} e^{-L_{ij}^{min,K_o}} \left\{ 1 - \sum_{K(\neq K_o)} e^{-L_{ij}^{cellK}/L_{ij}^{min,K_o}} \right\} \tag{2.141}$$

The second term in the above equation is expected to be much smaller than 1 for most configurations. Thus we can make the approximation:

$$\sum_{cells} \sum_{i=1}^{N'} \sum_{j=1}^{N'} f_{ij}^{(p,q)} \cong \sum_{i=1}^{N'} \sum_{j=1}^{N'} e^{-L_{ij}^{min}} \equiv \sum_{i=1}^{N'} \sum_{j=1}^{N'} \sum_{k=1}^{N'} \sum_{l=1}^{N'} f_{ij,min}^{(p,q)} \tag{2.142}$$

Furthermore, by applying this approximation to exchange processes that involve more than two electrons, we reduce the computational load from the evaluation of a $27N' \times 27N'$ matrix to a more tractable $N' \times N'$ matrix:

$$\det\left(E_{27N' \times 27N'}^{(p,q)}\right) \cong 1 - \sum_{i=1}^{N'} \sum_{j=1}^{N'} f_{ij,min}^{(p,q)} + \sum_{i=1}^{N'} \sum_{j=1}^{N'} \sum_{k=1}^{N'} g_{ijk,min}^{(p,q)} - \ldots \tag{2.143}$$

or

$$\det\left(E_{27N' \times 27N'}^{(p,q)}\right) \cong \det\left(E_{N' \times N'}^{(p,q)}\right) \tag{2.144}$$

2.3.6 Calculation of Forces

Up to this point we have not considered the calculation of the forces resulting from the effective potentials/classical Hamiltonian modeling the various quantum systems discussed above. These forces are calculated classically as the negative of the gradient of the effective potentials. Details concerning the derivation of these forces are given in Appendix 2 for the electron plasma/alkali metal systems and Appendix 3 for the single electron in a non-local pseudo-potential.

For example, the effective force derived from the effective quantum potential of a N-electron system has two contributions: $\boldsymbol{f_{eff} = f_{harm} + f_{exch}}$ where $\boldsymbol{f}_{harm}$ and $\boldsymbol{f}_{exch}$ are the harmonic force and the exchange force, respectively. Let's consider the effective force on the bead (k) of an electron (i). The harmonic force may be written easily as

$$f_{i,harm}^{(k)} = \frac{mP}{\beta^2\hbar^2}\left\{\left(r_i^{(k)} - r_i^{(k+1)}\right) - \left(r_i^{(k-1)} - r_i^{(k)}\right)\right\} \tag{2.145}$$

The harmonic force is calculated using the closed necklace condition. The exchange force can be derived by calculating of $f_{i,exch}^{(k)} = -\Delta_i^{(k)} V_{exch}$ with the following matrix algebra:

$$\det A = \sum_{i=1}^{n} a_{ij} A_{ij},$$
$$\frac{\partial}{\partial x}\det A(x) = \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial a_{ij}}{\partial x} A_{ij}$$

where A is an$(n \times n)$ square matrix, a_{ij} is an element of matrix A, and A_{ij} is the cofactor of the element a_{ij}. The exchange force on the bead (k) of an electron (i) is given by:

$$f_{i,exch}^{(k)} = \frac{1}{\beta}\sum\nolimits_{\mu=1}^{P}\sum\nolimits_{v=1}^{P}\frac{\theta_{\mu v}^{+}/P^{*}}{\det\left(E^{(\mu,v)}\right)}\sum\nolimits_{p=1}^{N\prime}\sum\nolimits_{q=1}^{N\prime}\frac{\partial\left(E^{(\mu,v)}\right)_{pq}}{\partial r_i^{(k)}}B_{pq}^{(\mu,v)} \tag{2.146}$$

where $B_{pq}^{(\mu,v)}$ is a cofactor of a matrix element $\left(E^{(k,\mu)}\right)_{pq}$. Subsequent details of the calculation are given in Appendix 3. The calculation of the forces during a MD simulation is restricted to the restricted paths.

2.3.7 Isothermal Molecular Dynamics

In our path integral MD simulations, we adopt the constant temperature MD method based on three types of thermostating procedures, namely the momentum rescaling approach [64], the Andersen thermostat [77] and the Nosé chain of thermostats [78, 79].

With the momentum rescaling constant temperature method, the momenta of the particles in a simulation cell are rescaled at each time step to maintain the total kinetic energy at a constant value. Here a particle may be a bead of an electron necklace and its momentum is dependent on the chosen artificial mass m^*. If T_I is an instantaneous temperature of the system, the equi-partition theorem states that:

$$KE = \sum_{i=1}^{N}\frac{p_i^2}{2m_i} = \frac{3N}{2}k_B T_I$$

where KE is the total kinetic energy and p_i is the momentum of the ith article. Using this relation, the momentum can be rescaled such that

$$p_i^{scaled} = \sqrt{\frac{T_{ref}}{T_I}} \cdot p_i$$

where $\sqrt{T_{ref}/T_I}$ is the scaling factor and T_{ref} is the desired temperature of the system.

The momentum rescaling procedure is known to deviate from the canonical momentum distribution by order of N^{-1} for a N particles system [78]. The classical isomorphism we use to model quantum particles results in a representation in the form of closed necklaces of beads connected via harmonic springs. The stiffness of the springs depends on the number of beads and the temperature. It is known that the simulation of stiff harmonic chains suffers from non-ergodicity. To alleviate this problem, we can thermostat the electron degrees of freedom by coupling every group of N_{elec} beads corresponding to the same time slice identical label "n" to a Nosé chain of thermostats [79]. With this procedure, the first thermostat directly coupled to the electrons degrees of freedom. This first thermostat is then coupled to a second thermostat. The second thermostat is coupled to a third one, and so on. The dynamics of the last thermostat is driven by a force calculated as the difference between the desired kinetic energy and it actual kinetic energy. In this chapter when we use a Nosé chain of thermostat composed of five thermostats. Each thermostat possesses a "mass". The mass of the first thermostat is taken as 100 amu and that of the subsequent thermostats to 10 amu. We have observed that thermostating the electrons with a Nosé chain still yields non-ergodic behavior for low density quantum systems (weakly interacting) or electrons in quickly varying potentials. To overcome this difficulty we have also performed simulations by coupling individual necklace to an Andersen's thermostat [77]. An Andersen's thermostat is implemented by assigning random velocities distributed according to a Maxwell-Boltzmann distribution (corresponding to the desired temperature) to a necklace selected randomly over some preselected period of time or over a period of time selected at random.

2.3.8 Calculation of Properties

In this section, we address briefly the calculation of properties with the PIMD method. Special attention is paid to the calculation of the energy of a quantum system.

2.3.8.1 Time Averages

Using the molecular dynamics method, one can obtain the phase space trajectories, $(r(t), p(t))$, and subsequently any microscopic physical quantities, say $A(r(t), p(t))$, as functions of time. If the system is ergodic, we may compute the time average of $A(t)$ to find its corresponding macroscopic property. In other words, the average of the physical quantity A can be calculated as

$$<A \geq \lim_{T\to\infty} \frac{1}{T} \int_0^T A(r(t), p(t))dt$$

2.3.8.2 Energy Estimator

Let us consider the canonical partition function for a many-electron system with P intermediate states in the imaginary time. To evaluate the mean energy, let's rewrite the partition function:

$$Z = \frac{1}{N!}\left(\frac{Pm}{2\pi\beta\hbar^2}\right)^{3NP/2} \int \prod_{i=1}^{N}\prod_{k=1}^{P} dr_i^{(k)} e^{-\beta V_{eff}} e^{-\beta V_2}$$

where the potential energy V_2 includes all classical potentials such as the short range interatomic potential and/or local pseudo-potential between an electron and an ion as well as the long range Coulomb potential. The non-local quantum effective potential, V_{eff}, is given as

$$V_{eff} = \frac{Pm}{2\hbar^2\beta^2} \sum_{j=1}^{N}\sum_{l=1}^{P}\left(r_j^{(k)} - r_j^{(k+1)}\right)^2 + V_{eff}^{exch},$$

where

$$V_{eff}^{exch} = -\frac{1}{\beta}\sum_{s=up}^{down}\sum_{k=1}^{P}\sum_{l=1}^{P}\frac{1}{P^*}\ln\left(\det\left(E^{(k,l)}\right)\right)\theta_{kls}^{+}.$$

Now we can evaluate the energy estimator by

$$\langle E\rangle = -\frac{\partial}{\partial\beta}\ln Z \tag{2.147}$$

The estimator for large enough P becomes

$$\langle E\rangle = \frac{3PN}{2\beta} - \frac{Pm}{2\hbar^2\beta^2}\left\langle \sum_{j=1}^{N}\sum_{k=1}^{P}\left(r_j^{(k)} - r_j^{(k+1)}\right)^2\right\rangle + \left\langle \frac{\partial}{\partial\beta}\left(\beta V_{eff}^{exch}\right)\right\rangle + \langle V_2\rangle \tag{2.148}$$

In (2.148), the angled brackets denote ensemble average and these quantities can be evaluated by time average calculation. Since the last term of the energy estimator equation can be interpreted as a potential energy estimator, the kinetic energy estimator, (*KE*), becomes

$$\langle KE\rangle = \frac{3PN}{2\beta} - \frac{Pm}{2\hbar^2\beta^2}\left\langle \sum\nolimits_{j=1}^{N}\sum\nolimits_{k=1}^{P}\left(r_j^{(k)} - r_j^{(k+1)}\right)^2\right\rangle + \langle KE_{exch}\rangle \tag{2.149}$$

where $\langle KE_{exch}\rangle$ is defined by

$$\begin{aligned}\langle KE_{exch}\rangle &= \left\langle \frac{\partial}{\partial\beta}\left(\beta V_{eff}^{exch}\right)\right\rangle \\ &= \left\langle \frac{\partial}{\partial\beta}\sum_{s=up}^{down}\sum_{k=1}^{P}\sum_{l=1}^{P}\frac{1}{P^*}\ln\left(\det\left(E^{(k,l)}\right)\theta_{kls}^{+}\right)\right\rangle\end{aligned} \tag{2.150}$$

The first term in the left-hand side of (2.149) is a constant for given *P, N,* and *T*. The second term of the equation is the harmonic effective potential contribution to the kinetic energy estimator. Both terms are easily calculated with minor amount of computing time during a MD sampling of the systems' configurations. However, the third term in the equation, i.e. $\langle KE_{exch}\rangle$, which is the exchange potential contribution, cannot be calculated unless we pay a major computing cost, because the number of required operations is the order of $\sim N^3P^2$. We give a full derivation $\langle KE_{exch}\rangle$ term in Appendix 4 and simply write the result here:

$$\langle KE_{exch}\rangle = \sum\nolimits_{s=up}^{down}\sum\nolimits_{k=1}^{P}\sum\nolimits_{l=1}^{P}\frac{1}{P^*}\frac{1}{\det(E^{(k,l)})}\sum\nolimits_{i=1}^{N}\frac{Pm}{2\beta^2\hbar^2}\det\left(H_i^{(k,l)}\right)\theta_{kls}^{+} \tag{2.151}$$

where $H_i^{(k,l)}$ is an $N \times N$ matrix and its element $\left(H_i^{(k,l)}\right)_{st}$ is given by

$$\left(H_i^{(k,l)}\right)_{st} = \begin{cases} L_{st}^{(k,l)}\exp\left(-\frac{Pm}{2\beta^2\hbar^2}L_{st}^{(k,l)}\right) & \text{if } i = t \\ \exp\left(-\frac{Pm}{2\beta^2\hbar^2}L_{st}^{(k,l)}\right) & \text{if } i \neq t \end{cases}$$

and

$$L_{st}^{(k,l)} = \left(r_s^{(k)} - r_t^{(l)}\right)^2 - \left(r_s^{(k)} - r_s^{(l)}\right)^2.$$

Because of the prohibitive computational cost of the calculation of this contribution, results reported in this chapter do not account for this kinetic energy. We will see that in the case of the degenerate electron plasma, neglecting this contribution is justified. In fact Hall has shown that the exchange potential does not contribute directly to the actual energy estimator but indirectly through the equilibrium configurations [9–11].

The kinetic energy is therefore estimated using the following:

$$\langle KE \rangle = \frac{3PN}{2\beta} - \frac{Pm}{2\hbar^2\beta^2}\left\langle \sum_{j=1}^{N}\sum_{k=1}^{P}\left(r_j^{(k)} - r_j^{(k+1)}\right)^2 \right\rangle \tag{2.152}$$

The main drawback of this estimator is that it is expressed as the difference between two large quantities. This estimator presents the drawback of a variance increasing with P. Other forms of energy estimators have been proposed [13], for instance, we can cite an estimator with a better variance based on the virial method of Herman et al. [80].

The energy estimator for a single electron in a non-local pseudo-potential is derived in Appendix 5.

2.3.8.3 Pair Correlation Function

An important structural physical quantity that we can easily compute during a simulation is the pair correlation function (also called the radial distribution function), $g(r)$, which tells us the actual spatial distribution of one kind of particles with respect to another kind of particles. With the pair correlation function, we can compare simulation results with the experimental structural data, like that obtained from X-ray diffraction. The pair correlation function between the (i) and (j) types of particles, $g_{i-j}(r)$, is defined as:

$$g_{i-j}(r) = \frac{\Omega}{N_j}\frac{n_i(r)}{4\pi r^2 \delta r}$$

where Ω is the volume of a simulation cell, N_j is the number of the type j particles, and $n_i(r)$ the number of the type i particles situated at a distance r and $r + \delta r$ from a type j particle. In the case of the unpolarized electron plasma, one can calculate the partial pair correlation function for iso and heterospin electrons. In the case of the alkali metals and electrons evolving in some field of ionic pseudopotentials, in addition to electron-electron pair correlations functions, one may get insight into the electronic structure of the system by calculating electron-ion pair correlation functions.

2.4 Applications of the Pimd Method

2.4.1 Electron Plasma

We have tested the PIMD method on an unpolarized electron plasma composed of $N_{el} = 30$ electrons (number of electrons with spin up, $N_{\uparrow} = 15$ and number of electrons with spin down, $N_{\downarrow} = 15$). The simulation cell is a fixed cubic box with edge length L = 13.3 Å, 19.95 Å or 26.6 Å, which correspond to electronic densities with r_s/a_o = 5, 7.5 and 10, respectively. Periodic boundary conditions are used and the system is constituted of a simulation cell and 26 image cells. Under these conditions the matrix $[E_{n,m}]$ for each spin should be a $(27N_{\uparrow,\downarrow}x27N_{\uparrow,\downarrow})$ matrix. To make the calculation more tractable, as discussed previously, we make the numerical approximation $\det[E_{n,m}] \approx C.\det[F_{n,m}]$ where $[F_{n,m}]$ is a $N_{\uparrow,\downarrow}xN_{\uparrow,\downarrow}$ matrix which off-diagonal elements give the maximum contribution of pair exchange to the determinant among the possible combinations of exchange between the electrons in the simulation cell and the electrons in all periodic cells. $C \approx \overline{\det[F]^{26}}$ is a constant representing the mean contribution of the image cells to the overall determinant. Within this approximation, the constant, C, drops out of the force calculation. In this form, the computing time for the calculation of the exchange potential scales as $P^2N^3_{\uparrow,\downarrow}$. This scaling at present limits the applicability of the method to systems with a fairly small number of fermions. However, one may exploit the natural parallelizability of the approximate effective exchange potential over the number of beads to reduce the computational cost to a nearly linear scaling with respect to P.

We solve the equation of motion with a leap frog scheme and an integration time step of $\sim 10^{-16}$ s. Most simulations were run for an average of 30,000 time steps. In some cases for the low and intermediate density plasmas, we have ran simulations up to 50,000 steps for better equilibration, this at low temperature.

Because of the large computational cost of the calculation of the exchange effective potential and forces, the exchange forces are calculated and updated every 10 MD time steps. The values for the exchange forces are used subsequently during the 10 steps following their calculation. We have compared the calculated average energies during a simulation where the exchange forces were calculated every step and the more efficient scheme described above and found no significant statistical difference in their values.

The chosen time step is small enough to resolve the high frequency oscillations of the harmonic springs. We have elected to rescale the temperature of each chain of P beads independently of each other via a simple momentum rescaling thermostat. With this procedure we do not obtain a true canonical distribution, but most thermal averages will be accurate to orders N^{-1}. We have also verified that with this approach over the length of our simulations the chains would sample a large region of configuration space and therefore resolve not only the fast but also the slow dynamical scale.

The calculation of the Coulomb energy is handled by the usual Ewald method of summation. The Coulomb potential energy of the electron plasma is made convergent by the introduction of a positively charged background of density $-30e/\Omega$ where Ω is the cell volume. The discretized nature of the quantum particle introduces some pecularity in the calculation of the Coulomb potential. The potential at a bead n of an electron i is therefore given by:

$$u_{i,n} = \frac{1}{4\pi\varepsilon_0}\frac{(-e)}{P}\left\{ \sum_{\substack{j \neq i \\ if\ j\ \text{incell}}}^{N_{el}} \frac{1}{r_{ij,n}} erfc\left(\eta r_{ij,n}\right) - 2\frac{\eta}{\sqrt{\pi}} + \sum_{k\neq 0}^{\infty} \frac{4\pi}{\Omega}\frac{e^{-k^2/4\eta^2}}{k^2} \sum_{\substack{j \neq i \\ j\ \text{incell}}}^{N_{el}} e^{i\vec{k}.(\vec{r}_{j,n}-\vec{r}_{i,n})} - \frac{\pi(N_{el}-1)}{\Omega\eta^2} \right\}$$

where k is a reciprocal space vector associated with the periodicity of the simulated and image systems. *erfc* stands for the usual complementary error function. The first three terms in equation are the usual real and reciprocal space contributions to the potential. The last term is the contribution of the background. We chose an Ewald parameter $\eta = \frac{5.741}{L}$ for which satisfactory convergence is obtained with truncation of the real space sum at $\frac{1}{2}L$ and truncation of the reciprocal sum at $k^2 \leq 49$.

We calculate the kinetic energy with the usual energy estimator derived from $\partial \ln Z/\partial\beta$ (2.149) and neglecting the exchange contribution to the kineric energy). With this estimator, the kinetic energy is given as a small quantity, difference between two larger quantities, with a variance growing with P. This estimator, therefore, introduces an error on the calculated values of the kinetic energy which we have estimated to be on the order of 0.1 eV.

Every simulation reported starts with a different initial configuration obtained from randomly generated bead position in every electron chain. The initial bead-bead distance is determined by the temperature. In order to reduce the time for the system to reach equilibrium from its initial configuration, we have taken some care as to construct initial configurations closely related to the anticipated equilibrium state. Specifically, we have found that localized compact necklaces take more time to reach equilibrium compared to more open necklaces which beads are distributed uniformly throughout the simulation cell. This observation is more important for low density systems where the extent of space to sample is large.

In Fig. 2.10, we present results on the convergence of the discretized restricted path integral molecular dynamics. The kinetic energy (Fig. 2.10a) and potential energy (Fig. 2.10b) of the high density plasma (r_s = 5a_o) at T = 1800 K and the intermediate density plasma (r_s = 7.5a_o) at a temperature of 700 K are reported as a

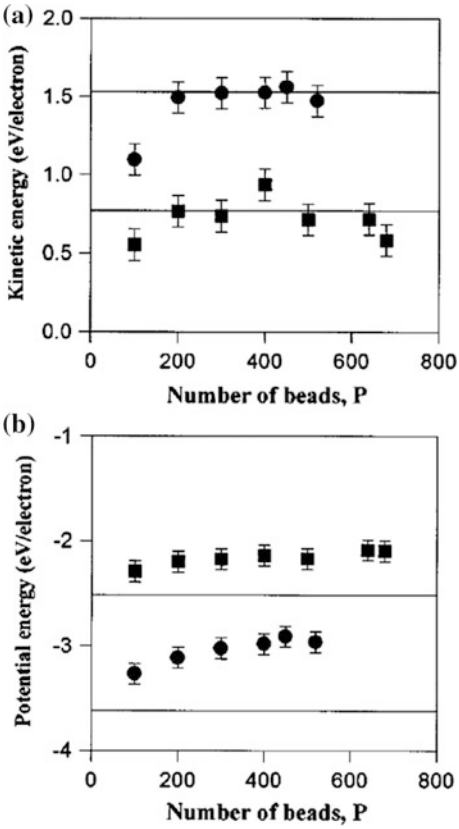

Fig. 2.10 Kinetic energy (**a**) and potential energy (**b**) as functions of number of beads in the necklace representation of quantum particles. The *circles* and *squares* refer to the high density ($r_s = 5a_o$, T = 1800 K) and medium density ($r_s = 7.5a_o$, T = 700 K) electron plasma, respectively

function of the number of beads, P. We note that the energies converge to some asymptotic value for necklaces containing as few as 200–300 beads even for the electron plasma near metallic density. In addition, to PIMD energies, we have also indicated the 0 K kinetic and potential energies of electron plasma with same density of [12, 13]. At the temperatures of 1800 and 700 K, the high and intermediate density systems are in the degenerate regime and the electronic energies should be comparable to the 0 K values. The kinetic energies are in very good agreement but some discrepencies exist between the potential energies as the restricted PIMD appears to over-estimate them. In order to further the validation of the restricted PIMD, we have conducted a series of calculations at several temperatures for the three densities. For the high and medium density systems we have used 450 and 300 beads, respectively. These number of beads fall within the region of convergence. Electrons in the low density electron plasma are discretized over 360, 380, 450, 680, 720, 780 for the temperatures 1100, 900, 450, 575 and 550, 500 and 450, 400 and 350 K.

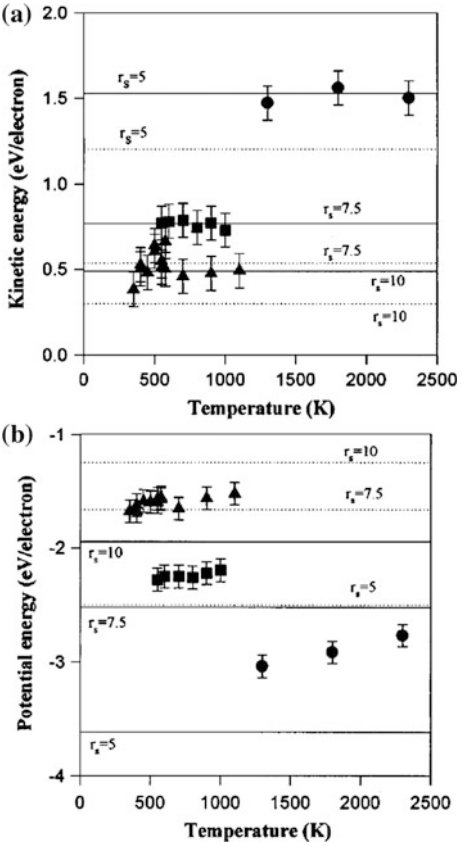

Fig. 2.11 Kinetic energy (**a**) and potential energy (**b**) as functions of temperature. The electron plasma with $r_s = 5$, $r_s = 7.5$ and $r_s = 10$ Bohr radius are refered to by *circles*, *squares* and *triangles*, respectively. The *horizontal lines* correspond to the correlated energies of Ceperley [16]. The *dotted lines* indicate the Hartree-Fock energies

The calculated kinetic energies of Fig. 2.11a are in excellent agreement with the variational Monte Carlo results of Ceperley [12, 13] for correlated plasma. We note that the kinetic energy is not varying significantly over the range of temperature studied as is expected for these plasma at the border of the degenerate and the semi-degenerate regimes. At low temperature, the low density system with large numbers of beads takes a very long time to equilibrate and sampling of phase space is not very efficient. In this case, calculation of reliable energies require very long simulations. Another difficulty in calculating reliable energies when large number of beads are used results from the fact, as was noted before, that the variance of the kinetic energy increases with P. We did not need to use so many nodes for the low density plasma even at low temperature, however, these simulations illustrate the need to use as few beads as possible within the interval of convergence. In addition to the 0 K correlated energies, we have indicated the Hartree kinetic energy ($2.21/r_s^2$ in Rydberg) with a dotted line. Figure 2.11a shows that the non-local form of the density matrix introduces some electron correlation. This is also apparent in the results for the temperature dependence of the potential energy. The calculated

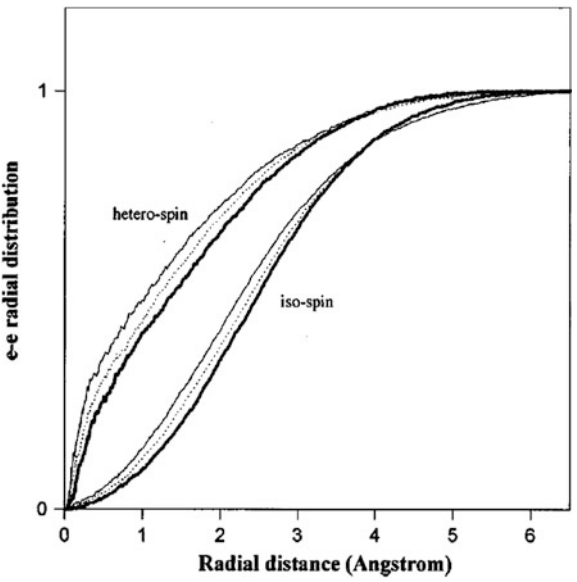

Fig. 2.12 Iso-spin and hetero-spin electron-electron pair distributions for the high density electron plasma at T = 1300 K (*thick solid line*), T = 1800 K (*dotted line*), and T = 2300 K (*thin solid line*)

potential energy falls between the fully correlated results of Ceperley and the electron/electron interaction contribution to the Hartree-Fock energy (given by $-0.916/r_s$ in Rydberg). We also note that the potential energy increases weakly with temperature and that extrapolation toward 0 K should result in potential energies in better agreement with the correlated potential energies than the uncorrelated ones. In the present model, however, the non-local effective potential introduces electron correlation between electrons with identical spins only. The present potential energies are over-estimated as correlations between electrons with opposite spins are not accounted for.

Finally as a demonstration of the effectivenes of the exchange potential, the pair correlation for iso-spin and hetero-spin electrons is reported in Fig. 2.12 in the case of the high density plasma at the three temperatures studied.

The difference between iso- and hetero-spin radial distributions is striking. In order to satisfy Pauli exclusion principle, the non-local exchange potential keeps the electrons with identical spin away from each other while electrons with different spins can approach each other quite closely. The Coulomb repulsive force is the only force keeping electrons with different spins from approaching. The non-local exchange potential is quite short range as it does not appear to affect the electron distribution beyond 5 Å. The exclusion is particularly important in the interval [0.3 Å]. The major effect of a rising temperature is the increase in pair correlation at shorter distance or in other words, the shrinkage of the exchange-correlation hole.

2.4.2 Alkali Metal

After showing that the non-local restricted PIMD can simulate with reasonable accuracy electron plasma near metal density, we apply the method to the simulation of an alkali metal from first-principle. We show that the simple metal undergoes a phase transformation upon heating. Calculated dynamic properties indicate that the atomic motion changes from a vibrational to a diffusive character identifying the transformation as melting. Calculated structural properties further confirm the transformation. Ionic vibrations in the crystal state and the loss of long-range order during melting modify the electronic structure and in particular localize the electrons inside and at the border of the ion core.

We have chosen potassium because (1) it is a prototype free-electron metal which has been studied previously using semi-empirical pair potentials, (2) there exist experimental data for the pair-correlation function of the liquid state, thermodynamic, and vibrational properties. The simulation cell contains 54 K^+ and 54 non-polarized electrons. The number of electrons with spin up and spin down is 27, respectively. In the crystal structure, the potassium ions and electrons are arranged on a body-centered-cubic (BCC) lattice. Electrons necklaces with spin up are placed on the lattice with nearest neighbors of opposite spin. Every simulation reported starts with a different initial necklace configuration obtained from randomly generated-bead positions in every electron chain. The initial bead-bead distance is chosen according to the temperature. However, in order to reduce the time for the system to reach equilibrium from its initial configuration, we have taken some care as to construct necklace initial configurations, i.e., necklace spatial extent closely related to the anticipated equilibrium state.

As temperature is varied, the dimensions of the cell are adjusted to match the experimental density of the K crystal, this over the entire range of temperature studied. Unexpectedly, our simulations have shown that the potassium model system melts at a temperature below the experimental value of the melting point, thus the density of the liquid system reported below conforms to the value of the density of the crystal at the corresponding temperature. Periodic boundary conditions (PBC) are applied to the simulation cell.

In the model, the ions K^+ are dealt with in a purely classical manner and interact through a Born-Mayer potential and the parameters of the potential are those fixed by Sangster and Atwood [81]. As to the electron/ion interaction, we have used a simple empty core local pseudopotential with a core radius R_c = 2.22 Å [82]. To optimize the calculation of the long-range Coulomb energy and forces, we do not use the Ewald summation but simply replace the long-range Coulomb potential and by a shorter range effective potential of the form $(1/r)erfc(\eta r)$ where $\eta = 5.741/L_0$ and $L_0 = 16$ Å. The physical potassium ion mass is M_I = 71 830 me, leading to an extreme disparity in electronic and ionic time scales. For practical reasons, we use a ratio of the ion mass, M_I, to the electron's bead artificial mass, m^*, of 39.1:1. The dynamics of the electrons is still significantly faster than the dynamics of the ions. We solve the equation of motion with a leapfrog scheme and an integration time

step of 2.8×10^{-16} s. Most simulations were run for a minimum of 70,000 time steps or 20 ps. In some cases for the calculation of vibrational properties, we have run simulations up to 120,000 steps. Because of the large computational cost of the calculation of the exchange effective potential and forces, the exchange forces are calculated and updated every 10 MD time steps. The values for the exchange forces are used subsequently during the 10 steps following their calculation. We have chosen to rescale the temperature of each chain of P beads independently of each other via a simple momentum rescaling thermostat.

We have studied the potassium model system at temperatures in the interval [10, 298] K. The simulation of the electronic degrees of freedom as discrete necklaces at these low temperatures would necessitate a prohibitively large number of beads for convergence with respect to P. The PIMD simulations of the electron plasma have shown that at metallic density the electron system conserves a nearly degenerate character up to a temperature of 2300 K. Since temperature does not affect significantly the electronic states at the metal density, we have thermally decoupled the classical ionic degrees of freedom and the quantum electronic degrees of freedom. The electronic necklaces are attached to a thermostat set at a temperature of 1300 K, while the ionic temperature is adjusted independently with another thermostat. At the electron temperature of 1300 K, it is sufficient to employ a reasonably small number of beads for convergence of the algorithm. We calculate the kinetic energy of the electrons with the standard estimator (P dependent variance). We have estimated the error on the calculated kinetic energy by calculating the standard deviation on the running cumulative average over the last 30,000 time steps of the simulations. This error is estimated to be on the order of 0.01 eV per electron.

In a first stage, we have investigated the convergence of the algorithm with respect to the number of beads in the electron necklaces, namely, P. For this we have calculated the electron kinetic energy at an ion temperature of 273 K for systems with varying values of P. It is important to note again that each simulation starts from different initial necklace configurations. Figure 2.13 presents the results of these calculations.

It is clearly seen that the electron kinetic energy converges to an asymptotic value of approximately 1.23 eV/electron. The algorithm appears to have nearly converged for number of beads exceeding 240. As a trade-off between accuracy and efficiency, we have chosen $P = 260$ for all subsequent simulations. The total energy of the potassium system as a function of temperature is reported in Fig. 2.14. The energy shows two regions separated by an apparent discontinuity of approximately 0.025 eV/atom. In Fig. 2.14, we have also drawn as a guide for the eyes best second-order and first-order polynomial fits to the low-temperature and high-temperature energies, respectively. The slope of the fitted curves increases from the low- to the high-temperature region indicative of larger energy fluctuations in the high-temperature systems.

The dotted line in the low-temperature region is constructed from the experimental value of the constant pressure heat capacity with the constraint that it gives the calculated energy at 76 K. In view of the error on the energy, the simulated metal is in reasonable agreement with its experimental counterpart.

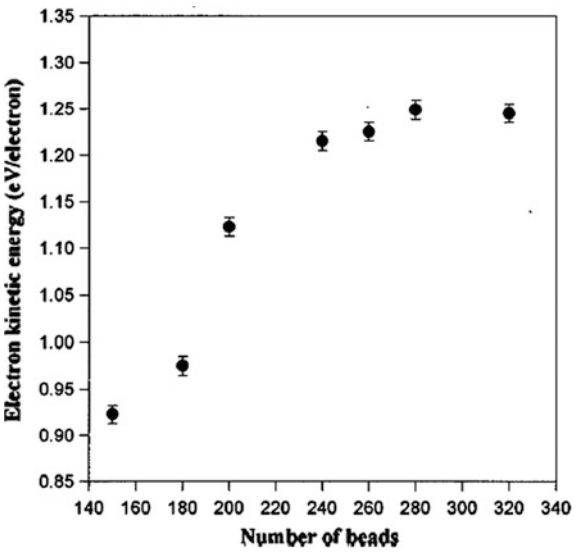

Fig. 2.13 Electron kinetic energy as a function of the number of necklace beads *P*. The ion temperature is 273 K and the electron temperature is 1300 K

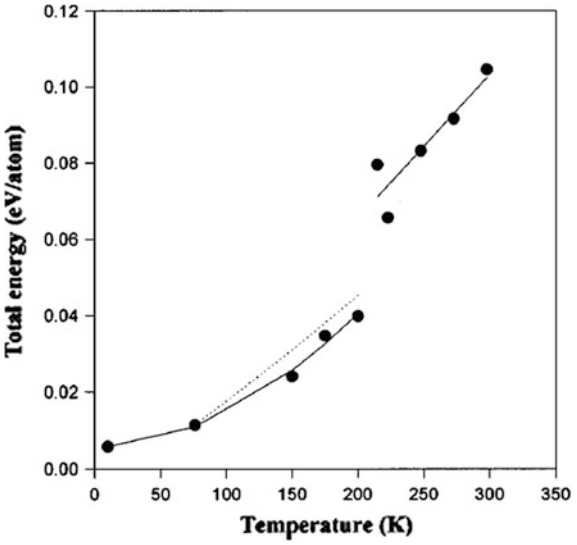

Fig. 2.14 Total energy versus temperature. The *solid lines* are fits to the data in the low- and high-temperature regions. The *dotted line* is constructed from the experimental constant pressure heat capacity such that it intersects the calculated energy at 76 K

We recall that the density of the simulated potassium system varies continuously as a function of temperature as it is set to the temperature-dependent density of the solid. Therefore, the discontinuity is not associated with any discontinuous change in the volume of the system but can only result from a structural transformation. This structural transformation takes place around 210 K. As we will see later from structural data, this is a solid to liquid transformation. The calculated transformation, therefore, underestimates the melting point by nearly 120 K as potassium melts at 333 K under atmospheric pressure. This difference cannot be assigned to

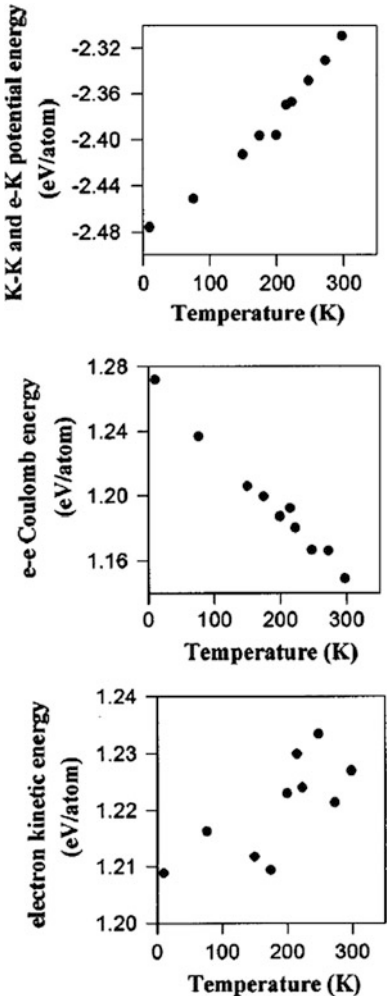

Fig. 2.15 Various contributions to the total energy of the potassium system as functions of temperature

the fact that the density of the simulated system is constrained since such a constraint should have the opposite effect of raising the melting point. The difference between experimental and simulation melting point can only result from the computer model that underestimates the strength of the K–K bond and in particular we believe that it is a consequence in part of the approximation made to reduce the range of the Coulomb interaction. In that respect it is predominantly a size effect.

To gain further insight into the energetics of the transformation, we have graphed in Fig. 2.15 some of the contributions to the energy of the system.

The only energy that is not plotted is the classical kinetic energy of the ions. Since the temperature of the ions is maintained constant by a thermostat, the ion kinetic energy is a simple linear function of temperature and cannot account for the discontinuity in the total energy. Apart from an isolated point at 200 K, the potential

energies vary reasonably continuously with temperature. In contrast, it appears that the electron kinetic energy data is separable in two groups, namely, a low-temperature group and a high-temperature group. Since the electrons in the potassium system are nearly degenerate, their kinetic energy should not be temperature dependent provided the atomic structure remains the same. Within each group the kinetic energy does not show any systematic variation. We should remember that the standard deviation on the electron kinetic energy is approximately 0.01 eV/electron. The difference between the energies of the two groups amounts to approximately 0.015–0.02 eV electron and appears to be a significant contribution to the total-energy discontinuity. The rise in kinetic energy as one crosses the discontinuity from the low-temperature to the high temperature is indicative of a change to an electronic state of higher localization in the high-temperature metal. This observation is consistent with the expected behavior of electrons in a liquid structure in contrast to a crystalline solid. As the structure disorders from crystalline to liquid, one anticipates a narrowing of the electronic band. However, since the short-range local atomic environment does not change drastically between the liquid and the solid above and below the transformation temperature, the extent of the electronic localization should be small. We characterize the atomic structure of the simulated system via the K–K pair-distribution function. The distributions calculated at several temperatures are drawn in Fig. 2.16.

The very low-temperature ion pair-distribution function shows a first nearest-neighbor peak at approximately 4.6 Å and a well-defined second-nearest shoulder at 5.2 Å. This second distance represents the lattice parameter of the crystal phase. The third-nearest-neighbor peak occurs near 7.4 Å. These features are characteristic of the body-centered-cubic structure of crystalline potassium.

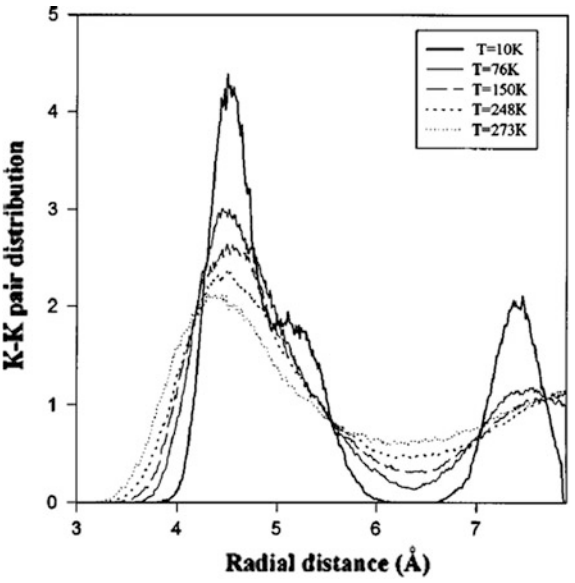

Fig. 2.16 Ion pair-distribution functions at the temperatures of 10 K (*thick solid line*), 76 K (*thin solid line*), 150 K (*dashed line*), 248 K (*thick dotted line*), and 273 K (*thin dotted line*)

As temperature increases, the second-nearest-neighbor shoulder fades away and merges with the first-nearest-neighbor peak forming a broad asymmetric peak because of the larger amplitude of atomic motion. At the temperature of 76 K, the third-nearest-neighbor peak retains its identity. At 150 K, this peak consists only of a vague shoulder part of a much broader peak that should encompass higher-order nearest neighbors. However, due to the size of our simulation cell, we cannot resolve with much confidence the pair-distribution function beyond one half the length of the edge of the simulation cell. On the same figure, we have also plotted the K–K pair distribution functions at the temperatures of 248 and 273 K. We have not represented the distribution at 298 K since it is practically identical to the one at 273 K. The maximum of the first-nearest-neighbor peak shifts toward lower values as temperature increases. At 273 K, this maximum occurs at a distance of approximately 4.3 Å. This distance is an underestimation of the experimental K–K first-nearest-neighbor distance [83] but the calculated liquid distribution is in good qualitative accord with available experimental data. To supplement the structural information provided by the ion pair-distribution functions, we report in Fig. 2.17 two-dimensional projections of the trajectories of the K ions at several temperatures.

Figure 2.17a, b correspond to the crystalline state. The ionic species vibrate about clearly well-defined equilibrium lattice positions. At the two high

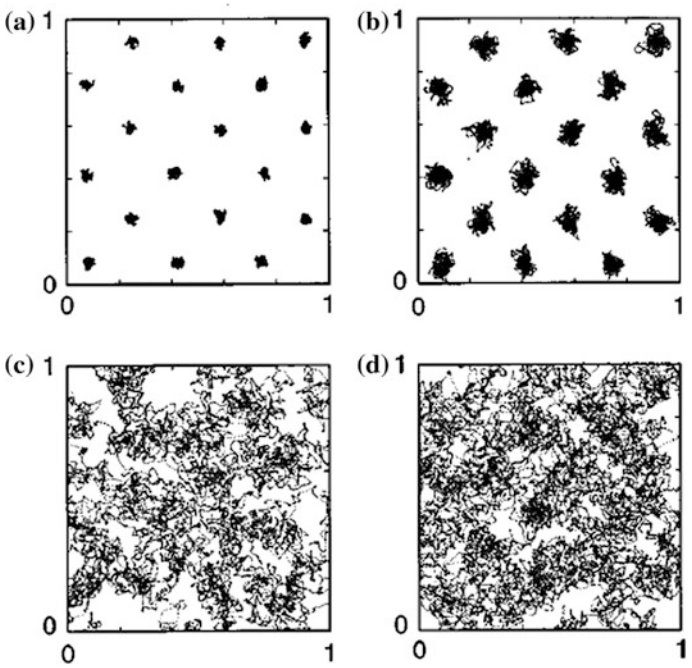

Fig. 2.17 Trajectories of the K ions at **a** T = 10 K, **b** T = 76 K, **c** T = 248 K and **d** T = 298 K

temperatures, Fig. 2.17c, d, one cannot identify lattice positions anymore. Although one may still identify some vibrational component to the ionic motion in the form of some localization in the trajectories, ionic motion is not predominantly oscillatory but also possesses a diffusive character.

More quantitative information concerning ionic motion is available from the analysis of the mean-square displacement (MSD).

The MSD is defined as

$$MSD(\tau) = \lim_{T\to\infty} \int_0^T (r(t+\tau) - r(\tau)))^2$$

with τ being a correlation time and r being the position of a particle. By calculating the MSD from the trajectories of particles, we may determine whether they behave according to diffusion processes or vibrational ones. For normal diffusion processes, i.e. one that occurs slowly with respect to microscopic times and for which spatial variations are smooth, we have $D = \frac{1}{6\tau} \cdot MSD(t)$ where D is the diffusion coefficient.

Figure 2.18 shows the mean-square displacement (MSD) of K ions as a function of time and temperature. In terms of the MSD, diffusive motion is identified by linear variation with time in the limit of large time. Vibrational motion is characterized by a time-independent MSD. At the three lowest temperatures (10, 76, and 150 K), the MSD indicates that ionic motion is vibrational. At the highest temperatures of 248, 273, and 298 K, the ions exhibit diffusive motion. It is somewhat more difficult with the present data to identify on the basis of the MSD only the

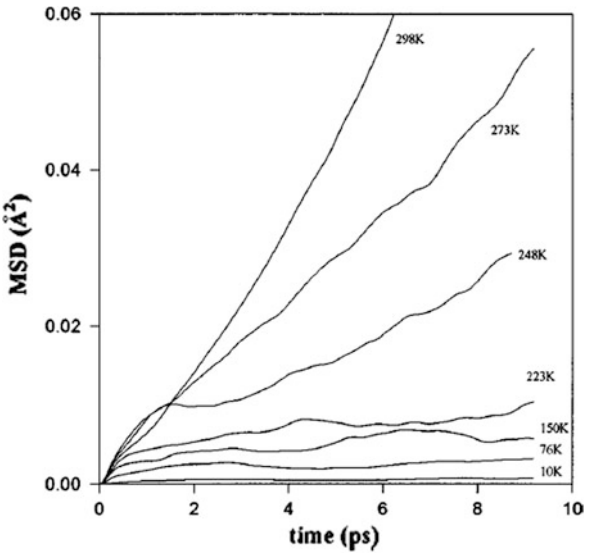

Fig. 2.18 Mean-square displacement (MSD) as a function of time and temperature

nature of atomic motion at the temperature of 223 K. However, because the density of the system is constrained to conform to that of the solid, it is not surprising that at 223 K, atoms in this liquid may display essentially vibrational motion.

Further information on the ionic motion is obtained from the calculation of the normalized velocity autocorrelation function (NVAF).

The NVAF is defined as $NVAF(\tau) = \frac{VAF(\tau)}{VAF(0)}$ where the velocity autocorrelation function (VAF) is given by $VAF(\tau) = \lim_{T\to\infty} \int_0^T (v(t+\tau) - v(\tau)))^2$. v in that equation is the velocity of a particle. Using the velocity autocorrelation function, we can also calculate the power spectrum or spectral density $G(f)$ of the NVAF. By the Wiener-Khintchine theorem [84], the power spectrum is the Fourier cosine transform of NVAF, i.e.

$$G(f) = \lim_{T\to\infty} 4 \int_0^T NVAF(\tau) \cos(2\pi f \tau) d\tau$$

where f is the frequency of a vibrating particle. The power spectrum is nothing but the vibrational density of states of a collective ionic motion.

We therefore also consider the power spectrum of the NVAF, defined as its Fourier transform. The NAVF's and the associated power spectra have to be analyzed in a qualitative manner because the time over which they are calculated is not long enough for quantitative characterization. In Fig. 2.19, the NVAF's at the two high temperatures of 273 and 298 K show features of the crystalline state (10 K), with oscillations representative of thermally excited phonons in crystal lattices. The contrast in ionic motion between the liquid and the crystal is also quite apparent in the power spectrum and in particular in the low-frequency modes. At 10 K, the power spectrum drops to zero at zero frequency. The liquid systems at 273 and more evidently at 298 K, exhibit nonzero values of the power spectrum at zero frequency. This observation is in accord with a diffusive ionic motion [85]. The peaks in the power spectra of the liquid metal are consequences of the oscillations in the NVAF's and may thus be regarded as remnants of the phonon structure observed in the crystal state. The fact that the density of the liquid is constrained to that of the crystal may accentuate this effect. As temperature increases or density decreases, these peaks should disappear with the decay of the oscillations. It is not possible to extract detailed information from the fine structure of the power spectra because of the finite time used in their calculation. However, one may compare qualitatively the calculated power spectrum at 10 K with that deduced from experimental measurements at 9 K [85]. The experimental phonon density of states possesses a major peak near 2.1×10^{12} Hz. Vibrations in the PIMD model of the crystal potassium have lower frequencies in the range $0.8–1.3 \times 10^{12}$ Hz suggestive of weaker bonds. This observation correlates closely with the observation of a calculated melting temperature underestimating the experimental melting point.

Finally, we consider the change in the electronic structure of the metal upon melting. This change is associated with an increase in electronic kinetic energy of

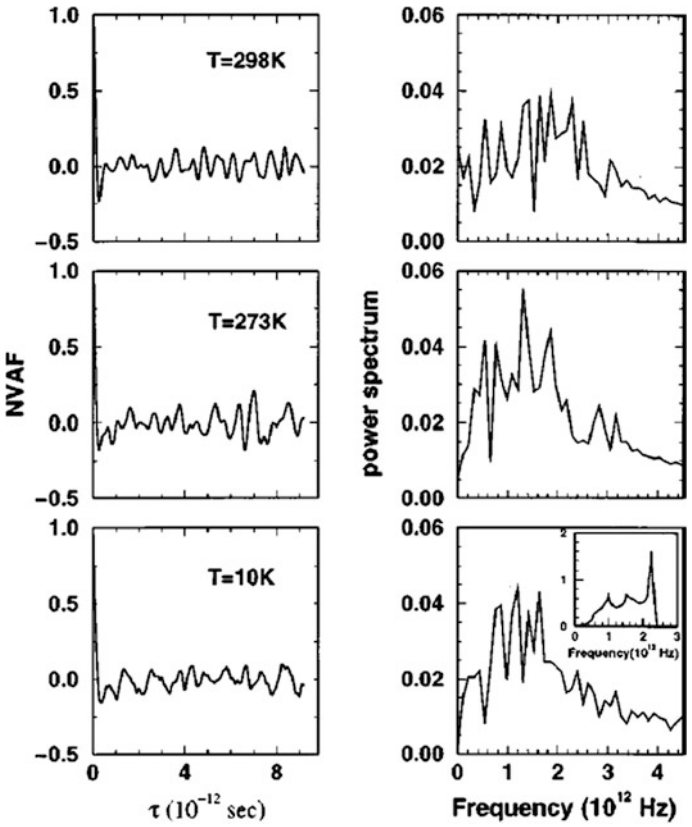

Fig. 2.19 Normalized velocity autocorrelation function (NVAF) and associated power spectrum for crystalline potassium ($T = 10$ K) and liquid metal ($T = 273$, 298 K). The *inset* in the $T = 10$ K power spectrum is the experimentally deduced phonon density of states at 9 K of [85]

approximately 0.02 eV/electron. This energy is small and thus one expects only a slight modification of the electronic structure. Such a variation is observable in the electron pair-distribution functions of Fig. 2.20.

The partial distribution functions show that the major difference between the low-temperature crystal and the liquid is an increase of the maximum in the heterospin pair correlation between 3 and 4 Å and an expansion of the exchange-correlation hole as seen in the isospin distribution. In our previous study of the effect of temperature on the electron density of an electron plasma with density near that of the present potassium system, we had shown that increasing temperature shrinks the exchange-correlation hole. However, the direct effect of temperature on the electronic structure cannot be a factor as it is maintained constant by a thermostat. Here, the expansion in the parallel-spin electron correlation may thus simply be a result of volume expansion. On the other hand, upon melting the first-nearest-neighbor and second-nearest-neighbor shells of the crystal structure

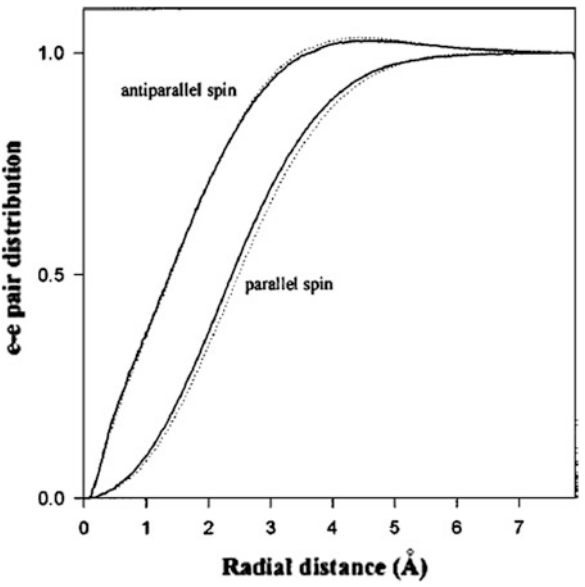

Fig. 2.20 Partial electron pair-correlation functions. The *solid line* and *dotted line* refer to the crystal at $T = 510$ K and the liquid at 273 K, respectively

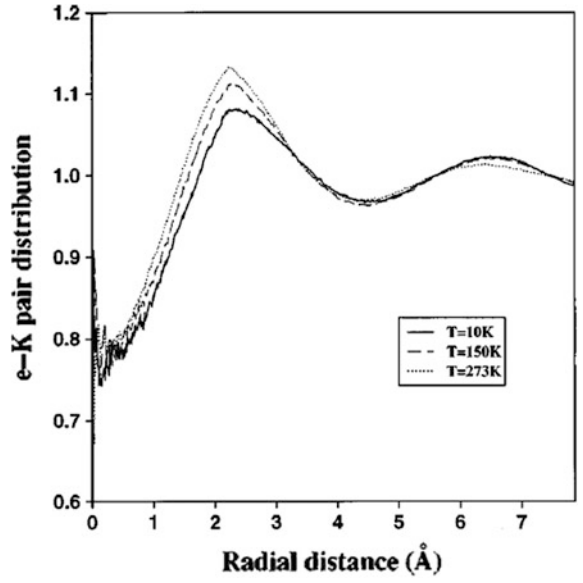

Fig. 2.21 Electron-ion pair-distribution functions at several temperatures

collapse and the ion coordination number in the liquid increases. The exchange-correlation force between neighboring isospin electrons may then induce further localization. The resulting localization within and at the border of the ionic core is seen best in the electron-ion radial distribution of Fig. 2.21.

At low temperature the ion-electron pair distribution shows a significant first maximum at a distance of 2.2 Å. This distance corresponds to approximately one half the first-nearest-neighbor interatomic distance. The ion-electron correlation reaches a minimum between 4.3 and 4.5 Å followed by a second maximum near 6.5 Å. The ion-electron pair distribution is therefore complementary of the ion-ion distribution. In other words, high electron-ion correlation is expected where there is low ion-ion correlation. The high electron density between ions is indicative of bonding. Considering now the midpoints between ionic sites as consisting of electronic sites, we can estimate the electron-electron distance in the potassium body-centered-cubic structure to be on the order of $1/\sqrt{2}$ times the lattice parameter. The electron-electron distance thus calculated amounts to approximately 3.67 Å. This number is in excellent accord with the observed maximum in the heterospin electron-electron pair distribution. With this information, we may construct a simplified picture of the electronic structure in the crystal phase. The electron density is the highest between the ions thus leading to bonding and the electronic sites are occupied alternatively by electrons with differing spin. An increase in electron localization at the electron sites occurs even in the solid state at the higher temperature of 150 K. This shows that atomic vibrations have a significant effect on the nature of the electronic states in crystalline potassium. Vibrations tend to localize the electron density. Similar observations by other investigators were made for the case of sodium clusters [9–11]. The electron density localizes further with disordering of the structure at even higher temperature. In the liquid, the electron density increases near 2.2 Å. This increase is compensated by a reduction in the electron-ion pair correlation at longer range as seen by the loss of electron-ion correlation near 6.5 Å. Since the calculated dynamical properties support the retention of vibrational motion in addition to diffusive motion in the liquid state, it is unclear at this stage which of the two processes: vibration versus disorder, contributes principally to the localization. The electron localization at the edge of the ionic cores should lead to an increase in heterospin correlation between 3 and 4 Å which is observed in Fig. 2.20. Finally, we note that the larger electron density between nearest-neighbor ions is consistent with the shorter K–K bond length in the liquid structure.

2.4.3 *Expanded Body Centered Cubic (BCC) Hydrogenoid Crystal*

We have also used the restricted path integral molecular dynamics method to study the correlated electronic structure of a half-filled expanded three-dimensional hydrogenoid body-centered cubic (BCC) lattice at finite temperatures. Starting from a paramagnetic metallic state with electron gas character, we explore the formation of bound electron states form upon expansion of the lattice.

The cubic simulation cell contains N_I = 54 ions arranged on a BCC lattice (27 unit cells) and N_{el} = 54 non-polarized electrons ($N_{s=\uparrow} = 27$ and $N_{s=\downarrow} = 27$). Periodic boundary conditions are applied. The long-range Coulomb potentials are replaced by a shorter-range screened potential of the form $(1/r)erfc(\eta r)$ where $\eta = 0.382$ Å^{-1}. With the chosen value for η, the reciprocal-space part of the Ewald sum is negligible compared to the real-space parts. For a fixed value of η, the point self-energy in the Ewald sum is independent of density and is not presented here. All potentials are truncated at half the length of the edge of the simulation cell.

We solve the equations of motion with a leap-frog scheme and an integration time step of 2.8×10^{-16} s. For the electron-ion pseudopotential, we employ an empty core local pseudopotential with core radius, R_c = 1.5 Å. As will be seen in this section, R_c is small enough to lead to a significant electronic density inside the ionic core providing the on-site electron-electron interaction necessary to observe the formation of bound states. We use P = 400 beads for the electron necklaces in order to ensure convergence of the path integral at the temperatures and densities studied. We eliminate all phonons by holding the ions at fixed positions. The temperature of the electrons is maintained at high enough values such that for the high densities paramagnetic metal state studied the electrons are in a nearly degenerate regime. We thermostat the electrons with a Nosé-Hoover chain of five thermostats. We have also coupled each necklace to an Andersen's thermostat to ensure ergodicity and satisfactory exploration of the potential surface (especially at low densities). We calculate the kinetic energy with the standard kinetic energy estimator (with P dependent variance) and Heerman et al.'s estimator.

Note that in the present model, the spin state is permanently attached to an electron. Frustration in magnetic ordering will arise from the small difference in the distance between first and second nearest neighbors in the BCC lattice compared to the deBroglie wavelength of electrons. Under the conditions of our simulations, a thermodynamic transition associated with magnetic ordering of the entire simulations cell is unlikely. Considering the spatial extend of bound electrons, the BCC lattice cannot accommodate a periodic antiferromagnetic structure with identical spin states on the first and second nearest neighbors of each site. However, the formation of small antiferromagnetic clusters is not precluded.

We conducted series of simulations of the electronic structure of the hydrogenoid BCC lattice at several temperatures with lattice parameter ranging from 5.3 to 13.3 Å. Every simulation uses as starting configuration the final equilibrated configuration from a preceding simulation at either a lower temperature and identical density or same temperature but lower or higher density. For numerical reasons associated with arithmetic limitations on the magnitude of det[E], we could not simulate expanded systems beyond r_s = 12.41. The total number of MD integration time steps for the simulations ranges from 25,000 for the highest densities where equilibrium is reached fairly rapidly and up to an excess of 200,000 steps at lower density. Because of limits in computational resources, each simulation is a sequence of shorter runs of approximately 10,000–20,000 steps. Consequently, we estimate

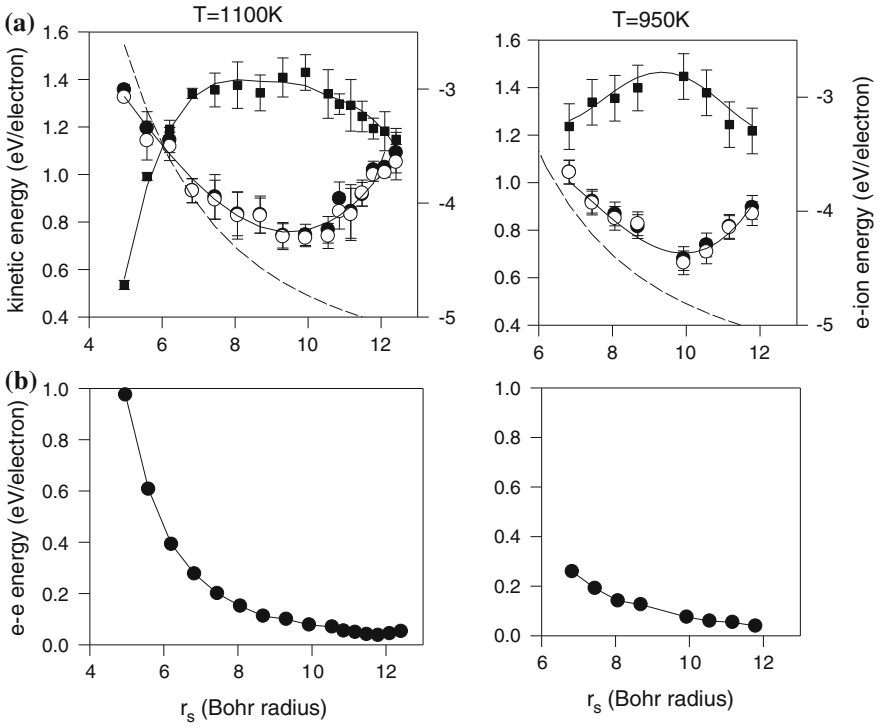

Fig. 2.22 **a** Average electron kinetic (*circles*) and electron-ion potential energy (*squares*) versus the electron sphere radius, r_s, at two temperatures. The *open* and *closed circles* are calculated with the energy estimators of [32, 38]. The *solid lines* are simple polynomial fits to serve as guides to the eye. The *thin dashed line* is the kinetic energy of the unpolarized electron plasma [19]. **b** Electron-electron (e-e) coulomb potential energy as a function of r_s

the standard deviation of the average energies from the set of individual average energies obtained from the short runs constituting each simulation.

To verify that our simulations had indeed reached thermodynamics equilibrium, we simulated a low-density systems, at $r_s = 11.17$, starting from a randomly generated paramagnetic insulator with only singly occupied lattice sites. The temperature was T = 1000 K. After the system reached equilibrium, structural and thermodynamics properties approached those of the expanded systems at the same density providing evidence that the simulations reported in this paper correspond to stable thermodynamic equilibrium.

The electron kinetic energies, electron-ion potential energy and electron-electron coulomb energy for two temperatures, T = 950 and 1100 K, are presented as functions of the electron sphere radius, r_s, in Fig. 2.22.

The kinetic energies calculated with the two estimators agree well with each other indicating that ergodicity is satisfied. The kinetic energy is U-shaped. At high density, the calculated kinetic energy follows the trend of the unpolarized uniform

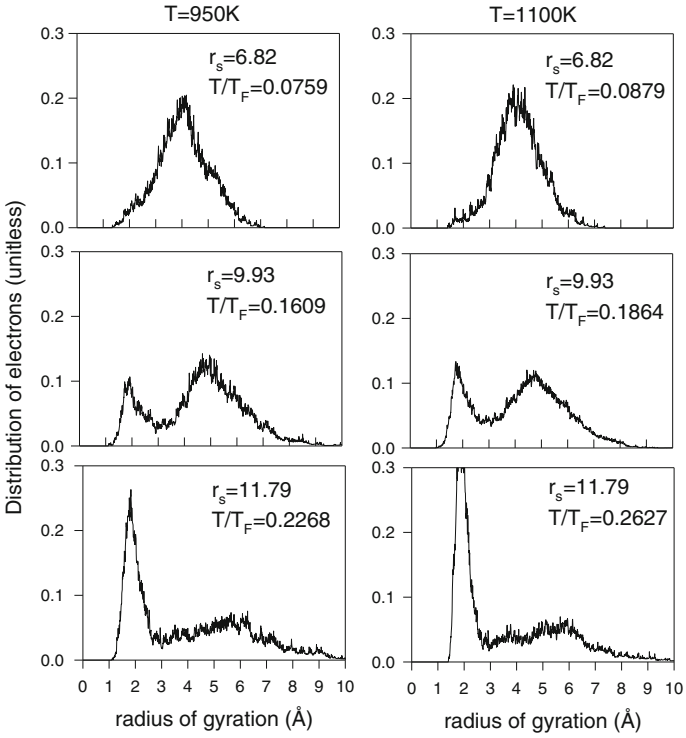

Fig. 2.23 Distributions of electron radius of gyration (see text for definition) for several densities at two temperatures

electron gas. As density decreases i.e. r_s increases, the electrons in the hydrogenoid lattice deviate from the electron gas behavior. The rise in kinetic energy at low density is associated with the formation of compact electrons i.e. highly localized electron necklaces. Figure 2.22a clearly shows that beyond r_s = 10, electron localization results from a competition between the rising kinetic energy and the decreasing electron-ion potential energy. The compact electrons with high kinetic energy are localizing inside the core of the ions to minimize their potential energy. The electron-electron coulomb energy (Fig. 2.22b) decays monotonously over the range of densities studied. Beyond r_s = 10, the electron-electron potential energy decreases only slightly (or even appears to have reached a minimum for T = 1100 K) indicating that the localized electrons have effectively excluded other electrons from the ion cores.

The formation of compact electronic states is also clearly seen in Fig. 2.23 where we report the distribution of electrons' radius of gyration (RG) at several densities and two temperatures. The radius of gyration of an electron (k) is calculated with the following expression:

$$RG^{(k)} = \frac{1}{P}\sqrt{\sum_{i=1}^{P}\left(r_i^{(k)} - \bar{r}^{(k)}\right)^2}$$

where $\bar{r}^{(k)}$ is the position of the center of mass of the electron. Parrinello and Rahman [5] have shown that $\left\langle RG^2 \right\rangle = \frac{\hbar^2/2m_e}{k_B T} C$. For a free particle, C is equal to 1 and <RG^2> is interpreted as the square of the average de Broglie wavelength for electrons with average momentum square, $m_e k_B T$.

At high density, the RG distribution takes a nearly Gaussian form with a mean RG $\sim$ 4 Å. For T = 1100 K and r_s = 6.82, the value of C is estimated to be $\sim$0.45 which is nearly half that of a free particle. The electrons have some free particle character but this value for C reflects some compactness due to electron correlation as well as the interaction with the lattice. Upon a decrease in density, compact electrons form a shoulder in the RG distribution transforming into a narrow peak centered around RG $\sim$ 2 Å. At fixed density, the weight of bound states increases as temperature rises from 950 to 1100 K. At T = 1100 K, the value of C for the bound electrons is now one order of magnitude less than the free electron value. This value is comparable to that of a single F-center studied by Parrinello and Rahman with PIMD [5]. The compact electrons are representative of bound states since their center for the motion of gyration is located on ionic positions as illustrated by the electron-ion pair correlation functions (PCF)/radial distribution functions (RDF) of Fig. 2.24.

In Fig. 2.24a we show that at the highest density studied, $r_s \sim 4.96$, decreasing the pseudo-potential core radius from R_c = 2.2–1.5 Å converts our system from a simple metal to an atomic lattice with overlapping atomic wave functions. In a simple metal like potassium the e-ion RDF exhibits oscillations that have an inverse phase to the ion-ion RDF. Note that the smaller amplitude of our e-ion RDF is probably due to the *erfc* screening of the Coulomb interactions. At R_c = 2.2 Å, the e-ion RDF shows a minimum at $r/r_s \sim 1.9$ corresponding to the closely related first and second nearest neighbor distances of the BCC lattice, r/r_s = 1.76 and 2.03, respectively. For R_c = 1.5 Å, the e-ion RDF fills the ion core and shows a maximum in phase with the ion-ion distances. The minimum near r/r_s = 1, indicates that the electrons occupying atomic sites start to exclude other electrons. We have also included in Fig. 2.24a, the electron-proton RDF for a hot dense partially ionized atomic hydrogen fluid that also shows exclusion outside the atomic radius [88, 89]. Figure 2.24b demonstrates strong electron localization within the ion cores as density decreases. The growing depression at a radius of 1 is evidence for a reduction in the number of itinerant electrons. A pictorial representation of the spatial arrangement of the localized electrons and of the free carriers within the BCC lattice is given in Fig. 2.25. This figure provides information somewhat similar to that of Fig. 2.24 but supplements it by showing that at low density, the lattice is composed of separate regions rich in bound electrons and regions rich in free carriers.

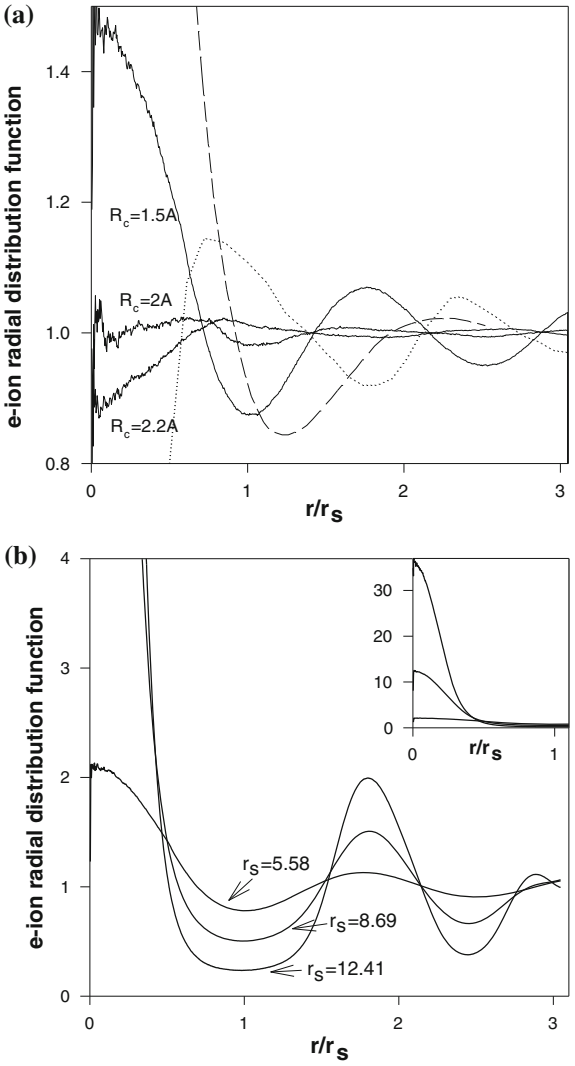

Fig. 2.24 **a** Electron-ion RDF at $r_s = 4.96$ for several values of the local pseudo-potential core radii at T = 1100 K. The *dotted line* is the RDF for potassium ($r_s = 5.024$) extracted from [86] (inner-core structure is omitted). The *dashed line* is a proton-electron RDF for the partially ionized hydrogen atomic fluid of [87] ($r_s = 2$). **b** Electron-ion RDF at T = 1100 K for several electron densities. A sharpening of the peaks and growing depressions at $r/r_s \sim 1$ and 2.5 indicates electron localization on the ion sites with decreasing density. The *inset* shows the complete distributions for radii less than the electron sphere radius

The evolution of the electronic structure of the expanded system at fixed temperature is also clearly seen in the heterospin, $g_{\uparrow\downarrow}$, and homospin, $g_{\uparrow\uparrow \text{and} \downarrow\downarrow}$, electron-electron RDF reported in Fig. 2.26.

The pair correlations are between beads of necklaces with the same label, that is, the same time slices along the discretized path representation. At high density, the electron-electron RDF calculated with the RPIMD method are characteristic of the uniform electron gas with a correlation hole in the heterospin RDF and a wider exchange-correlation hole in the homospin RDF. The inset of Fig. 2.26, compares our results at $r_s = 4.96$ with the RDF of the unpolarized uniform electron gas of

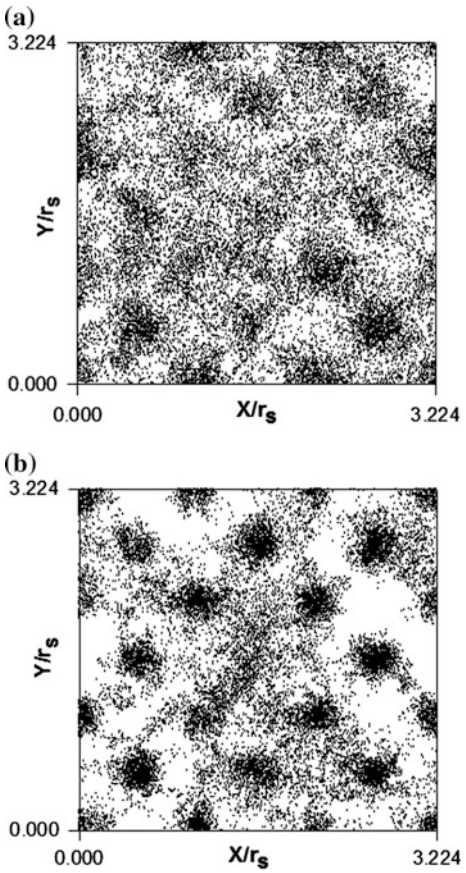

Fig. 2.25 Snapshot projections of the position of the nodes of all the electron necklaces on the (100) plane of the BCC lattice for two densities. The temperature is T = 1100 K. The electron sphere radius is **a** r_s = 7.44 and **b** r_s = 11.17

reference [90] at a similar density calculated with variational Monte Carlo (VMC) and diffusion Monte Carlo (DMC) methods. The agreement is quite good except near the origin. The PIMD heterospin RDF falls essentially between the VMC and DMC curves for most radial distances. At very short distance, the PIMD heterospin RDF does not take on the expected non-zero value at the origin but rapidly changes slope and dives toward 0. The discretized path integral representation of a quantum particle becomes exact as $P \rightarrow \infty$. In that limit the electron-electron Coulomb interaction, allows heterospin electrons to overlap to some extent. At finite P, this quantum effect is lost and the PIMD method overestimates the Coulomb force between electrons at very short distance. For P = 400 and $r_s \sim 5$, the RPIMD method does not represent the electron-electron RDF properly for radial distances shorter than $\sim$0.1 Å. This problem will not affect significantly the results presented in this paper since most of the simulations reported here are at low density for which $g_{\uparrow\downarrow}$ is very small at the origin.

As density decreases to r_s = 8.69, the heterospin RDF exhibits a shoulder at radial distances, $r/r_s \sim 0.4$. This shoulder corresponds to doubly occupied lattice

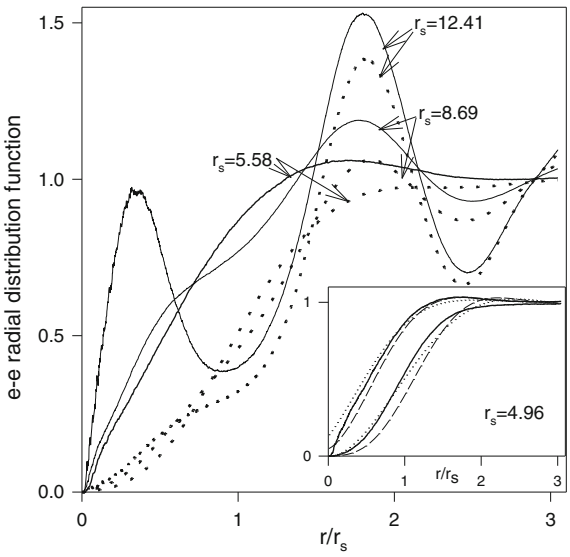

Fig. 2.26 Heterospin (*solid line*) and homospin (*dotted line*) electron-electron RDF for several densities. The temperature is T = 1100 K. The *inset* compares the RPIMD RDF for the hydrogenoid lattice at $r_s = 4.96$ (*solid line*) with the variational (*dotted line*) and diffusion Monte Carlo (*dashed line*) results of [6] for an unpolarized uniform electron gas at $r_s = 5$. The *inset* shows that quantum effects are overlooked by the discrete path integral approximation (P = 400) only at very short radii. Because of the finite P, the heterospin RDF changes slope below approximately $r/r_s \sim 0.1$ and decreases to zero in contrast to the expected non-zero value at the origin

sites (i.e. singlet). This shoulder converts into a peak for lower densities indicating further spatial localization of the electrons participating in doubly occupied sites. It is worth noting that the number of electrons in double occupancy calculated by integrating $g_{\uparrow\downarrow}(r)$ over the interval $0 < r/r_s < 1$ does not amount to more than a fraction of an electron. The homospin RDF shows also a shoulder at $r/r_s \sim 0.7$ that results from a non-zero but very small number of doubly occupied triplet sites. Upon dilation of the lattice the increasing number of singly occupied lattice sites yields a sharper peaks in the PCF encompassing both the nearest neighbor distance of the BCC lattice and the second-nearest neighbor distance.

The transfer of spectral weight from the free carrier states to the bound electron states as temperature is increased at fixed density ($r_s = 8.69$) is unambiguously confirmed in Fig. 2.27.

We note that there exists a similarity between the distribution at T = 1300, $r_s = 8.69$ and the distribution of Fig. 2.23 at T = 950 and $r_s = 9.93$. This similarity shows that the relevant energy scale for the formation of bound states upon expansion is the Fermi temperature T_F. Indeed, despite their different temperatures and densities, these two systems have very similar values of T/T_F. The inset of Fig. 2.27 summarizes the spectral reweighing between unbound and bound states. We calculated the number of bound electrons by integrating a Gaussian fit to the

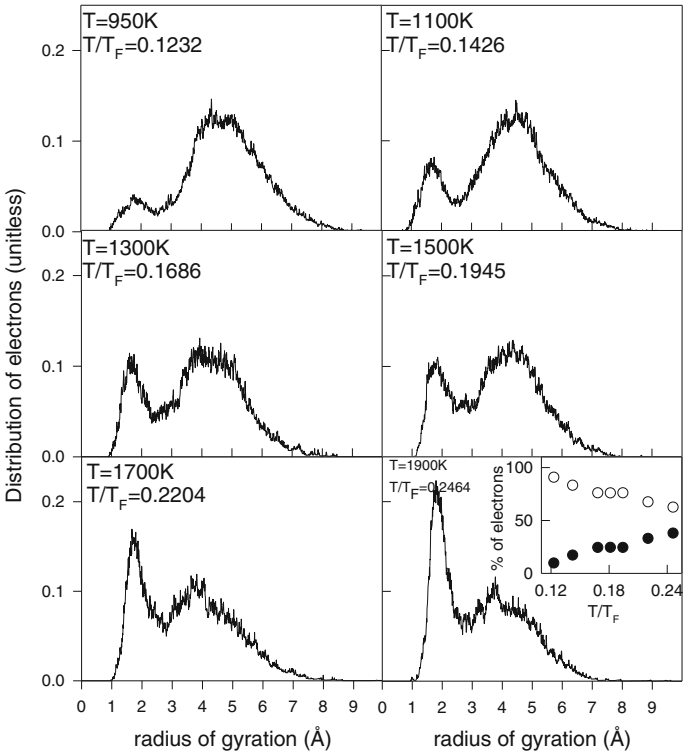

Fig. 2.27 Distributions of electron radius of gyration for a crystal with r_s = 8.69 at several temperatures. The *inset* reports the % of electrons in bound states (*closed circle*) and of free carriers (*open circle*) as a function of temperature

first peak in the distribution of RG. The number of free carriers is the difference between the total number of electrons (i.e. 54) and the number of bound electrons. At a constant density, the % of unbound electron decreases as T increases. This implies that for r_s = 8.69, the single particle density of state at the Fermi energy is non-zero and that it decreases as temperature increases.

Our observations are compatible with the predictions of the Hubbard model concerning the single particle excitation spectrum (SPES) [91, 92]. In the Hubbard model [35], a metal-insulator transition results from a competition between the electrons' potential energy that tends toward localization and kinetic energy that favors delocalization. On a lattice, the competition between the electron kinetic energy (quantified by a band width, W) and the intra-atomic energy of two electrons with antiparallel spins on a given site (interaction strength, U) may open a gap in the electronic energy spectrum leading to the formation of the so-called lower and upper Hubbard bands. At low values of U/W, the Hubbard model leads to a metallic state. At high values of U/W, the stable state is insulating. Upon increase of U/W (our simulations correspond to a fixed U and a W decreasing with increasing r_s)

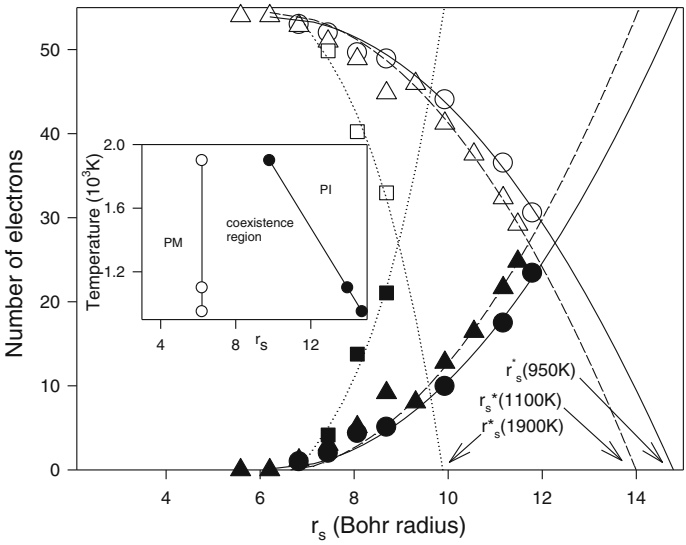

Fig. 2.28 Number of bound electrons (*filled symbols*) and unbound electrons (*open symbols*) at three temperatures (T = 950 K, *circles*; T = 1100 K, *triangles*; and T = 1900 K, *squares*). r_s^* corresponds to the density at which all the electrons ought to have transformed to bound states and is used to construct the temperature-density phase diagram in the inset. PM and PI stand for paramagnetic metal and paramagnetic insulator, respectively

from the paramagnetic metal, the SPES develops a central peak flanked by Mott's sidebands. The central peak is located at the Fermi liquid value. At equilibrium, electrons occupying the central peak are free carriers, while electrons occupying states in the lower sideband exhibit more bound-like behavior. Rare doubly occupied sites correspond to the electrons in the upper sideband. As temperature increases the spectral weight of the free-carrier central peak diminishes rapidly. This results in a transfer of spectral weight to the side bands.

Finally, we report in Fig. 2.28 the calculated number of bound and unbound electrons at the three temperatures (950, 1100, and 1900 K) at which density was varied systematically over a wide range of values.

The number of bound electrons is estimated from the first peak in the distributions of RG as mentioned before. At all three temperatures, the number of bound electrons increases with decreasing density, i.e. increasing r_s. Independently of temperature, at least to the level of resolution of our calculations, bound electronic states appear at $r_s \sim 6$ corresponding to an electron density $n = 6.75 \times 10^{-3}$ Å^{-3}. This density is in good agreement with Mott's criterion for the formation of bound states: $n^{1/3} r_H > 0.4$ [93].

The PIMD method did not allow us to expand the BCC lattice beyond $r_s \sim 12$. However, we can estimate the density at which all the electrons ought to be localized on ionic sites, that is, the density for completion of the transition from a paramagnetic metal to a paramagnetic insulator. For this, we fit the calculated

number of bound electrons at the three temperatures to monotonously increasing functions that reach the value 0 at $r_s = 6.2$. The choice of a continuous function for this fit is motivated by the observation of Kohn and Majumdar that the transition from unbound to bound states in a non-interacting Fermi gas occurs with continuous change of the properties [94]. These authors conclude, however, that the properties may not vary continuously if one has strong interaction between the electrons as is the case here. The inset of Fig. 2.28 summarizes our findings in the form of a temperature-density phase diagram. This diagram shows two phase-boundaries. A boundary, nearly independent of temperature, separates the paramagnetic metallic state (PM) from the region where bound and free carriers coexist. The second boundary separates the coexistence region and the paramagnetic insulting state (PI). The width of the coexistence region narrows as temperature increases suggesting that the transition may be of first-order in accord with several studies of the Hubbard model [91, 95–99]. Indeed, a cross over transition would imply that the width of the coexistence region increases with temperature [100]. The temperature dependence of the phase boundaries in our T versus r_s phase diagram shows similar trends observed in T versus U/W phase diagrams for the Hubbard model in infinite dimensions [98, 99]; namely a significantly weaker temperature dependence of the PM-coexistence region boundary compared to the boundary between the PI and the coexistence regions.

2.4.4 *Electron in Non-local Pseudo-potential*

Using the discretized path integral with non-local pseudopotential, we simulated using MD method the behavior of the valence electron in the field of a sodium ion. We considered the two cases: the 3s state, and the 3p state. The MD simulations were conducted at a constant temperature of 1006 K using a momentum rescaling thermostat. The electron was represented by a set of 384 nodes. The reason for picking this number of nodes was based on the work of Li and Broughton's [101]. In their work, they used 400 nodes to represent a 2s electron in the field of a lithium ion. The simulation method employed was based on path integral molecular dynamics with a high order correction to the Trotter expansion. The pseudopotential used in their case was of the BHS type [102]. One of the reasons why Li and Broughton's used the higher order correction to the Trotter expansion was to reduce the computational time by decreasing the number of nodes. The equations of motion were solved using the leapfrog algorithm with a time integration step of 1.3×10^{-17} s. Each simulation run was carried on for 4,00,000 time steps.

In Fig. 2.29 we show the pseudopotentials for sodium. We generated these potentials based on BHS work [102]. The steep barrier in the $l = 0$ pseudopotential represents the effect of Pauli exclusion due to the core electrons. In other words, this barrier prevents the 3s electron from staying in the core region. The pseudopotential for $l = 1$ is also shown in Fig. 2.29.

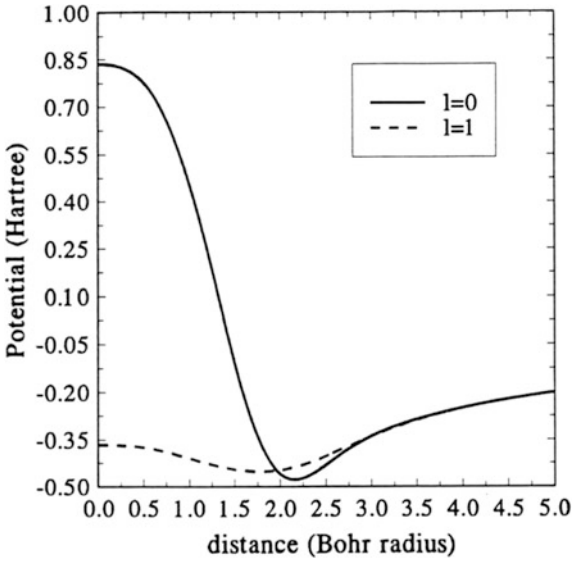

Fig. 2.29 Non-local pseudopotentials for Na that were generated using BHS method [102]. The $l = 0$, and $l = 1$ pseudo-potentials are shown

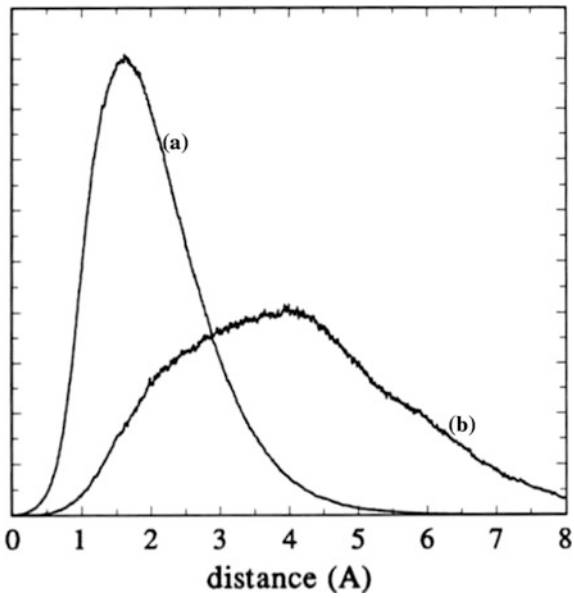

Fig. 2.30 The pseudo-density of the valence electron for two cases: **a** 3s state, and **b** pure 3p state

Figure 2.30 shows the pseudo-density for the 3s state, and the pure 3p state. In this case, the effects of the 3s pseudopotentials were filtered out to prevent the 3p state from acquiring any local 3s character. The centrifugal effects are very pronounced here and help to keep the 3p electron away from the core region.

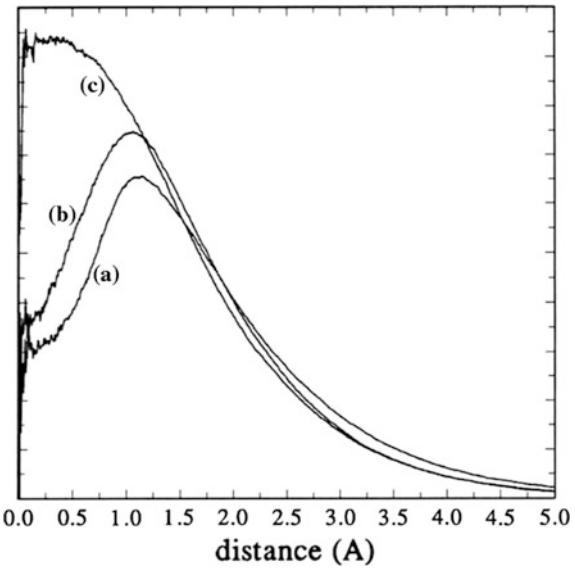

Fig. 2.31 The pseudo-amplitude for: **a** the 3s state, **b** the hybrid 3s-3p state, **c** the 3p state using local pseudo-potential only

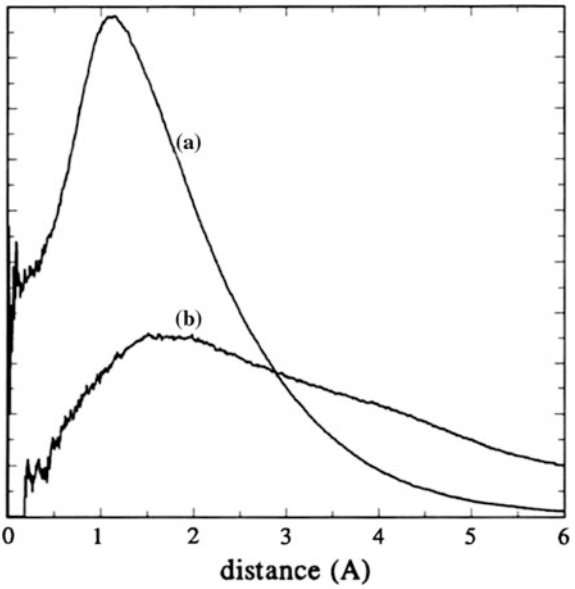

Fig. 2.32 The pseudo-amplitude for: **a** 3s state, **b** 3p state in non-local pseudo-potential

The difference in the pseudo-density curves between the 3s and the 3p is to be expected. This difference can be explained if we remember that the 3p orbit in non-hydrogenic atoms is more shielded from the nucleus attraction than the 3s orbit. Therefore, we expect the 3p electron in sodium to stay at larger distances

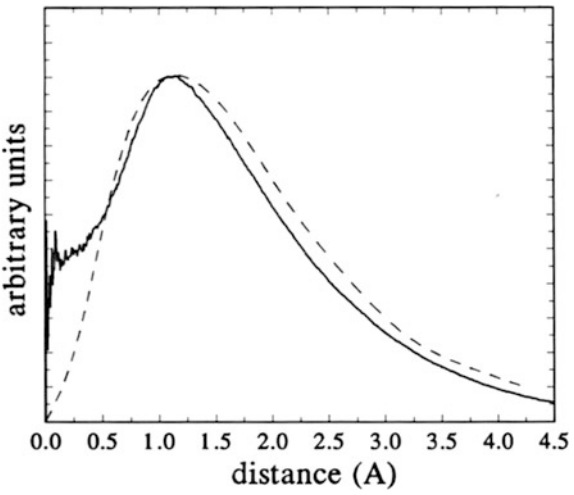

Fig. 2.33 The 3s amplitude using PIMD method with non-local pseudo-potential (*solid line*). The 3s amplitude in Na based on Melius' work (*dashed line*) [63, 103]

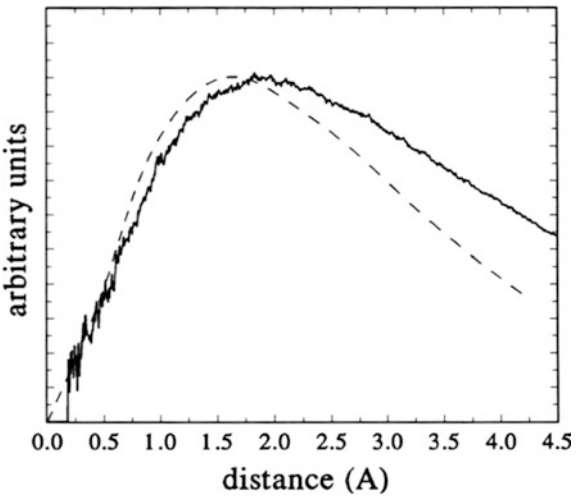

Fig. 2.34 The 3p amplitude using PIMD method using non-local pseudo-potential (*solid line*). Melius' result for the 3p amplitude (*dashed line*) [63, 103]

from the nucleus than the 3s. This can also be seen from the amplitude plots in Fig. 2.31.

The 3p pseudo-amplitude in the non-local pseudopotential is shown, along with the 3s state, in Fig. 2.32. The relative shift between the 3p and the 3s state is obvious in the figure. Comparison of the 3s amplitude generated using our PIMD method with the 3s amplitude in Fig. 2.33 [63, 103], which was generated by Melius et al. using the coreless Hartree-Fock method, shows a very good agreement between the two approaches. Figure 2.34 depicts the 3p orbit. As shown in this figure, our 3p orbit matches well with Melius' result up to a radius of 1.7 Å. The reason for the divergence of our results from Melius' beyond 1.7 Å is due to the

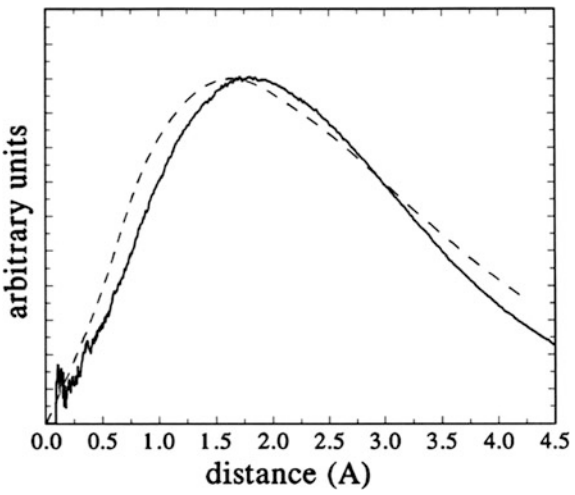

Fig. 2.35 Effect of the correction on the non-locality radius. The *solid line* is the amplitude obtained via PIMD method. The *dashed line* is due to Melius' [63, 103]

artificial constraint that we put on our system. The constraint allows a longer radius (20 Å) for non-locality. This is not physical, of course. But in any case, the centrifugal effect is still there. In order to more accurately include non-local effects, the mentioned radius must be shortened. In this case we used 3 Å (this is the point where the two pseudopotentials converge to the same curve). The amplitude in this case is shown in Fig. 2.35. We can see the enhancement in our data as compared to Melius'.

2.4.5 Conclusions

The PIMD method described in this chapter is made possible by the following features:

(a) the discretized path integral representation of quantum particles as closed "polymeric" chains (necklaces) of classical particles (beads) coupled through harmonic springs,
(b) the treatment of quantum exchange as crosslinking of the chains,
(c) the non-locality of crosslinking (exchange/correlation) along the chains (in imaginary time),
(d) the restricted path integral to resolve the problem of negative weights to the partition function resulting from the crosslinking of even numbers of quantum particles,
(e) the extension of the path integral to account for non-local pseudo-potentials.

The real space representation of quantum particles within the PIMD enables easy incorporation of quantum degrees of freedom in existing classical molecular dynamics. This ability may be particularly useful in the simulation of materials

from a concurrent multiscale point of view whereby a small quantum model of a material or chemical process may be embedded locally within a classical model. Furthermore, MD is a dynamical simulation method. Therefore, it allows for the calculation of dynamical properties (such as vibrational and transport properties). The PIMD method may enable simulation of a wide range of important materials problems previously unaccessible with conventional methods. For instance, new areas for advancement of knowledge include materials problems where structural and electronic behavior are intimately related at finite temperatures. More specifically, the PIMD method does not rely on the Born-Oppenheimer approximation that decouples the ionic and electronic degrees of freedom and can therefore shed light on dynamical coupling between electrons and ions.

The PIMD algorithm is ideally suited for parallel computation, because we can independently evaluate the exchange forces between electrons on P different nodes. This parallelizability makes it well suited for use on massively parallel computer platforms. With the speed of supercomputers and the accuracy of a first principle technique, the PIMD method may be used to shed light on even more complex problems and larger systems than those illustrated in this chapter.

Improvements to the PIMD method could be achieved by considering higher–order expansions of the density matrix. Good convergence of the path integral representation of the density matrix with a smaller number of beads, P, can be achieved by using better forms of the density matrix such as a higher-order correction to Trotter expression [101].

In the form presented in this chapter, the computing time for the computation of the exchange forces between electrons scales as P^2N^3 where N is the number of isospin electrons and P is the number of beads. This scaling restricts the current applicability of the method to small systems (small N) and high temperature (small P). It is possible, however, to improve significantly on the computational cost of the technique. One can take advantage of the short-range nature of exchange to achieve a better scaling with respect to N. For instance, a large simulation cell may be divided up into subcells which size is determined by some appropriate cut-off length. Consequently the number of operations would reduce to P^2M^3 N where M is the number of isospin electrons in a subcell.

In its present form, the PIMD method is not able to calculate dynamical and transport properties of the quantum electrons. This drawback can be overcome by using the path integral centroid MD method (CMD). Cao and Voth have first proposed this approach to semi-classical dynamics simulation of quantum systems as an alternative to solving the time-dependent Schrödinger equation [104–108]. The CMD is the real-time dynamical evolution of the centroid (the path integral averaged coordinate) of a quantum particle. The centroid moves on an effective potential surface of centroid mean force that reflects the effect of quantum fluctuations. This provides extensive applicability to semi-classical simulations of many-body quantum dynamics.

Another approach to the use of path-integral inspired algorithms for MD has frequently been applied to polymer simulations. In these approaches, the exchange forces are neglected and the MD is driven from the bead locations, velocities, and forces that may be supplied from empirical potentials. (see Chap. 3) The forces may also arise, as in the previous chapter, from DFT calculations, however it should be clear that the quantum statistics of the entities represented by the PI beads is neither fermionic nor bosonic in these simulations. Various computer codes are available for implementing these type of calculations including a python script [109].

Appendix 1: Free Particle Propagator

In this Appendix, we will prove equation (2.16)

$$\rho_0(r_n, r_{n+1}; \epsilon) = \left(\frac{m}{2\pi\epsilon\hbar^2}\right)^{3/2} \exp\left(-\frac{m(r_n - r_{n+1})^2}{2\epsilon\hbar^2}\right) \tag{2.153}$$

We assume that every state is defined by a set of plane waves. The free particle propagator in the position representation can be written as

$$\begin{aligned}\rho_0(r_i, r_j; \epsilon) &= \left\langle r_i \middle| e^{-\epsilon\widehat{\mathrm{T}}} \middle| r_j \right\rangle \\ &= \int dp \langle r_i|p\rangle \left\langle p \middle| e^{-\epsilon\widehat{p^2}/2m} \middle| r_j \right\rangle \\ &= \int dp \langle r_i|p\rangle \langle p|r_j\rangle^{e^{-\epsilon\widehat{p^2}/2m}}\end{aligned} \tag{2.154}$$

We used a completeness of momentum states, $1 = \int dp|p\rangle\langle p|$, in (2.154). If we use a plane wave basis $\langle r_i|p\rangle = \frac{1}{(2\pi\hbar)^{3/2}}\exp(ip \cdot r_i/\hbar)$, (2.154) becomes

$$\begin{aligned}\rho_0(r_i, r_j; \epsilon) &= \int \frac{dp}{(2\pi\hbar)^3} e^{ip\cdot(r_i - r_j)/\hbar} e^{-\epsilon p^2/2m} \\ &= \frac{4\pi}{(2\pi\hbar)^3} \int_0^\infty dp\, p \frac{\sin(pr_{ij}/\hbar)}{r_{ij}/\hbar} e^{-\epsilon p^2/2m}\end{aligned} \tag{2.155}$$

where $r_{ij} = |r_i - r_j|$. From an integral table, we have

$$\int_0^\infty x \sin(tx) e^{-ax^2} dx = \frac{1}{4a}\sqrt{\frac{\pi}{a}} t e^{-t^2/4a}$$

Thus if we set a = $\epsilon/2m$ and $t = r_{ij}/\hbar$, we obtain (2.153).

Appendix 2: Exchange Force Calculation

Here we calculate the exchange force resulting from the effective exchange potential in the Hamiltonian given by (2.146)

$$f_{i,exch}^{(k)} = \frac{1}{\beta}\sum_{\mu=1}^{P}\sum_{\nu=1}^{P}\frac{\theta_{\mu\nu}^{+}/P^{*}}{\det(E^{(\mu,\nu)})}\sum_{p=1}^{N'}\sum_{q=1}^{N'}\frac{\partial\left(E^{(\mu,\nu)}\right)_{pq}}{\partial r_i^{(k)}}\left(B^{(\mu,\nu)}\right)_{pq} \tag{2.156}$$

where

$$\begin{aligned}\left(E^{(\mu,\nu)}\right)_{pq} &= \exp\left\{-C_0\left(\left(r_p^{(\mu)} - r_q^{(\nu)}\right)^2 - \left(r_p^{(\mu)} - r_p^{(\nu)}\right)^2\right)\right\} \\ &= \exp\left\{C_0\left(\left(r_p^{(\nu)}\right)^2 - \left(r_q^{(\nu)}\right)^2 + 2r_p^{(\mu)}\cdot r_q^{(\nu)} - 2r_p^{(\mu)}\cdot r_p^{(\nu)}\right)\right\}\end{aligned} \tag{2.157}$$

with $C_0 = \frac{Pm}{2\beta\hbar^2}$. $\left(B^{(\mu,\nu)}\right)_{pq}$ is a cofactor of a matrix element $\left(E^{(\mu,\nu)}\right)_{pq}$. Using (2.157), we obtain

$$\begin{aligned}\frac{\partial\left(E^{(\mu,\nu)}\right)_{pq}}{\partial r_i^{(k)}} = 2C_0\Big(&r_p^{(\nu)}\delta_{i,p}\delta_{k,\nu} - r_q^{(\nu)}\delta_{i,q}\delta_{k,\nu} + r_q^{(\nu)}\delta_{i,p}\delta_{k,\mu} + r_p^{(\mu)}\delta_{i,q}\delta_{k,\nu} - r_p^{(\nu)}\delta_{i,p}\delta_{k,\mu} - r_p^{(\mu)}\delta_{i,p}\delta_{k,\nu}\Big) \\ &\left(E^{(\mu,\nu)}\right)_{pq}\end{aligned} \tag{2.158}$$

The first two terms of the left-hand side of (2.158) with (2.156) become

$$\begin{aligned}f_{i,exch}^{(k),1st,2nd} &= \frac{Pm}{\beta^2\hbar^2}\sum_{\mu=1}^{P}\sum_{\nu=1}^{P}\frac{\frac{\theta_{\mu\nu}^{+}}{P^{*}}}{\det(E^{(\mu,\nu)})}\sum_{p=1}^{N'}\sum_{q=1}^{N'}\left(r_p^{(\nu)}\delta_{i,p}\delta_{k,\nu} - r_q^{(\nu)}\delta_{i,q}\delta_{k,\nu}\right)E_{pq}^{(\mu,\nu)}B_{pq}^{(\mu,\nu)} \\ &= \frac{Pm}{\beta^2\hbar^2}\sum_{\mu=1}^{P}\frac{\frac{\theta_{\mu k}^{+}}{P^{*}}}{\det(E^{(\mu,k)})}\left(\sum_{q=1}^{N'} r_i^{(k)}E_{iq}^{(\mu,k)}B_{iq}^{(\mu,k)} - \sum_{p=1}^{N'} r_i^{(k)}E_{pi}^{(\mu,k)}B_{pi}^{(\mu,k)}\right) = 0\end{aligned} \tag{2.159}$$

In (2.159), we used the relation $\det\left(E^{(k,\mu)}\right) = \det\left(E^{(\mu,k)}\right)$ and the matrix algebra

$$\det A = \sum_{i=1}^{N} a_{ij}A_{ij} = \sum_{j=1}^{N} a_{ij}A_{ij}$$

From the third and the fifth terms of equation (2.158), we obtain:

$$f_{i,exch}^{(k),3rd,5th} = \frac{Pm}{\beta^2\hbar^2}\sum_{\mu=1}^{P}\sum_{\nu=1}^{P}\frac{\theta_{\mu\nu}^{+}/P^{*}}{\det(E^{(\mu,\nu)})}\sum_{p=1}^{N'}\sum_{q=1}^{N'}\left(r_q^{(\nu)}\delta_{i,p}\delta_{k,\mu} - r_p^{(\nu)}\delta_{i,p}\delta_{k,\mu}\right)E_{pq}^{(\mu,\nu)}B_{pq}^{(\mu,\nu)}$$
$$= \frac{Pm}{\beta^2\hbar^2}\sum_{\nu=1}^{P}\frac{\theta_{k\nu}^{+}/P^{*}}{\det(E^{(k,\nu)})}\sum_{q=1}^{N'}\left(r_q^{(\nu)} - r_i^{(\nu)}\right)E_{iq}^{(k,\nu)}B_{iq}^{(k,\nu)} \tag{2.160}$$

Similarly from the fourth and sixth terms, we get

$$f_{i,exch}^{(k),4th,6th} = \frac{Pm}{\beta^2\hbar^2}\sum_{\mu=1}^{P}\sum_{\nu=1}^{P}\frac{\theta_{\mu\nu}^{+}/P^{*}}{\det(E^{(\mu,\nu)})}\sum_{p=1}^{N'}\sum_{q=1}^{N'}\left(r_p^{(\mu)}\delta_{i,q}\delta_{k,\nu} - r_p^{(\mu)}\delta_{i,p}\delta_{k,\nu}\right)E_{pq}^{(\mu,\nu)}B_{pq}^{(\mu,\nu)}$$
$$= \frac{Pm}{\beta^2\hbar^2}\sum_{\mu=1}^{P}\frac{\theta_{\mu k}^{+}/P^{*}}{\det(E^{(\mu,k)})}\left(\sum_{p=1}^{N'}r_p^{(\mu)}E_{pi}^{(\mu,k)}B_{pi}^{(\mu,k)} - \sum_{q=1}^{N'}r_i^{(\mu)}E_{iq}^{(\mu,k)}B_{iq}^{(\mu,k)}\right)$$
$$= \frac{Pm}{\beta^2\hbar^2}\sum_{\mu=1}^{P}\frac{\theta_{\mu k}^{+}/P^{*}}{\det(E^{(\mu,k)})}\left(\sum_{p=1}^{N'}r_p^{(\mu)}E_{pi}^{(\mu,k)}B_{pi}^{(\mu,k)} - \sum_{p=1}^{N'}r_i^{(\mu)}E_{pi}^{(\mu,k)}B_{pi}^{(\mu,k)}\right)$$
$$= \frac{Pm}{\beta^2\hbar^2}\sum_{\mu=1}^{P}\frac{\theta_{\mu k}^{+}/P^{*}}{\det(E^{(\mu,k)})}\sum_{p=1}^{N'}\left(r_p^{(\mu)} - r_i^{(\mu)}\right)E_{pi}^{(\mu,k)}B_{pi}^{(\mu,k)} \tag{2.161}$$

From the above equations, the exchange force felt by the kth bead of the ith electron becomes

$$f_{i,exch}^{(k)} = \frac{Pm}{\beta^2\hbar^2}\sum_{\mu=1}^{P}\frac{\theta_{k\mu}^{+}/P^{*}}{\det(E^{(k,\mu)})}\left\{\det\left(f_i^{(k,\mu)}\right) + \det\left(G_i^{(\mu,k)}\right)\right\} \tag{2.162}$$

where the elements of matrix $f_i^{(k,\nu)}$ and $G_i^{(\nu,k)}$ are defined as

$$\left(F_i^{(k,\mu)}\right)_{pq} = \begin{cases} \left(r_q^{(\mu)} - r_i^{(\mu)}\right)\left(E^{(k,\mu)}\right)_{pq} & \text{if } p = 1 \\ \left(E^{(k,\mu)}\right)_{pq} & \text{if } p \neq 1 \end{cases} \tag{2.163}$$

and

$$\left(G_i^{(\mu,k)}\right)_{pq} = \begin{cases} \left(r_p^{(\mu)} - r_i^{(\mu)}\right)\left(E^{(\mu,k)}\right)_{pq} & \text{if } q = i \\ \left(E^{(\mu,k)}\right)_{pq} & \text{if } q \neq i \end{cases} \tag{2.164}$$

Appendix 3: Derivation of the Force on an Electron in a Non-local Pseudo-potential

We start with the expression for the effective potential of the electron necklace in a non-local pseudo-potential given by (2.125) which is repeated below

$$V_{eff} = -\frac{1}{\beta}\sum_{n=1}^{P} Ln\Big\{e^{-\frac{\beta}{P}V_{loc}}e^{-\beta C(r_n-r_{n+1})^2} \times \mathfrak{F}_{\frac{1}{2}}(2\beta C r_n r_{n+1})\Big[e^{-\frac{\beta}{P}V_0(r_n)} - e^{-\frac{\beta}{P}V_1(r_n)}\Big] + e^{-\frac{\beta}{P}[V_{loc}+V_1(r_n)]}e^{-\beta C(\vec{r}_n-\vec{r}_{n+1})^2}\Big\} \tag{2.165}$$

The total effective force on the electron is the vector sum of all the forces acting on the nodes. The force on the nth node of the electron is given by:

$$\vec{F}_n = -\frac{\partial V_{eff}}{\partial \vec{r}_n} \tag{2.166}$$

Using $x = 2\beta C r_n r_{n+1}$ and absorbing V_{loc} in the angular momentum, dependent potentials will help us to simplify the bookkeeping. The force on the nth node is the given by:

$$\begin{aligned}\vec{F}_n = &\frac{1}{<\ldots>}\Bigg\{-2C(r_n - r_{n+1})\frac{\vec{r}_n}{r_n}e^{-\beta C(r_n-r_{n+1})^2} \times \mathfrak{F}_{\frac{1}{2}}(\chi)\Big[e^{-\frac{\beta}{P}V_0(r_n)} - e^{-\frac{\beta}{P}V_1(r_n)}\Big] \\ &+ 2Cr_{n+1}\frac{\vec{r}_n}{r_n}\frac{e^{-\chi}}{\chi}\left[\cosh\chi - \sinh\chi\left(1+\frac{1}{\chi}\right)\right] \\ &\times e^{-\beta C(r_n-r_{n+1})^2}\Big[e^{-\frac{\beta}{P}V_0(r_n)} - e^{-\frac{\beta}{P}V_1(r_n)}\Big] \\ &- \frac{\vec{r}_n}{r_n}e^{-\beta C(r_n-r_{n+1})^2}\mathfrak{F}_{\frac{1}{2}}(\chi)\left[\frac{1}{P}e^{-\frac{\beta}{P}V_0(r_n)}\frac{\partial V_0(r_n)}{\partial r_n} - \frac{1}{P}e^{-\frac{\beta}{P}V_1(r_n)}\frac{\partial V_1(r_n)}{\partial r_n}\right] \\ &+ e^{-\beta C(\vec{r}_n-\vec{r}_{n+1})^2}e^{-\frac{\beta}{P}V_1(r_n)} \times \left[-2C(\vec{r}_n - \vec{r}_{n+1}) - \frac{1}{P}\frac{\partial V_1(r_n)}{\partial r_n}\right]\Bigg\}\end{aligned} \tag{2.167}$$

Where $< \ldots >$ is the argument of the Logarithmic function in the expression of V_{eff}. We would like to note that in deriving the above expression the modified Bessel function was used in its integral form [62]:

$$I_\tau(\chi) = \frac{\left(\frac{\chi}{2}\right)^\tau}{\Gamma\left(\tau+\frac{1}{2}\right)\Gamma\left(\frac{1}{2}\right)}\int_{-1}^{1}\left(1-y^2\right)^{\tau-\frac{1}{2}}e^{\pm\chi y}dy$$

where Γ is the familiar Gamma function. In our case this form could be simplified to become: $\mathfrak{F}_{\frac{1}{2}}(\chi) = \frac{e^{-\chi}}{\chi}\sin h\chi$ where we use the fact that $\Gamma(1/2) = \sqrt{\pi}\, and$ $\Gamma(1) = 1$.

The force on the $(n+1)$th node is calculated in a similar manner and given by:

$$\begin{aligned}\vec{F}_{n+1} = &\frac{1}{<\ldots>}\left\{2C(r_n - r_{n+1})\frac{\vec{r}_{n+1}}{r_{n+1}}e^{-\beta C(r_n - r_{n+1})^2} \times \mathfrak{F}_{\frac{1}{2}}(\chi)\left[e^{-\frac{\beta}{P}V_0(r_n)} - e^{-\frac{\beta}{P}V_1 r_n}\right]\right.\\ &+ 2Ce^{-\beta C(r_n - r_{n+1})^2}\frac{\vec{r}_{n+1}}{r_{n+1}}r_n\frac{e^{-\chi}}{\chi}\left[\cos h\chi - \sin h\chi\left(1+\frac{1}{\chi}\right)\right]\left[e^{-\frac{\beta}{P}V_0(r_n)} - e^{-\frac{\beta}{P}V_1(r_n)}\right]\\ &\left.+2Ce^{-\beta C(\vec{r}_n - \vec{r}_{n+1})^2}e^{-\frac{\beta}{P}V_1(r_n)}(\vec{r}_n - \vec{r}_{n+1})\right\}\end{aligned} \tag{2.168}$$

Appendix 4: Exchange Kinetic Energy Estimator for N-Electron System

In this section, we will calculate the exchange kinetic estimator, $\langle KE_{exch}\rangle$, which is given

$$\begin{aligned}KE_{exch} &= \left\langle \frac{\partial}{\partial\beta}\left(\beta V_{eff}^{exch}\right)\right\rangle\\ &= \left\langle \frac{\partial}{\partial\beta}\sum_{s=up}^{down}\sum_{k=1}^{P}\sum_{l=1}^{P}\frac{1}{P^*}\ln\left(\det\left(E^{(k,l)}\right)\right)\theta_{kls}^{+}\right\rangle\\ &= \left\langle \sum_{s=up}^{down}\frac{1}{P^*}\sum_{k=1}^{P}\sum_{l=1}^{P}\frac{1}{\det(E^{(k,l)})}\frac{\partial}{\partial\beta}\det\left(E^{(k,l)}\right)\theta_{kls}^{+}\right\rangle\end{aligned} \tag{2.169}$$

In order to differentiate a determinant, we use the following matrix algebra: if a $N \times N$ square matrix A is function of X, we have

$$\begin{gathered}\det A = \sum_{i=1}^{N} a_{ij}A_{ij} = \sum_{j=1}^{N} a_{ij}A_{ij}\\ (6.143)\\ \frac{\partial}{\partial X}\det A = \sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\partial a_{ij}}{\partial X}A_{ij}\end{gathered} \tag{2.170}$$

where a_{ij} is an element of the matrix A and A_{ij} is a cofactor of the element a_{ij}. Thus if we differentiate $\det\left(E^{(k,l)}\right)$ with respect to β, we have

$$\begin{aligned}\frac{\partial}{\partial\beta}\det\left(E^{(k,l)}\right) &= \sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\partial}{\partial\beta}\left(E^{(k,l)}\right)_{ij}\cdot A_{ij} \\ &= \sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\partial}{\partial\beta}\left[\exp\left\{-\frac{Pm}{2\beta\hbar^2}\left(\left(r_i^{(k)}-r_j^{(l)}\right)^2-\left(r_i^{(k)}-r_i^{(l)}\right)^2\right)\right\}\right]\cdot A_{ij} \\ &= \frac{Pm}{2\beta^2\hbar^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\left(\left(r_i^{(k)}-r_j^{(l)}\right)^2-\left(r_i^{(k)}-r_i^{(l)}\right)^2\right)\left(E^{(k,l)}\right)_{ij}\cdot A_{ij}\end{aligned} \tag{2.171}$$

If we apply (2.170) to the left-hand side of the last term of (2.171), we can write that equation in a simpler form:

$$\frac{\partial}{\partial\beta}\det\left(E^{(k,l)}\right) = \frac{Pm}{2\beta^2\hbar^2}\sum_{i=1}^{N}\det\left(H_i^{(k,l)}\right) \tag{2.172}$$

where $H_i^{(k,l)}$ is an $N\times N$ matrix and its element $\left(H_i^{(k,l)}\right)_{st}$ is given by

$$\left(H_i^{(k,l)}\right)_{st} = \begin{cases} L_{st}^{(k,l)}\exp\left(-\frac{Pm}{2\beta^2\hbar^2}L_{st}^{(k,l)}\right) & \text{if } i=t \\ \exp\left(-\frac{Pm}{2\beta^2\hbar^2}L_{st}^{(k,l)}\right) & \text{if } i\neq t \end{cases} \tag{6.147}$$

and

$$L_{st}^{(k,l)} = \left(r_s^{(k)}-r_t^{(l)}\right)^2-\left(r_s^{(k)}-r_s^{(l)}\right)^2$$

From (2.170) and (2.172), the exchange kinetic estimator can be written as

$$\langle KE_{exch}\rangle = \left\langle\sum_{s=up}^{down}\frac{1}{P^*}\sum_{k=1}^{P}\sum_{l=1}^{P}\frac{1}{\det(E^{(k,l)})}\sum_{i=1}^{N}\frac{Pm}{2\beta^2\hbar^2}\det\left(H_i^{(k,l)}\right)\theta_{kls}^{+}\right\rangle \tag{2.173}$$

Appendix 5: Energy Estimator for Electron in Non-local Pseudo-potential

The energy could be calculated easily using the canonical ensemble. In this case the energy is given by:

$$\langle E \rangle = -\frac{1}{Z}\frac{\partial Z}{\partial \beta} \tag{2.174}$$

Using the partition function given by (2.124) the ensemble average of the energy is:

$$\langle E \rangle = \left(\frac{3P}{2\beta}\right) - \left\langle \sum_{n=1}^{P} \frac{1}{\langle \ldots \rangle} \frac{\partial \langle \ldots \rangle}{\partial \beta} \right\rangle \tag{2.175}$$

Where $< \ldots >$ is the argument of the Ln function in the effective potential of an electron in the non-local pseudo-potential. For the sodium ion, the average energy is then given by the following equation:

$$\begin{aligned}
\langle E \rangle = \left(\frac{3P}{2\beta}\right) - \Bigg\langle \sum_{n=1}^{P} \frac{1}{<\ldots>} \Big\{ & C(r_n - r_{n+1})^2 \times e^{-\beta C(r_n - r_{n+1})^2} \mathfrak{F}_{\frac{1}{2}}(\chi) \left[e^{-\frac{\beta}{P} V_0(r_n)} - e^{-\frac{\beta}{P} V_1(r_n)} \right] \\
& + e^{-\beta C(r_n - r_{n+1})^2} \mathfrak{F}_{\frac{1}{2}}(\chi) \left[\frac{V_1(r_n)}{P} e^{-\frac{\beta}{P} V_1(r_n)} - \frac{V_0(r_n)}{P} e^{-\frac{\beta}{P} V_0(r_n)} \right] \\
& - e^{-\beta C(r_n - r_{n+1})^2} \left[e^{-\frac{\beta}{P} V_0(r_n)} - e^{-\frac{\beta}{P} V_1(r_n)} \right] \left(\frac{\chi}{\beta} \right) \left[\frac{e^{-2\chi}}{\chi} \left(1 + \frac{1}{2\chi} \right) - \frac{1}{2\chi^2} \right] \\
& + C(\vec{r}_n - \vec{r}_{n+1})^2 e^{-\frac{\beta}{P} V_1(r_n)} e^{-\beta C(\vec{r}_n - \vec{r}_{n+1})^2} - \frac{V_1(r_n)}{P} e^{-\frac{\beta}{P} V_1(r_n)} e^{-\beta C(\vec{r}_n - \vec{r}_{n+1})^2} \Big\} \Bigg\rangle
\end{aligned} \tag{2.176}$$

References

1. R. Car, M. Parrinello, Phys. Rev. Lett. **55**, 2471 (9185)
2. A. Selloni, P. Carnevali, R. Car, M. Parrinello, Phys. Rev. Lett. **59**, 823 (1987)
3. M. Parrinello, Solid State Comm. **102**, 107 (1997)
4. R.P. Feynman, A.R. Hibbs, *Quantum Mechanics and Path Integrals* (McGraw-Hill, New York, 1965)
5. M. Parrinello, A. Rahman, J. Chem. Phys. **80**, 860 (1984)
6. D. Marx, M. Parrinello, Z. fur Physik B **95**, 143 (1994)
7. A. Alavi, D. Frenkel, J. Chem. Phys. **97**, 9249 (1992)
8. A. Alavi, J. Kohanoff, M. Parrinello, D. Frenkel, Phys. Rev. Lett. **73**, 2599 (1994)
9. R.W. Hall, J. Chem. Phys. **89**, 4212 (1988)
10. R.W. Hall, J. Chem. Phys. **91**, 1926 (1989)
11. R.W. Hall, J. Chem. Phys. **93**, 5628 (1989)
12. D.M. Ceperley, Phys. Rev. Lett. **69**, 331 (1992)
13. D.M. Ceperley, Rev. Mod. Phys. **67**, 279 (1995)
14. S. Zhang, J. Carlson, J.E. Gubernatis, Phys. Rev. B **55**, 7464 (1997)
15. D. Chandler, P.G. Wolynes, J. Chem. Phys. **74**, 4078 (1981)
16. D.M. Ceperley, J. Stat. Phys. **63**, 1237 (1991)
17. D.M. Ceperley, Phys. Rev. Lett. **73**, 2145 (1994)
18. M. Gell-Mann, K. Bruekner, Phys. Rev. **106**, 364 (1957)
19. D.M. Ceperley, Phys. Rev. B **18**, 3126 (1978)

20. D.M. Ceperley, B.J. Alder, Phys. Rev. Lett. **45**, 566 (1980)
21. R.G. Dendrea, N.W. Ashcrofft, A.E. Carlsson, Phys. Rev. B **34**, 2097 (1986)
22. F. Perrot, M.W.C. Dharma-wardana, Phys. Rev. A **30**, 2619 (1984)
23. S. Tanaka, S. Mitake, S. Ichimaru, Phys. Rev. A **32**, 1896 (1985)
24. Oh Ki-Dong, P.A. Deymier, Phys. Rev. Lett. **81**, 3104 (1998)
25. Oh Ki-Dong, P.A. Deymier, Phys. Rev. B **58**, 7577 (1998)
26. P.A. Deymier, K.-D. Oh, Modell. Simul. MSE **12**, 197 (2004)
27. P.A. Deymier, K.-D. Oh, J. Non-cryst. Solids, **274**, 364 (2000)
28. Oh Ki-dong, P.A. Deymier, Phys. Rev. B **59**, 11276 (1999)
29. Y. Waseda, *The Structure of Non-Crystalline Materials* (McGraw Hill, New York, 1980)
30. R.A. Cowley, A.D.B. Woods, G. Dolling, Phys. Rev. **150**, 487 (1966)
31. P.E. Blochl, M. Parrinello, Phys. Rev. B **45**, 9413 (1992)
32. T. Arima, Y. Tokura, J.B. Torrance, Phys. Rev. B **48**, 17006 (1993)
33. J.S. Ahn, J. Bak, H.S. Choi, T.W. Noh, J.E. Han, Y. Bang, J.H. Cho, Q.X. Jia, Phys. Rev. Lett. **82**, 5321 (1999)
34. N.F. Mott, *Metal-Insulator Transitions*, 2nd edn. (Taylor and Francis, London, 1990)
35. J. Hubbard, Proc. R. Soc. London A **276**, 238 (1963)
36. Oh Ki-Dong, P.A. Deymier, Phys. Rev. B **69**, 155101 (2004)
37. G.E. Jabbour, P.A. Deymier, Modelling and Simulation in MSE **2**, 1111 (1994)
38. R.P. Feynman, Phys. Rev. **76**, 769 (1949); ibid **90**, 1116 (1953); ibid **91**, 1291 (1953); and ibid **91**, 1301 (1953)
39. R.P. Feynman, *Statistical Mechanics* (Benjamin, New York, 1972)
40. L.D. Fosdick, H.F. Jordan, Phys. Rev. **143**, 58 (1966); ibid **171**, 128 (1968)
41. E.L. Pollock, D.M. Ceperley, Phys. Rev. B **30**, 2555 (1985); ibid **36**, 8343 (1987)
42. M.F. Herman, E.J. Bruskin, B.J. Berne, J. Chem. Phys. **76**, 5150 (1982)
43. C.B Alcock, J. Phys. Chem. Ref. Data, **23**, 385 (1994)
44. B.J. Alder, J.E. Wainwright, J. Chem. Phys. **27**, 1208 (1957)
45. B.J. Alder, J.E. Wainwright, J. Chem. Phys. **31**, 456 (1958); ibid, **33**, 1439 (1960)
46. J.B. Anderson, J. Chem. Phys. **63**, 1499 (1975); ibid **65**, 4121 (1976); and ibid **73**, 3897 (1980)
47. A. Muramatsu et al., Int. J. Mod. Phys. C **3**, 185 (1992)
48. S.R. White, Phys. Rev. B **41**, 9031 (1990)
49. N. Metropolis, S. Ulam, J. Am. Stat, Assoc. **44**, 335 (1949)
50. S. Ulam, *A Collection of Mathematical Problems* (Interscience, New York, 1960)
51. R.W. Hall, J. Chem. Phys. **93**, 8211 (1990)
52. H. Kleinert, *Path Integral Quantum Mechanics, Statistics, and Polymer Physics* (World Scientific, Singapore, 1990)
53. S.F. Edwards, Y.V. Gulayaev, Proc. Roy. Soc. Lond. **A279**, 229 (1964)
54. D. Peak, A. Inomata, J. Math. Phys. **10**, 1422 (1969)
55. F. Basco, C.C. Bernido, M.V. Carpio-Bernido, Phys. Lett. A **157**, 461 (1991)
56. M.J. Morauscsik, The two-Nucleon Interaction, Oxford: Clarendon Press, p. 52 (1963)
57. D. Kiang, Amer. J. Phys. **39**, 996 (1971)
58. J.C. Phillips, L. Klienman, Phys. Rev. **116**, 287 (1959)
59. M. Cohen, V. Heine, Phys. Rev. **122**, 1821 (1961)
60. I.V. Abrenkov, V. Heine, Phil. Mag. **12**, 529 (1965)
61. J.M. Ziman, *Principles of the Theory of Solids, Cambridge* (University Press, England, 1972)
62. I.S. Gradshteyn, I.M. Ryzhik, *Table of Integrals, Series and Products* (Academic Press, San Diego, 1992)
63. C.F. Melius, W.A. Goddard III, Phys. Rev. A **10**, 1528 (1974)
64. D.W. Heermann, *Computer simulation methods in Theoretical Physics* (Springer-Verlag, Berlin, 1986)
65. K. Ohno, K. Esfarjani, Y. Kawazoe, Computational Materials Science: From Ab-initio to Monte Carlo Methods, p. 129. In: Springer Series in Solid State Sciences, Berlin (1999)

66. G. Ciccotti, D. Frenkel, I.R. McDonald (eds.), *Simulation of Liquids and Solids: Molecular Dynamics and Monte Carlo Methods in Statistical Mechanics* (North Holland, Amsterdam, 1987)
67. D.C. Rapaport, *The Art of Molecular Dynamics Simulation*, 2nd edn. (Cambridge University Press, Cambridge, 2004)
68. L. Verlet, Phys. Rev. B **159**, 98 (1967)
69. E. Teller, J. Chem. Phys. **21**, 1087 (1953)
70. W.W. Wood, F.R. Parker, J. Chem. Phys. **27**, 720 (1957)
71. F.G. Fumi, M.P. Tosi, J. Phys. Chem. Solids **25**, 31 (1964)
72. P.P. Ewald, Ann. Physik. **64**, 253 (1921)
73. B.R.A. Nijboer, F.W. De Wette, Physica **23**, 309 (1957)
74. S.G. Brush, H.L. Sahlin, E. Teller, J. Chem. Phys. **45**, 2102 (1966)
75. M.J.L. Sangster, M. Dixon, Adv. Phys. **25**, 247–342 (1976)
76. P. Linse, H.C. Andersen, J. Chem. Phys. **85**, 3027 (1986)
77. H.C. Andersen, J. Chem. Phys. **72**, 2384 (1980)
78. S. Nosé, J. Chem. Phys. **81**, 511 (1984)
79. G.J. Marthyna, M.L. Klein, J. Chem. Phys. **97**, 2635 (1992)
80. M.E. Herman, E.J. Bruskin, B.J. Berne, J. Chem. Phys. **76**, 5150 (1982)
81. M.J.L. Sangster, R.M. Atwood, J. Phys. C **11**, 1541 (1978)
82. W.A. Harrison, *Pseudopotentials in the Theory of Metals* (Benjamin, New York, 1966)
83. Y. Waseda, *The Structure of Non-Crystalline Materials* (McGraw-Hill, New York, 1980)
84. C. Kittel, *Elementary Statistical Physics* (Wiley, New York, 1967)
85. J. Theilhaber, Phys. Fluids B **4**, 2044 (1992)
86. S. Kambayashi, J. Chihara, Phys. Rev. E **53**, 6253 (1996)
87. W.R. Magro, B. Militzer, D.M. Ceperley, B. Bernu, C. Pierleoni, in *Strongly Coupled Coulomb Systems*, ed. by G.J. Kalman, J.M. Rommel, K. Blagoev. Plenum Press, New York, (1998)
88. W.R. Magro, B. Militzer, D.M. Ceperley, B. Bernu, C. Pierleoni, in *Strongly Coupled Coulomb Systems*, ed. by G.J. Kalman, J.M. Rommel, and K. Blagoev, Ed., Plenum Press, New York, (1998)
89. B. Militzer, E.L. Pollock, Phys. Rev. E **61**, 3470 (2000)
90. G. Ortiz, P. Ballone, Phys. Rev. B **50**, 1391 (1994)
91. J. Schlipf, M. Jarrell, P.G.J. van Dongen, N. Blumer, S. Kehrein, Th Pruschke, D. Vollhardt, Phys. Rev. Lett. **82**, 4890 (1999)
92. G. Kotliar, E. Lange, M. Rozenber, Phys. Rev. Lett. **84**, 5180 (2000)
93. N.F. Mott, Philos. Mag. **6**, 287 (1961)
94. W. Kohn, C. Majumdar, Phys. Rev. **138**, A1617 (1965)
95. M. Rozenberg, G. Kotliar, X.Y. Zhang, Phys. Rev. B **49**, 10181 (1994)
96. A. Georges, G. Kotliar, W. Krauth, M.J. Rozenberg, Rev. Mod. Phys. **68**, 13 (1996)
97. G. Kotliar, J. Low Temp. Phys. **126**, 1009 (2002)
98. Ning-Hua Tong, Shun-Qing Shen, Pu Fu-Cho, Phys. Rev. B **64**, 235109 (2001)
99. R. Bulla, T.A. Costi, D. Vollhardt, Phys. Rev. B **64**, 045103 (2001)
100. F. Gebhard, *The Mott Metal-Insulator Transition: Models and Methods* (Springer, Berlin, 1997)
101. X.P. Li, J.Q. Broughton, J. Chem. Phys. **86**, 5094 (1987)
102. G.B. Bachelet, D.R. Hamann, M. Schluter, Phys. Rev. **26**, 4199 (1982)
103. C.F. Melius, PhD Dissertation, California Institute of Technology, 136 Pasadena, CA 91109 (1972)
104. J. Cao, J. Voth, J. Chem. Phys. **99**, 10070 (1993)
105. J. Cao, J. Voth, J. Chem. Phys. **100**, 5093 (1994)
106. J. Cao, J. Voth, J. Chem. Phys. **100**, 5106 (1994)
107. J. Cao, J. Voth, J. Chem. Phys. **101**, 6157 (1994)
108. J. Cao, J. Voth, J. Chem. Phys. **101**, 6168 (1994)
109. M. Ceriotti, J. More, D.E. Manolopoulos, Chem. Phys. Comm. **185**, 1019 (2014)

Chapter 3
Interatomic Potentials Including Chemistry

S.M. Valone, Krishna Muralidharan and Keith Runge

Abstract Beginning from two theories, classical and quantum mechanical, as realized in terms of Newton's second law and the time-independent Schrödinger equation, we put forth a framework for understanding the development of atomistic potentials that include chemistry. Our analysis introduces, explains, and exploits the Fragment Hamiltonian approach to the electronic structure of molecular and condensed matter systems. Illustrations of the Fragment Hamiltionian display the roles of various physical concepts in the formation of these atomistic potentials. Electron density fluctuations are clearly seen as essential to the realistic description of interatomic interactions over a large range of nuclear (ionic) configurations. Finally, we present a novel approach to the parameterization of interatomic potentials that explicitly include the effect of charge fluctuations, the environment-dependent dynamic charge potential.

3.1 Background

Molecular dynamics (MD) methods, as mentioned in Chap. 2, have seen numerous applications to the atomistic simulations. Two theoretical concepts that have guided the modeling of atomistic simulation are quantum materials theory and Newtonian dynamics. In quantum theory, the Born-Oppenheimer (B-O) approximation is

S.M. Valone (✉)
Materials Science and Technology Division, Los Alamos National Laboratory, Los Alamos, NM 87545, USA
e-mail: smv@lanl.gov

K. Muralidharan · K. Runge
Department of Materials Science and Engineering, University of Arizona, Tucson, AZ 85721, USA
e-mail: krishna@u.arizona.edu

K. Runge
e-mail: krunge@u.arizona.edu

P.A. Deymier et al. (eds.), *Multiscale Paradigms in Integrated Computational Materials Science and Engineering*, Springer Series in Materials Science 226,
DOI 10.1007/978-3-319-24529-4_3

frequently chosen as a developmental framework. This approximation, sometimes called the 'clamped nuclei' approximation, takes the electronic wave function to depend parametrically on the nuclear configuration at any given time. The B-O approximation decouples the quantum degrees of freedom (the electrons) from the (largely) classical degrees of freedom (the nuclei). Another approach to understanding the B-O approximation will be described later in which distinguishing quantum electrons and classical nuclei can be done without approximation using the eikonal representation. In this representation, the B-O approximation results from the neglect of some terms in the full electron-nuclear coupling.

Why do we need an atomic scale theory? The reason is that one is often interested in the larger scale phenomena associated with that scale. Resolution of the changing states of the electrons with the dynamics is not always necessary. Frequently, at the atomistic scale, the dynamics is classical and can be described by Newton's equations.

In Newtonian dynamics, the atoms, now considered nuclei with associated electrons, interact with one another through interatomic potentials, that is the ground-state *E's* from (3.4). The gradient of the interatomic potential provides forces $\vec{F} = -\nabla E$ that accelerate the atoms by Newton's second law, $\vec{F} = m\vec{a}$, where m is the mass of some atom and $\vec{a}$ is its acceleration. Modeling of the interatomic potential is the challenge for realizing Newtonian dynamics in the context of MD simulations. Simple potentials, like the Lennard-Jones 6–12 potential, incorporate knowledge of long-range and short-range interactions, but are seldom appropriate for the description of chemical or material interactions.

More broadly, an interatomic potential describes the set of interactions among the atoms in a molecule or material. However, at the heart of an interatomic potential are the *electronic states of the atoms*. These states are not the same as the ground states of these same atoms when isolated. It is the interactions among these atoms in their altered electronic states that hold our central focus at this theoretical scale. It is this point of view—electronic states of atoms—that makes the present approach to interatomic potentials unique and a tuned to chemical environment.

The importance of the states of atoms is brought out in the common practice of discussing states of atoms in a material or molecule. Take for instance the use of oxidation states or valence states. One example of among many is captured in an STM-EELS spectrum in Fig. 3.1. An Fe crystallite contacts a TiO_2 crystallite. The spectra indicate that the valence state of the Ti is being modulated by the presence of the Fe crystallite [1] in the mixed region of the interface and for some distance away from the mixed region.

While we may often talk about states of atoms, the way that atomistic properties are computed from first principles is quite different. Most often, in the present day and age, electronic properties of electrons are calculated from an effective one-electron Hamiltonian and then the properties of atoms are deduced through a variety of methods. In another usage, the states of atoms are often translated into the properties of its highest occupied orbital, again a one-electron property. This progression and the alternative presented here are detailed in this chapter.

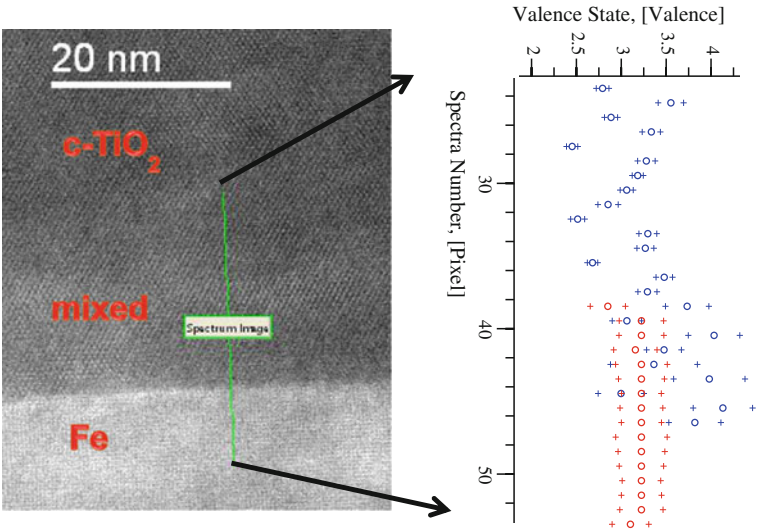

Fig. 3.1 STM-EELS spectrum of a Fe–TiO_2 interface. The spectra are interpreted with the help of electronic structure theory into valence states for Ti (*red*) and Fe (*blue*)

Consequently, our plan for the development of this chapter is a top-down one. We start with some basic quantum mechanics related to materials, develop our general notion of interatomic potentials, and then look back at the vast literature of interatomic potentials to see how many of these concepts can be understood within the general notion advanced here. The more common approach of developing atomistic models from one-electron, electronic structure methods, such as DFT or tight binding (TB), is not pursued here. A many-electron approach is enlisted instead. After all, atoms are many-electron entities.

To reiterate from the earlier chapters, the full Hamiltonian $\widehat{H}$ for the material or chemical system in a spatial representation can be written as the decomposition:

$$\begin{aligned}\widehat{H} = &\sum_i \left(-\frac{\Delta_i}{2} - \sum_A \frac{Z_A}{|r_i - R_A|} + \frac{1}{2}\sum_{j\neq i} \frac{1}{|r_i - r_j|} \right) \\ &+ \sum_A \left(-\frac{\Delta_A}{2m_A} + \frac{Z_A}{2}\sum_{B\neq A} \frac{Z_B}{|R_A - R_B|} \right)\end{aligned} \tag{3.1}$$

The first term in the sum over i is the kinetic energy operator for electron i. The next term is the Coulombic operator between electron i at position $\vec{r}_i$ and each nucleus A at position $\vec{R}_A$. The third term is the Coulombic operator between electron i and another electron j. The second line of (3.1) is composed of the kinetic energy and Coulomb operators for nucleus A and interactions between A and another nucleus B.

It will often be more convenient to use a representation-free notation for these operators. One can also write, in the same order as in (3.1),

$$\widehat{H} = \widehat{T}_e + \widehat{V}_{eN} + \widehat{V}_{ee} + \widehat{T}_N + \widehat{V}_{NN} \tag{3.2}$$

$$\widehat{H} = \widehat{H}_{el} + \widehat{T}_N \tag{3.3}$$

where the five terms in the Hamiltonian represent the electronic kinetic energy, the nuclear-electron electrostatic attraction, the electron-electron electrostatic repulsion, the nuclear kinetic energy and the electrostatic nuclear repulsion, respectively. For convenience of notation, Planck's constant, ħ, and the electron charge have been set equal to unity (i.e. this and all other equations in this chapter are written in atomic units). $\widehat{H}_{el}$ is defined to contain all of the terms other than the nuclear kinetic energy.

Over the last 40 years, a number of researchers and groups have established an approach to factoring the wavefunction of the full chemical system, that is including both nuclei and electrons, to elucidate those terms in the electronic potential for the nuclei which are neglected in the Born-Oppenheimer approximation [2–6]. While each of these groups has their own approaches and purposes, the resulting potential takes on the same form. To obtain the result, let us consider eigenvalue problem from (3.3) using the full Hamiltonian $\widehat{H}$

$$\widehat{H}\Psi\left(\left\{\vec{R}_N, \vec{r}_i\right\}\right) = E\Psi\left(\left\{\vec{R}_N, \vec{r}_i\right\}\right) \tag{3.4}$$

and choose a total wavefunction of the form;

$$\Psi\left(\left\{\vec{R}_N, \vec{r}_i\right\}\right) = \exp\left[iS\left(\left\{\vec{R}_N\right\}\right)\right]\psi\left(\left\{\vec{r}_i\right\}, \left\{\vec{R}_N\right\}\right) \tag{3.5}$$

where $\{\ldots\}$ indicates the collection of all coordinates of the nuclei or electrons. This has been referred to as the 'eikonal' representation in [4]. Substituting (3.5) into (3.4), we obtain

$$\left[\sum\nolimits_{N=1}^{N_n} \frac{1}{2M_N}\left(\frac{1}{i}\frac{\partial}{\partial\vec{R}_N + \frac{\partial S}{\partial\vec{R}_N}}\right)^2 + \widehat{H}_{el} - E\right]\psi = 0 \tag{3.6}$$

where M_N is the mass of the Nth nucleus and N_n is the total number of nuclei. Taking S to be real and projecting (3.6) on ψ and taking the real part, we arrive at an equation for S

$$\sum\nolimits_{N=1}^{N_n} \frac{1}{2M_N}\left[\frac{\partial S}{\partial\vec{R}_N}\right]^2 + V_{qu}\left[\frac{\partial S}{\partial\vec{R}_N}, \vec{R}\right] = E;\ V_{qu} = V + V' + V'' \tag{3.7}$$

$$V = \frac{\langle \psi | \widehat{H}_{el} | \psi \rangle}{\langle \psi | \psi \rangle} \tag{3.7a}$$

$$V' = \sum\nolimits_{N=1}^{N_n} \frac{1}{2M_N} \frac{\partial S}{\partial \vec{R}_N}^{\circ} \left[\langle \frac{\partial \psi}{\partial \vec{R}_N} | \psi \rangle - \langle \psi | \frac{\partial \psi}{\partial \vec{R}_N} \rangle \right] / \langle \psi | \psi \rangle \tag{3.7b}$$

$$V'' = \sum\nolimits_{N=1}^{N_n} \frac{-1}{2M_N} \frac{1}{2} \left[\langle \frac{\partial^2 \psi}{\partial \vec{R}_N^2} | \psi \rangle + \langle \psi | \frac{\partial^2 \psi}{\partial \vec{R}_N^2} \rangle \right] / \langle \psi | \psi \rangle \tag{3.7c}$$

If we were to set V' and V'' to zero, we would recover the potential that comes from the B-O approximation. In this sense, these two pieces of the quantum potential, V_{qu}, represent non-Born-Oppenheimer contributions.

It is the quantum potential V_{qu}, that interatomic potentials are approximating, and we note that for any configuration of the nuclei, there is a solution to (3.7a), while the contributions to the quantum potential from (3.7b) and (3.7c) are dependent on the details of the nuclear trajectories. For simplicity, in the remainder of this chapter we will examine the characteristics of the electronic Hamiltonian $\widehat{H}_{el}$ which relates to the nuclear potential in the B-O approximation. The set of ground-state energy solutions of $\widehat{H}_{el}$ for all nuclear configurations is often called an energy landscape or an interatomic potential.

The field of quantum chemistry can be understood as a series of modeling efforts that allow one to find approximate eigensolutions (i.e. eigenvalues plus eigenvectors or eigenfunctions) to the time-independent Schrödinger equation with the electronic Hamiltonian $\widehat{H}_{el}$. Quantum chemical models can be characterized broadly as either 'wave function' methods or 'density functional' methods. This chapter is framed in terms of wave function methods, but can also be related to the density functional theory (DFT). Of course, one does not know exactly what the interatomic potential is for any material, much less materials undergoing some change through a chemical process. This chapter presents considerations that should be taken into account when approximating the interatomic potential for any material.

Consider (3.7a) and let $|\psi>$ be treated as a trial state instead of an eigenfunction. This approach further simplifies the problem of estimating ground-state values for E. This is useful for atomistic potentials because ordinarily one takes information or properties from specific ground-state configurations of the system of interest to build an interatomic potential. One then interpolates or extrapolates from these known configurations to estimates for E in configurations of the system for which no direct information is known from other sources. In this strategy, one does not possess the ground-state eigenfunctions for the vast majority of atomic configurations.

The Fragment Hamiltonian, presented here, serves as a theoretical route toward informing interatomic potentials for atomistic models. This is quite different from the empirical considerations that have frequently been used in the development of potentials, such as the Lennard-Jones 6–12 for various materials.

3.2 The Fragment View of Hamiltonians for Materials and Molecules

To provide a unified framework for most interatomic potentials, we need to adopt a starting point that differs somewhat from the traditional one embodied in (3.7a), DFT methods, or state decomposition methods. Any of these points of view are based on operators that treat a system of interest as a single entity. Equation (3.4) poses the very general problem of many-electron physics and chemistry, although a decomposition that separates kinetic, external potential, and electron repulsion operators is invoked. This approach is far too complex for even the simplest chemical and materials problems that we wish to address.

Any implementation of DFT or Hartree-Fock theory simplifies the problem by decomposing or "fragmenting" a system into individual electrons. There is a one-electron Hamiltonian for each electron. These one-electron approximations generally introduce nonlinearities and discontinuities that do NOT appear in either (3.4) or (3.7a). In spite of the nonlinearities, one-electron approximations do achieve a satisfactory level of simplification in many, many instances. With modern computers, one can often determine forces needed in solving Newton's equations of motion directly from the one-electron methods. These sorts of compromises seem to be a common feature of describing a total system/material in terms of the properties of its constituents—we shall see that, in the approach to interatomic potentials to be pursued here, nonlinearities and discontinuities also appear. However, if a one-electron view is taken as a starting point for an atomistic model, then the one-electron theory or model has to be coarsened back to the atomic scale. The nonlinearities of the one-electron theory present a problem to any coarsening procedure as a result.

The third strategy focuses on decomposing the state or states of a system into constituents for each fragment. By "state decomposition," we mean that the state of the total system is decomposed into states pertaining to each atom. Most often the state of a system is adequately represented by its total electron density (or "density" for brevity), and it is ρ that is decomposed [7–9]. That is, one assumes that one can find ρ_A for each fragment A such that $\rho = \Sigma_A \rho_A$. One popular strategy for decomposing a total density scales that density by the Hirshfeld partition [10–12] to generate densities ρ_A for each fragment A. In the Hirshfeld decomposition, the density ρ is scaled by the ratio of the isolated fragment A density $\rho_A^{(0)}$ to the, sum of the isolated fragment densities for all of the fragments in the system $\rho_0 = \Sigma_B \rho_B^{(0)}$, with the sum being over all fragments in the system. Thus, ρ_A is defined as $\rho_A = (\rho/\rho_0)\rho_A^{(0)}$. While important and completely legitimate, this strategy presupposes knowledge of ρ.

Still some simplification or other is most certainly required. A better choice, in our minds, for atomistic modeling is based on atoms-in-molecules (AIM) theory, developed by Moffitt [13] and others [14–17]. We will call our broad framework the Fragment Hamiltonian (FH) approach [18]. The approach requires unification of

three essential components: (1) certain characteristics of many-electron wave functions, (2) decomposition of many-electron Hamiltonians into fragments, and (3) variables for the charge states of each fragment. The charge-state variables pertain to the states of the fragments and create the connection back to the many-electron wave functions and Hamiltonian from which the FH originates. Before developing the FH, a series of topics are reviewed briefly: Subsystems, electron indistinguishability, averages over fragments, and the concept of atoms-in-molecules for interacting systems.

To address these issues, imagine two or more atoms infinitely separated from one another. The many-electron Hamiltonian for each atom is agreed upon. Coulombic interactions among the atoms are not considered important. No one would include these terms in the Hamiltonian of (3.1). Instead we choose one of the atoms, say atom A, and write its Hamiltonian as

$$\widehat{H}_A = \sum\nolimits_{i_A} \left(-\frac{\Delta_{i_A}}{2} - \frac{Z_A}{\left|\vec{r}_{i_A} - \vec{R}_A\right|} + \frac{1}{2}\sum\nolimits_{j_A \neq i_A} \frac{1}{\left|\vec{r}_{i_A} - \vec{r}_{j_A}\right|} \right) \tag{3.8}$$

where now the sums are just over some number of electrons N_A that we have assigned to A. It is not necessary to specify which of the electrons is assigned to A, just the number. We will say more about N_A later. Thus, what is meant by an "atom" is broadly agreed upon. A similar argument applies to a very dilute gas of molecules, where the atoms in a single molecule interact among one another, but not with the atoms from any other molecule. If we have just two, infinitely-separated atoms in our system, the total electron Hamiltonian $\widehat{H}_{AB}^{\infty}$ is just

$$\widehat{H}_{AB}^{\infty} = \widehat{H}_A + \widehat{H}_B \tag{3.9}$$

where each component Hamiltonian has the structure of (3.8). Next we amplify on the issue of electron indistinguishability. The idea of identifying atoms or groups of atoms in a material might seem to clash with the indistinguishability of electrons. Indistinguishability of electron is not an issue because expectation values over all of the electrons are all that is needed in the present approach. For instance, it is common practice to recognize that the electron kinetic energy and electron-nuclear interactions are one-electron operators. For a single electron-nuclear interaction, one may write the expectation value,

$$\left\langle -\frac{Z_A}{\left|\vec{r} - \vec{R}_A\right|} \middle| \rho(\vec{r}) \right\rangle \tag{3.10}$$

as a contribution to the external potential energy for an atom A with nuclear charge Z_A at position $\vec{R}_A$, and where $\vec{r}$ is an electron coordinate and $\rho(\vec{r})$ is the electron

density for the system. Of course, one can retreat to the N-electron state $|\psi(N)\rangle$ from which $\rho(\vec{r})$ is derived and express the same quantity as

$$N\left\langle \psi(N)\left|-\frac{Z_A}{|\vec{r}-\vec{R}_A|}\right|\psi(N)\right\rangle \tag{3.11}$$

which is also

$$\sum_{i=1}^{N}\left\langle \psi(N)\left|-\frac{Z_A}{|\vec{r}_i-\vec{R}_A|}\right|\psi(N)\right\rangle \tag{3.12}$$

where the sum is over all electrons in the system. Indistinguishability of electrons is not an issue. The same arguments apply to other subsystem averages for the kinetic energy of an electron and external potential. By extension, one can also calculate expectation values for the two electron operators needed for $\widehat{V}_{ee}$ in (3.2). Similar statements hold for the kinetic and electron-electron repulsion integrals. Consequently, one can assemble the energy for a subsystem or fragment from combinations of these quantities and indistinguishability is not an issue.

To this end, suppose that some subsystem S has Hamiltonian $\widehat{H}_S$ with N_S electrons. Further suppose that S is part of a larger system of N electrons. Again denote an admissible, anti-symmetric, normalized state of the total system as $|\psi(N)\rangle$. Generalizing Rychlewski and Parr [19], the corresponding energy of S with respect to the state $|\psi(N)\rangle$ may be taken as the expectation value of subsystem Hamiltonian $\widehat{H}_S$,

$$E_S(N_S)=\left\langle \psi(N)\middle|\widehat{H}_S\middle|\psi(N)\right\rangle \tag{3.13}$$

This type of subsystem average will be applied extensively. Clearly $\widehat{H}_S$ could be a Hamiltonian for an atom or it could be some collection of atoms.

Finally, understanding interactions between atoms can be approached again by considering atoms infinitely separated from each other. Now let just two of those atoms, A and B, form a diatomic molecule. None of the other atoms are allowed to interact with each other or with the diatomic. The diatomic Hamiltonian $\widehat{H}_{AB}$ is a special case of (3.1), but with two nuclei, A and B, in the sum. Compared to the sum of Hamiltonians for separated A and B in (3.9), $\widehat{H}_{AB}$ now has a Coulombic interaction between atoms A and B and their electrons,

$$\widehat{V}_{AB}=\widehat{H}_{AB}-\widehat{H}_{AB}^{\infty} \tag{3.14}$$

an interaction operator to be explained in more detail later. All of the other atoms have Hamiltonians like (3.8). Everything about the FH is comprised of simple expansions of these basic ideas.

3.3 A Different Way to Decompose a System

The present point of view focuses on decomposing $\widehat{H}_{el}$ into atomic-scale components. The basis for choosing this strategy is that, in atomistic models, the electronic states of the *atoms* hold the position of central importance. This view was formalized in the (AIM) theory of Moffitt [13]. The essential observation behind this point of view is that the Hamiltonians of atoms in the gas phase *substantially* survive when the same atoms appear in a molecule or solid. We say "substantially" because there are some caveats to this statement, as will be explained shortly.

More generally, arbitrary groupings of atoms can be regarded as the basic units or building blocks for a particular system, and we will call these *fragments*. This decomposition of $\widehat{H}_{el}$ is what we will mean by the Fragment Hamiltonian (FH). There are several reasons for choosing a different term than "AIM" to describe the approach. One reason is to recognize the generality of the decomposition. Another reason is that the states of the system are specified somewhat differently. Still another is that we want to construct a general form of potential energy surface from the underlying theory, rather than a route to molecular scale calculations. Finally, the traditional properties of electron affinity and chemical hardness or Hubbard U are recovered in the model, providing a link to many other models.

In AIM theory, Moffitt and others [13–17] recognize that atoms can be in various charge states. As a result the number of electrons associated with each atom can vary. That variation should be tied to the state that is acting on $\widehat{H}_{el}$. Olson and Garrison [17], for instance, use this flexibility to introduce an anionic state for the W_2 pair in their treatment of sodium scattering from tungsten surfaces.

In the FH, these notions are carried further in a several ways, to produce a point of view with several novel features. At the current state of development in the FH approach, only a few key properties of each possible state of the system are specified. Other properties are not. In particular, the electronic states of the electrons (i.e. the orbitals) are not resolved as in most other AIM-based efforts.

To understand the underlying observation of decomposability of $\widehat{H}_{el}$, first consider two neutral atoms, A and B, in the gas phase, well separated from each other, so that there is no interaction. See top panel of Fig. 3.1. To simplify further, the two atoms are allowed to be the same kind of atom. With the two atoms neutral by assumption, the number of electrons, N_{A}, on any one atom equals Z_{A}, the nuclear charge of A. Each atom is described by their gas-phase Hamiltonian, $\widehat{H}_A$ and $\widehat{H}_B$, respectively, each being a copy of (3.8). Without any interactions, the total Hamiltonian for these two atoms is the sum of the Hamiltonians for the isolated atoms, $\widehat{H}_A(N_A = Z_A) + \widehat{H}_B(N_B = Z_B)$.

The same picture holds for an arbitrary number of isolated atoms. The total Hamiltonian $\widehat{H}^{(f)}$ is just the sum of the Hamiltonians for all of atoms:

$$\widehat{H}^{(f)} = \sum\nolimits_A \widehat{H}_A(N_A = Z_A) \tag{3.15}$$

where N_A is the number of electrons on A with nuclear charge Z_A. No approximations have been made at this point, given the stipulation that the atoms are not interacting with each other.

Each Hamiltonian fragment is the same operator as that of the fragment in isolation, with one exception to be explained below. Each Hamiltonian fragment $\widehat{\boldsymbol{H}}_{\boldsymbol{A}}$ has the same structure as $\widehat{\boldsymbol{H}}_{\boldsymbol{el}}$. It is the sum of kinetic energy, electron-nuclear, and electron-electron operators for its electrons,

$$\widehat{H}_A = \widehat{T}_A + \widehat{V}_A^{en} + \widehat{V}_A^{ee} \tag{3.16}$$

where

$$\widehat{T}_A = \sum\nolimits_{j \in electrons\ of\ A} \widehat{T}_j \tag{3.17a}$$

$$\widehat{V}_A^{en} = \sum\nolimits_{j \in electrons\ of\ A} \sum\nolimits_{K \in nuclei\ in\ A} \widehat{V}_{jK}^{en} \tag{3.17b}$$

and

$$\widehat{V}_A^{ee} = \frac{1}{2} \sum\nolimits_{j \in electrons\ of\ A} \sum\nolimits_{k \in electrons\ of\ A \neq j} \widehat{V}_{jk}^{ee} \tag{3.18}$$

It would seem that electron indistinguishability would pose a problem for the FH method. However, it does not. Electron indistinguishability is accounted for through the properties of the wave functions or states in (3.4). For Fragment Hamiltonians, it is perfectly acceptable to order the number of electrons such that electrons 1 through N_A to A and electrons $N_A + 1$ through N to B. Taking the kinetic energy operators as an example, for A,

$$\widehat{T}_A(N_A) = \sum\nolimits_{j=1}^{N_A} \widehat{T}_j \tag{3.19}$$

and, for B,

$$\widehat{T}_B(N_B) = \sum\nolimits_{j=N_A+1}^{N} \widehat{T}_j \tag{3.20}$$

Similar expressions can be constructed for the other operators in the fragment Hamiltonian. The details of the numbering are not important here. What is important is the number of electrons assigned to a given fragment.

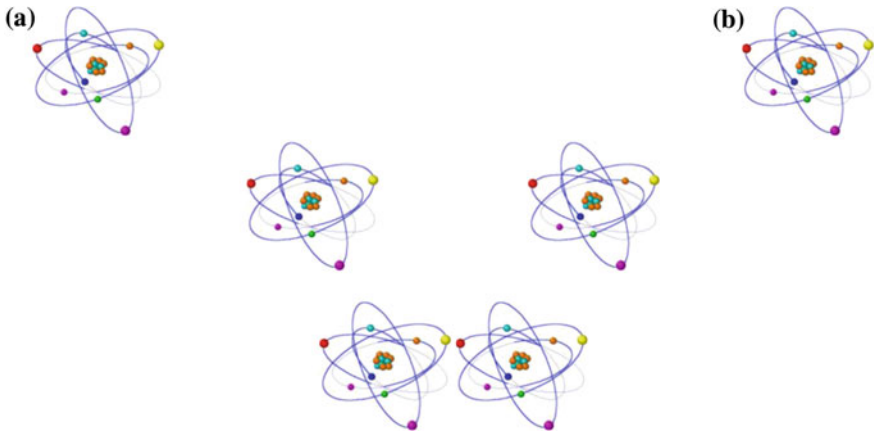

Fig. 3.2 Two isolated atoms labeled "*A*" and "*B*". In the text we call them Ni atoms for specificity. The pictures represent the Hamiltonians, $\widehat{H}_A$ and $\widehat{H}_B$, for the atoms as well. Electrons require large amounts of energy (relative to the energies of chemical bonds) to transfer (or hop) from one atom to the other. *Top* In the gas phase, transfer of charge would produce a type of plasma state between the two atoms. *Middle* Two atoms, beginning to approach each other, interact. All three components of the total Hamiltonian become important, two Hamiltonian fragments $\widehat{H}_A$ and $\widehat{H}_B$, as well as the Coulombic interactions between the two fragments $\widehat{V}_{AB}$. *Bottom* Two atoms in close enough proximity to begin interacting strongly with each other. Electrons begin to hop from atom center to atom center. The number of electrons on each atom is undefined, but we will limit the range

To illustrate (3.15), consider two, well-separated hydrogen atoms. The Hamiltonian fragments can be written out explicitly for this case. The Hamiltonian decomposition, using the notation from Fig. 3.2 is

$$\widehat{H}^{(f)} = \widehat{H}_A(N_A = 1) + \widehat{H}_B(N_B = 1) \tag{3.21}$$

where N_A and N_B are the number of electrons assigned to the two atoms, respectively. The individual contributions consist of Hamiltonian fragments

$$\widehat{H}_A(N_A = 1) = -\frac{1}{2}\Delta_{\vec{r}'} - \frac{1}{\left|\vec{r}' - \vec{d}_A\right|} \tag{3.22}$$

and

$$\widehat{H}_B(N_B = 1) = -\frac{1}{2}\Delta_{\vec{r}''} - \frac{1}{\left|\vec{r}'' - \vec{d}_B\right|} \tag{3.23}$$

where $\vec{r}'$ and $\vec{r}''$ are the electronic coordinates of the two electrons, and the $\vec{d}_A$ and $\vec{d}_B$ are the positions of to the two atoms, respectively. The assignment of the

electrons with spatial position designated by $\vec{r}'$ to the left-hand atom and $\vec{r}''$ to the right-hand atom, respectively, is arbitrary. The assignments could be revised. Only the number of electrons is important: Each atom is neutral at this point in the development.

To motivate this next step, let us return to our two-atom example. If the two atoms are allowed to approach each other (middle panel, Fig. 3.2), the electrons of one atom start to interact with the electrons and nucleus of the other through Coulombic forces. The opposite interactions occur as well. The Hamiltonians of the isolated atoms are preserved at this stage, but new interactions begin to contribute appreciably as the result of the approach of the atoms to each other. These interactions are captured in the Coulombic operator $\widehat{V}_{AB}(Z_A, Z_B)$. The numbers of electrons for the two interacting atoms are designated in parenthesis. All of the kinetic energy operators are retained in $\widehat{H}_A$ and $\widehat{H}_B$.

Again the picture holds for an arbitrary number of atoms beginning to interact with each other. Based on this picture, the total Hamiltonian of (3.3) can be reformulated as

$$\widehat{H}^{(f)} \equiv \sum\nolimits_A \left(\widehat{H}_A(Z_A) + \frac{1}{2} \sum\nolimits_{B \neq A} \widehat{V}_{AB}(Z_A, Z_B) \right) \tag{3.24}$$

where A is the index for the nuclear positions, just as in (3.15). This decomposition $\widehat{H}^{(f)}$ is the Fragment Hamiltonian or FH form. It has the $\widehat{H}_A$ for each atom A, which is the same Hamiltonian as in the gas phase. Each of these atoms interacts with all of the other atoms, as represented by the operators $\widehat{V}_{AB}$. The Coulombic interactions in $\widehat{V}_{AB}$ are only those that occur between electrons of A and the nuclei of B and electrons of B, between electrons of B and electrons and nuclei of A, and between the nuclei of A and B. These Coulombic terms account for all of the interaction operators not appearing in the Hamiltonians of the fragment themselves. In the end, formulation of the FH model is a bookkeeping exercise. All of the expressions to this point are exact.

We again illustrate the above with the H_2 molecule. The total Hamiltonian has the form

$$\widehat{H}^{(f)} = \widehat{H}_A(N_A = 1) + \widehat{H}_B(N_B = 1) + \widehat{V}_{AB}(N_A = 1; N_B = 1) \tag{3.25}$$

The expressions for the Hamiltonian fragments $\widehat{H}_A$ and $\widehat{H}_B$ from (3.21) remain valid. Now the Coulombic interaction must be added. Under our assumption that the two fragments are neutral, the fragment-fragment interaction may be written as

$$\widehat{V}_{AB}(N_A = 1; N_B = 1) = -\frac{1}{\left|\vec{r}' - \vec{d}_B\right|} - \frac{1}{\left|\vec{r}'' - \vec{d}_A\right|} + \frac{1}{\left|\vec{r}' - \vec{r}''\right|} + \frac{1}{\left|\vec{d}_A - \vec{d}_B\right|} \tag{3.26}$$

Notice that the electron labeled $\vec{r}'$ interacts with nucleus B, while the electron labeled $\vec{r}''$ interacts with nucleus A.

The example of a dimer, with the two atoms interacting, begs the question of how many electrons to assign to each fragment. Let us return one more time to our two-atom example, but now these two atoms are allowed to approach closely enough to interact strongly and form a dimer (bottom panel, Fig. 3.1). By "strongly" we mean that the interaction between the two atoms can promote electrons from one atom to hop to the other and vice versa. Now the number of electrons on each atom need not be equal to its nuclear charge.

The hops of electrons between atomic centers initiate two effects. First, at the level of the Hamiltonians for the two atoms, the number of electrons on one atom N_A is not always equal to Z_A. That is, the number of electrons on each atom fluctuates from time to time at the level of the Hamiltonians (see Fig. 3.3). We will say more about what we mean by this in a little while. Second, new ionic states of the dimer of A atoms, that were not accessible when the atoms were isolated or weakly interacting, now contribute to the energy. These states will be defined and described in the next subsection.

As a more general example, suppose the total system consists of two fragments, A and B. The total Hamiltonian is $\widehat{H}^{(f)} = \widehat{H}_A + \widehat{H}_B + \widehat{V}_{AB}$. Assume that the system is neutral, so that $N = Z_A + Z_B$, where N is the total number of electrons and Z_A and Z_B are the *total* nuclear charges on fragments A and B, respectively. Even though our immediate interests in this Chapter are atomistic in nature, the ideas behind the FH approach apply to arbitrary decompositions of molecules or materials into substituent pieces or fragments. The two fragments here might be two coins, a bicrystal, or a water molecule and a semiconductor surface. Since there are only two fragments in this example, $\widehat{V}_{AB}$ always involves all of the electrons and nuclei. However, $\widehat{V}_{AB}$ has only the those Coulombic interactions not already accounted for in either $\widehat{H}_A(N_A)$ or $\widehat{H}_B(N_B)$. As in the earlier examples, $\widehat{H}_A(N_A)$ is the Hamiltonian for isolated A with N_A electrons and $\widehat{H}_B(N_B)$ is the Hamiltonian for isolated B with N_B electrons. To maintain neutrality, it is only necessary that $N = N_A + N_B$. It is *not* necessary that $N_A = Z_A$ at all times.

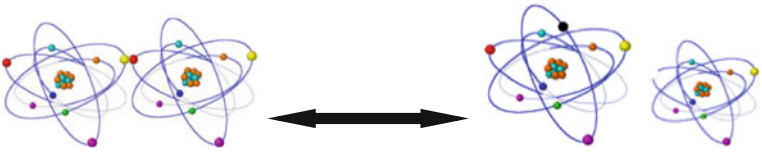

Fig. 3.3 A pair of atoms reacting to form a cation-anion pair by a charge-transfer process. The process is identified in different disciplines as a charge-separation reaction in some areas of materials science and in chemistry, or electron-hole production in other areas of materials science and in condensed-matter physics

As a very specific example, we again appeal to the H_2 molecule, with the atomic sites labeled by A and B, as in the bottom panel of Fig. 3.2. Suppose that a charge fluctuation occurs so that both electrons are associated with atom A. In that case, the Fragment Hamiltonian reads

$$\widehat{H}^{(f)} = \widehat{H}_A(N_A = 2) + \widehat{H}_B(N_B = 0) + \widehat{V}_{AB}(N_A = 2; N_B = 0) \tag{3.27}$$

Now the Hamiltonian fragments are:

$$\widehat{H}_A(N_A = 2) = -\frac{1}{2}\Delta_{\vec{r}'} - \frac{1}{2}\Delta_{\vec{r}''} - \frac{1}{\left|\vec{r}' - \vec{d}_A\right|} - \frac{1}{\left|\vec{r}'' - \vec{d}_A\right|} + \frac{1}{|\vec{r}' - \vec{r}''|} \tag{3.28}$$

for the left-hand atom and

$$\widehat{H}_B(N_B = 0) = 0 \tag{3.29}$$

for the right-hand atom, while the Coulombic interaction changes to

$$\widehat{V}_{AB}(N_A = 2; N_B = 0) = -\frac{1}{\left|\vec{r}' - \vec{d}_B\right|} - \frac{1}{\left|\vec{r}'' - \vec{d}_B\right|} + \frac{1}{\left|\vec{d}_A - \vec{d}_B\right|} \tag{3.30}$$

Note that, in $\widehat{V}_{AB}$, both electrons are interacting with the nucleus of the right-hand atom, while in $\widehat{H}_A$, the electrons are interacting with each other.

The same argument applies to the case where both electrons are associated with the Hamiltonian fragment for the right-hand atom. Only the indices are changed. The form of (3.30) still consists of a two-electron fragment H, a zero-electron fragment H, and the electrons on the two-electron fragment interacting with the nucleus of the zero-electron fragment.

Once again the same picture holds for an arbitrary number of atoms interacting strongly with each other. Equation (3.15) must be modified so that the number of electrons on the different Hamiltonian fragments can vary. This is the key step that does not appear in the atoms-in-molecules and diatomics-in-molecules approaches. Thus, in the general case, $\widehat{H}^{(f)}$ must be rendered as

$$\widehat{H}^{(f)} \equiv \sum\nolimits_A \left(\widehat{H}_A(N_A) + \frac{1}{2}\sum\nolimits_{B \neq A} \widehat{V}_{AB}(N_A, N_B) \right) \tag{3.31}$$

However, this expression for does not tell how to deal with the variation in the numbers of electrons on each fragment, only that it is undetermined. This variation is dealt with next.

3.4 Selected Many-Electron, Valence Basis States

Electrons may or may not fluctuate among atomic sites in a molecule or solid, depending on the environment provided by other atoms or by external fields. Both types of events need to be included, in order that a given atom (or fragment) be allowed to adapt to that environment. This requirement places a minimal set of demands on the wave functions that are employed in the development of the FH approach.

There are two major categories of minimal, many-electron wave functions. (Note that we very much mean many-electron wave functions, rather than one-electron orbitals.) One category is the very simplest kind of valence bond wave function. By this we mean the analogs of the Heitler-London wave function for an H_2 molecule. The key feature of this kind of wave function is that the atoms are individually neutral or nearly so. The other category is the single-determinant/LCAO[1]-based state, most commonly associated with molecular orbitals. Neither of these minimal categories is adequate for the FH approach. The valence state category is most successful in describing systems where the electrons are highly localized on atomic centers, but fails for systems that require delocalized states. The LCAO-based states can describe delocalized states well, but fails when atoms dissociate from a molecule or solid, or when electrons are strongly correlated.

For these reasons, a broader set of valence states is sought that forms a hybrid between these two minimal sets. The basic motivation for our choice is apparent again from the neutral H_2 molecule. First, assume that a definition of charge on each atom has been chosen. One distribution of the electrons conforms to the Heitler-London or covalent state, which places one electron on each atom. For example,

$$|00\rangle = (\varphi_A(\mathbf{1})\varphi_B(\mathbf{2}) + \varphi_A(\mathbf{2})\varphi_B(\mathbf{1}))/\sqrt{2} \tag{3.32}$$

where φ is an arbitrarily chosen orbital and A and B label the hydrogen nuclei of the molecule, respectively. The labels "**1**" and "**2**" signify the spatial coordinates of electrons 1 and 2, respectively. Spin dependencies are neglected for the time being. Alternatively, either of two ionic states places both electrons on one atom and none on the other. For example, when A is anionic and B is cationic,

$$|-+\rangle = \varphi_A(\mathbf{1})\varphi_A(\mathbf{2}) \tag{3.33}$$

There is another ionic state, $|+-\rangle = \varphi_B(\mathbf{1})\,\varphi_B(\mathbf{2})$, where both electrons are on the opposite nucleus.

[1]LCAO stand for 'linear combination of atomic orbitals' which is commonly used in wave function calculations.

An arbitrary combination of these three states comprises a trial state, $|\psi\rangle$, for a diatomic molecule or, more generally, a two-fragment system. Using coefficients with indices that mimic the basis states,

$$|\psi\rangle = C_{00}\,|00\rangle + C_{+-}|+-\rangle + C_{-+}|-+\rangle \tag{3.34}$$

Normalization, $C_{00}^2 + C_{+-}^2 + C_{-+}^2 = 1$, is assumed. For conciseness, we will call this form of wave function a "Mulliken state". This state forms the basis for Mulliken electronegativity [20].

Applying this example to the identification of the fragment states, one makes the identifications $|A^0\rangle = |B^0\rangle = |00\rangle$ for the neutral state, $|A^+\rangle = |B^-\rangle = |+-\rangle$, and $|A^-\rangle = |B^+\rangle = |-+\rangle$ for the cation-anion and anion-cation combinations of the two atomic centers.

From this state, four extensions are possible. First, note that the state $|00\rangle$ corresponds to the covalent state and has been used, since the inception of valence bond theory, to define of the notion of covalency. That is, if $|C_{00}| = 1$, the bond is completely covalent. Second, note that we have not restricted the orbitals in (3.33) and (3.34) to be 1s orbitals. They could be any combination of, say for H atoms, 1s and 2p orbitals. For early transition metal centers, 3p and 3d orbitals might play the dominant role.

Third, any of these states may be augmented with factors that include explicit electron correlation without changing the basics of the assignment to fragment wave functions. Two prominent candidates for explicit correlation functions would be Jastrow factors [21] or Kutzelnigg forms [22]. The latter kind of wave function played a central role in the development of the Lee-Yang-Parr correlation functional in DFT [23–25]. As one explicit example, Rychlewski and Parr [19] apply very sophisticated James-Coolidge wave functions to the neutral H_2 molecule in their calculation of their definition of fragment energies. This remark pertains only to their (3.22) and (3.23).

Fourth and finally, Mulliken electronegativity can be defined from a combination of the isolated-atom Hamiltonian in (3.15) and the Mulliken state. When AB is dissociated, each of the three states contributing to the Mulliken state factor into a product of wave functions for each atom. In the dissociation limit, only the wave function of one atom needs to be applied to the corresponding Hamiltonian for that atom. Thus, one can define the atom energies

$$E_A(\zeta) = \left\langle A^{\zeta}(N_A)\middle|\widehat{H}_A(N_A = Z_A - \zeta)\middle|A^{\zeta}(N_A)\right\rangle \tag{3.35}$$

where ζ is the charge state of A and N_A is the number of electrons on A.

From these atom energies, the ionization energy for the neutral A atom may be defined as $\mathcal{I}_A = E_A(\zeta = +1) - E_A(\zeta = 0)$, while the electron affinity for the neutral A atom may be defined as $\mathcal{E}_A = E_A(\zeta = +1) - E_A(\zeta = 0)$. Now the Mulliken electronegativity is defined as the average of these two energies [20],

$$\chi_A = (I_A + \mathcal{E}_A)/2 \tag{3.36}$$

Similarly, the Parr-Pearson hardness [26] is defined as the difference

$$\eta_A = (I_A - \mathcal{E}_A)/2 \tag{3.37}$$

These two energies scales have played a central role in any number of interatomic potentials.

From these arguments, one can distill the essential requirements for the allowed valence states in the FH model. Besides being suitable for bound states of systems of fermions [27], the three immediate requirements for wave functions within the FH approach are that (1) integer charges for each fragment can be defined from each many-electron wave function, (2) the integer charges allowed in the set of many-electron wave functions must span the range of charge states considered important for a particular problem of interest, and (3) variables appropriate to the charge states of each fragment can be defined. The charge-state variables pertain to the *states* of the fragments, rather than the states of the electrons, and create the connection back to the original electronic Hamiltonian problem in (3.4), from which the FH approach originates.

First, we consider those essential characteristics of many-electron wave functions that allow us to compose the basic features of the FH approach. Begin with an arbitrary, trial wave function $|\psi\rangle$, with N electrons in the system. It is expanded in a basis set of many-electron valence states $|i\rangle$, with the properties that they be normalized, antisymmetric, and assign an integer number of electrons to each and every fragment in the system of interest. Each of these valence states involves all N electrons in the system. These are our only requirements at this point. The trial wave function is then expressed as

$$|\psi\rangle = \sum\nolimits_i c_i |i\rangle \tag{3.38}$$

where the c_i are expansion coefficients. The individual states are assumed to be orthonormal, $\langle i \mid j\rangle = \delta_{ij}$. This assumption is adopted for convenience, not out of necessity. Requiring normalization of $|\psi\rangle$ yields

$$\langle \psi \mid \psi\rangle = \sum\nolimits_i c_i^2 = 1 \tag{3.39}$$

In traditional quantum mechanics, the coefficients in (3.38) would be determined by solving the linear eigenvalue problem. We are going to take a different path though. Instead choose some fragment A, and define a set of states that describes its essential properties. The states for A will be derived from the chosen basis states $|i\rangle$. The essential property to be applied here to select which $|i\rangle$ are to be associated with a given state of A will be based on the *allowed integer charge states* of A. Specifying charge states for each fragment is sufficient for the present purposes.

In what follows, we will further assume that these states become factorable into products of wave functions for the isolated fragments. This stipulation ensures that exact energies are obtained at dissociation. In essence, the basis states change with bond lengths and/or volume. For this reason, they are neither adiabatic or diabatic states, but are referred to simply as valence bond states. If refinement of a model beyond the charge states and dissociation behavior is required, then other properties of the valence basis states will need to be specified.

With regard to charges on atoms or fragments, there are numerous definitions that might be adopted. The FH approach does not say anything about which definition should be adopted. That choice poses a separate issue. For modeling the FH, it is only necessary to choose one definition and to apply it consistently in constructing the valence states. It is important that the integer charges be defined for the chosen definition, since these charges provide input for the model.

As for which charge states are to be included, that is a problem- or system-dependent question. The majority of this chapter only requires that two or three charge states to be allowed for each fragment. If three states are allowed, most often they will be the cation (+), neutral (0), and anion (–) states, or in other words, a Mulliken state *for each fragment*. However, there is no theoretical restriction on the allowed number of charge states for a given fragment coming from the FH itself. Strictly speaking, one could include all cation states up to the nuclear charge of any fragment, plus the bound anion states. One example where this level of inclusion was applied appears in Gyftopoulos and Hatsopoulos [28].

Using the index ζ to indicate a chosen charge state, a trial state of A, $|A\rangle$, is expanded as

$$|A\rangle = \sum\nolimits_{\zeta} C_{A^\zeta} \left|A^\zeta\right\rangle \tag{3.40}$$

where the sum runs over the allowed charge states of A and the C_{A^ζ} are a new set of coefficients to be defined shortly. By $\left|A^\zeta\right\rangle$ we mean the wave function describing the state of A when it has integer charge ζ. The coefficients C_{A^ζ} contain the information about how much the ζ charge state of A contributes to its wave function.

Because we have stipulated that each $|i\rangle$ assigns a specific number of electrons on each fragment, $\left|A^\zeta\right\rangle$ may be represented by the projection

$$\left|A^\zeta\right\rangle = \sum\nolimits_{i} \delta_{\zeta\zeta_i^A} c_i \, |i\rangle / \sqrt{C_{A^\zeta}^* C_{A^\zeta}} \tag{3.41}$$

The index ζ_i^A denotes the charge state of A in state $|i\rangle$. The Kronecker delta functions $\delta_{\zeta\zeta_i^A}$ select the valence electronic states contributing to the ζ charge state of A. For any set of fragments comprising the total system, one has available the identities $|\psi\rangle = |A\rangle$, for any fragment A.

The new expansion coefficients C_{A^ζ} are defined in relation to the expansion coefficients of the original trial wave function by

$$C^*_{A^\zeta} C_{A^\zeta} = \sum\nolimits_i \delta_{\zeta\zeta_i^A} c_i^* c_i \tag{3.42}$$

It is essential to understand that the products $C^*_{A^\zeta} C_{A^\zeta}$ or the moduli $|C_{A^\zeta}|$ can be defined, but not the individual expansion coefficients C_{A^ζ} and $C^*_{A^\zeta}$. Because all of the original coefficients c_i appear in one charge state of A or another, we have the normalization condition,

$$\sum\nolimits_\zeta C^*_{A^\zeta} C_{A^\zeta} = 1 \tag{3.43}$$

for any fragment A. This is equivalent to saying that A must to been in one of its allowed charge states. $C^*_{A^\zeta} C_{A^\zeta}$ must then be the probability for A to occupy charge state ζ.

Equation (3.40) for $|A\rangle$ will be referred to as a projection because the multiplicity of valence states $|i\rangle$ have been transformed into a few states characterizing A, rather than the individual electrons. The individual charge states of A retain all of the residual detail in the original valence states. One may or may not want to account for these additional details, depending on the materials problem to be solved. Ignoring the detail results in the simplest models of the FH.

What are most important to the atomistic models of the FH are the products $C^*_{A^\zeta} C_{A^\zeta}$. These products could be identified as occupations numbers for the different charge states of A by the relationship

$$n_{A^\zeta} = C^*_{A^\zeta} C_{A^\zeta} \tag{3.44}$$

The occupation numbers n_{A^ζ} quantify the probability for finding A in the ζ-charge state, that is, the probability for *occupying* the state $\left|A^\zeta\right\rangle$. As an outcome of these definitions, the fractional number of electrons assigned to A, N_A, can be written as

$$N_A = \sum\nolimits_\zeta N_{A^\zeta} n_{A^\zeta} \tag{3.45}$$

where N_{A^ζ} is the integer number of electrons for the ζ charge state. This definition is essential for later development of models.

For convenience, density matrices for the fragment states Γ_A composed of the matrix elements $C^*_{A^\zeta} C_{A^\zeta}$ are introduced. In vector notation, let the set of coefficients for the states of A be the column vector $\boldsymbol{C}_A$. Then Γ_A is defined as the vector (outer) product

$$\Gamma_A = \boldsymbol{C}_A \boldsymbol{C}_A^\dagger \tag{3.46}$$

Notice that the column vector is first in the product while the row vector (transpose) is second. This order is the opposite of what applies to a dot (inner) product.

Again focus on one fragment A with the usual three charge states. Suppose that we have constructed the principal charge states for A, $|A^0\rangle$, $|A^+\rangle$, and $|A^-\rangle$, producing a trial state

$$|A\rangle = c_{A^0}|A^0\rangle + c_{A^+}|A^+\rangle + c_{A^-}|A^-\rangle \tag{3.47}$$

This expression will be referred to as the Mulliken state for fragment A. The occupation numbers are obtained from the fragment expansion coefficients as defined in (3.44). From these occupation numbers, the corresponding three-state version of Γ_A is

$$\Gamma_A(n_{A^+}, n_{A^-}) = \begin{pmatrix} n_{A^+} & \sqrt{n_{A^+} n_{A^0}} & \sqrt{n_{A^+} n_{A^-}} \\ \sqrt{n_{A^+} n_{A^0}} & n_{A^0} & \sqrt{n_{A^-} n_{A^0}} \\ \sqrt{n_{A^+} n_{A^-}} & \sqrt{n_{A^-} n_{A^0}} & n_{A^-} \end{pmatrix}_A \tag{3.48}$$

By normalization $n_{A^0} = 1 - n_{A^+} - n_{A^-}$. For notational convenience, the substitution is not made in the previous equation. Strictly speaking, the square roots should have ± signs in front of them, or more generally, a phase factor. This stage of development has not yet been pursued.

3.4.1 The Fragment Hamiltonian Approach

The crucial and unique notion of the present approach is that each state $|i\rangle$ *defines an electron distribution* on $\widehat{H}^{(f)}$. Symbolically,

$$\widehat{H}^{(f)}|i\rangle = \widehat{H}_i^{(f)}|i\rangle \tag{3.49}$$

with the meaning that $\widehat{H}_i^{(f)}$ (more explicitly each $\widehat{H}_{A^{\zeta_i}}$) assigns $N_{A^{\zeta_i}}$ electrons, the number of electrons appropriate to A, for state i. The contracted representation of the wave function for any fragment A leads to its energy $\bar{E}_A$ being expressed as

$$\bar{E}_A \equiv \sum\nolimits_{\zeta'\zeta} C^*_{A^{\zeta'}} C_{A^\zeta} H_A^{\zeta'\zeta} \tag{3.50}$$

where the energy matrix element is

$$H_A^{\zeta'\zeta} = \left\langle A^{\zeta'}(N) \middle| \widehat{H}_A^{\zeta'\zeta}\left(N_{A^{\zeta'}}, N_{A^\zeta}\right) \middle| A^\zeta(N) \right\rangle \tag{3.51}$$

with

$$\widehat{H}_A^{\zeta'\zeta}\left(N_{A^{\zeta'}}, N_{A^{\zeta}}\right) = \left(\widehat{H}_{A^{\zeta'}}\left(N_{A^{\zeta'}}\right) + \widehat{H}_{A^{\zeta}}(N_{A^{\zeta}})\right)/2 \tag{3.52}$$

It is essential to recognize that each fragment wave function is still a function of all N electrons in the system. The fragment operators are functions of the number of electrons on the fragment being operated on. To maintain hermiticity in the matrix elements, we symmetrize the definition of the fragment operator for two different charge-state wave functions, $\left|A^{\zeta}(N)\right\rangle$ and $\left|A^{\zeta'}(N)\right\rangle$, where ζ and ζ' are not necessarily equal. Such a situation appears in the off-diagonal matrix elements. These matrix elements play an essential role in the behavior of the FH approach.

It is convenient to convert this expression for $\bar{E}_A$ into density matrix form. That is,

$$\bar{E}_A = \mathrm{tr}(\mathbf{H}_A \Gamma_A) \tag{3.53}$$

where tr is the matrix trace operator, the density matrix Γ_A has elements $C^*_{A^{\zeta'}} C_{A^{\zeta}}$ [29], and $\mathrm{tr}\,\Gamma_A = 1$ is assumed. Averaging the fragment Hamiltonians in the energy matrix elements ensures that they are hermitian. The fragment-fragment interactions can be brought to an analogous form, namely

$$V_{AB} = \mathrm{tr}(\mathbf{V}_{AB} \Gamma_{AB}) \tag{3.54}$$

Summing these contributions leads to the general form of the FH total energy,

$$\bar{E} = \sum_A \bar{E}_A + \frac{1}{2}\sum_A \sum_{B \neq A} \bar{V}_{AB} \tag{3.55}$$

Let us return once more to the diatomic example. In the definition of $\widehat{H}_A^{\zeta'\zeta}$, the kinetic energy for the fragment Hamiltonian of A becomes

$$\widehat{T}_A^{\zeta'\zeta} = \frac{1}{2}\left(\sum_{j=1}^{Z_A - \zeta_A} \widehat{T}_j + \sum_{j=1}^{Z_A - \zeta'_A} \widehat{T}_j\right) \tag{3.56}$$

It depends explicitly on the values of ζ_A and ζ'_A. The kinetic energy operators for B are balanced with those of A so that charge neutrality is maintained at all times. As before, analogous expressions may be constructed for the other operators contributing to $\widehat{H}_A^{\zeta'\zeta}$. The Coulombic operators for the A-B atom interaction may be expressed as

$$\begin{aligned}\widehat{V}_{AB}^{\zeta'\zeta} = \frac{1}{2}\Big[&\left(\sum_{j=1}^{Z_A-\zeta} \widehat{V}_{jB} + \sum_{j=1}^{Z_A-\zeta'} \widehat{V}_{jB}\right) + \left(\sum_{j=Z_A-\zeta+1}^{N} \widehat{V}_{jA} + \sum_{j=Z_A-\zeta'+1}^{N} \widehat{V}_{jA}\right) \\ &+ \left(\sum_{j=1}^{Z_A-\zeta} \sum_{k=Z_A-\zeta+1}^{N} \widehat{V}_{jk}^{ee} + \sum_{j=1}^{Z_A-\zeta'} \sum_{k=Z_A-\zeta'+1}^{N} \widehat{V}_{jk}^{ee}\right)\Big]\end{aligned} \tag{3.57}$$

Thus, V_{AB} itself (3.54) is determined by the matrix of distributions of electrons between A and B, reflected in the $\widehat{V}_{AB}^{\zeta'\zeta}$ operators, even though AB is neutral by assumption. Only the charge state of A is needed, because of the neutrality assumption. In the more general, multi-fragment case, more complex notation is required. Summing the fragment contributions for a given distribution of electrons recovers the total system Hamiltonian.

More elaborate examples of the valence states may be found elsewhere [30–36]. Notice that when there are more than two fragments, then the identifications of the contracted states is not so trivial. In addition, these examples from the literature show extensions of the ideas to arbitrary systems and numbers of valence states.

For our canonical, three-charge state example, the Hamiltonian energy matrix for fragment A can be written as

$$\mathbf{H}_A = \begin{pmatrix} H^{++} & H^{+0} & H^{+-} \\ H^{0+} & H^{00} & H^{0-} \\ H^{-+} & H^{-0} & H^{--} \end{pmatrix}_{\boldsymbol{A}} \tag{3.58}$$

The matrix elements have been organized so that the number of electrons on atom A is increasing along the diagonal. For the sake of familiarity, we express the diagonal matrix elements by the notation

$$E_A(\zeta_A) = H_A^{\zeta\zeta} \tag{3.59}$$

where ζ_A is any one of the integer charge states that fragment A can occupy. These last equations will be central to the discussion in the rest of the chapter.

3.4.2 *New Variables for the Charge States of Each Fragment*

The preferred variables used to describe the states of atoms in a material are net charge and background density. The net charge q appears in many models of ionic materials, and understanding how this is to be done on a physically sound basis has been the subject of many theoretical studies [26, 37–45]. These models begin with simple pair potentials such as simple electrostatic potentials, the Buckingham potential, and the like. More sophisticated models, such as EEM [46], QEq [47], ReaxFF [48], COMB [49], split charge potentials [50–54], and others, introduce many-body effects that account for the changes in the atomic site (*a.k.a.* fragment) energies. The basic notion of how q appears in the fragment energy was first advanced formally by Iczkowski and Margrave [55]. Values of q from an electronic structure calculation or X-Ray measurement of electron density are equally applicable, as long as one is consistent—choose a definition and apply it throughout the development of an FH model.

The second variable, background density $\bar{\rho}$ appears in models of metal and alloy materials. The fragment energies of the atoms in these materials, called the

embedding energy in these models, are regulated by changes in the local background density defined for each atom for any configuration of its neighbors. The actual identity of the background density in terms of electronic structure information is not known.

This situation with the choice of variables presents a problem for modeling the FH. At this point in the development of the model, the principle variables are the occupancies n_{A^ζ} for various integer charges ζ of the allowed charge states of the different fragments. In the case where cation, neutral, and anion states are allowed on fragment A, the occupation numbers are n_{A^+}, n_{A^0}, and n_{A^-}. To proceed, these variables must be related to the preferred variables q and $\bar{\rho}$.

First, the n_{A^ζ} are not independent because of the normalization condition (3.43). Second the charge q_A on A may be defined as $q_A = Z_A - N_A$ [18, 39, 40, 44, 45]. Applying (3.43)—normalization—and (3.45)—the definition of N_A—to the definition of q_A, leads to the general relationship in any model of the FH,

$$q_A = Z_A \sum\nolimits_\zeta n_{A^\zeta} - \sum\nolimits_\zeta N_{A^\zeta} n_{A^\zeta} = \sum\nolimits_\zeta (Z_A - N_{A^\zeta}) n_{A^\zeta} \tag{3.60}$$

For the case of our usual three charge states being accessible to fragment A, $Z_A - N_{A^+} = +1$, $Z_A - N_{A^0} = 0$, and $Z_A - N_{A^-} = -1$. Thus q_A reduces to

$$q_A = n_{A^+} - n_{A^-} \tag{3.61}$$

That is, the net charge is the difference between cation and anion occupation numbers. If the charges pertain to a metal oxide, one might want to expand the charge range to say [−2, +2], or more generally to [−ζ, ζ]. One then simply scales (3.60) by ζ to obtain

$$q_A = \zeta (n_{A^{+\zeta}} - n_{A^{-\zeta}}) \tag{3.62}$$

Still more generally, but still with in a three-state model, suppose that atom can takes on charges states ζ_-, 0, and ζ_+. The net charge becomes

$$q_A = \zeta_+ n_{A^{\zeta_+}} + \zeta_- n_{A^{\zeta_-}} \tag{3.63}$$

Here the sign of ζ_- is assumed to be negative. Note that the occupation numbers pertain to the charges states ζ_+ and ζ_-. Any of these variations can be obtained by application of (3.60).

Third, we must define variable besides net charge that we called the ionicity τ_A [56]. The simplest and most obvious choice for ionicity is that it is orthogonal to the net charge q_A in the three-state model. The general definition of τ_A is taken as the sum of the occupation numbers,

$$\tau_A = \sum\nolimits_\zeta n_{A^\zeta} \tag{3.64}$$

Other definitions are certainly possible, but this definition is the simplest and suits our present needs. For ionicity in the three-state model, it reduces to

$$\tau_A = n_{A^+} + n_{A^-} \tag{3.65}$$

With the definition in (3.63), τ_A in (3.65) is indeed orthogonal to q_A in the "coordinates" of the occupation numbers.

Transforming the occupation numbers into charges and ionicities requires the relations

$$n_{A^\pm} = (\tau_A \pm q_A)/2 \tag{3.66}$$

Substituting these relations into (3.46) transforms the occupancy density matrix from $\Gamma_A(n_{A^+}, n_{A^-})$ into $\Gamma_A(q_A, \tau_A)$,

$$\Gamma_A(q_A,\tau_A) = \begin{pmatrix} (\tau+q)/2 & \sqrt{(1-\tau)(\tau+q)/2} & \sqrt{(\tau^2-q^2)/2} \\ \sqrt{(1-\tau)(\tau+q)/2} & 1-\tau & \sqrt{(1-\tau)(\tau-q)/2} \\ \sqrt{(\tau^2-q^2)/2} & \sqrt{(1-\tau)(\tau-q)/2} & (\tau-q)/2 \end{pmatrix}_A \tag{3.67}$$

In the present context, these two variables account for different properties of the fragment. Further details are available elsewhere [18]. Like (3.58), (3.67) will play a central role in the discussion for the remainder of the chapter.

To this point, four sets of variables have been defined, starting with the coefficients for the valence states in (3.38) and there have been three variable transformations. The first one contracts the coefficients of the valence states into coefficients for the states of the fragments C_{A^ζ}. The second converts the fragment state coefficients into occupation numbers, and the third transforms occupation numbers into net charge and ionicity. It is important to understand how the normalization constraint is transformed with these variable changes. Within the three-state model, the second and third transformations can be visualized. In Fig. 3.4 (left) the C_{A^ζ} appear for $\zeta = +1$, 0, or -1. Their rms value is unity so that all allowed values of these

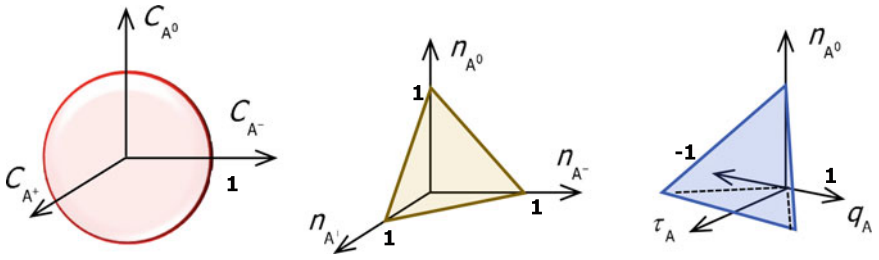

Fig. 3.4 Transformations among variable sets for charge states of fragments, (*left*) expansion coefficients, (*middle*) occupation numbers, and (*right*) charge and ionicity

coefficients must fall on a unit sphere. The occupation numbers transform the entire unit sphere into a triangle (Fig. 3.4 (middle)), defined by the points for unit occupancy for any one charge-state. Finally, Fig. 3.4 (right) shows the same triangle, with the cation and anion occupancies, n_{A^+} and n_{A^-}, rotated by 45° around the n_{A^0} axis. This outcome makes sense, since (3.61) and (3.65) constitute a unitary transformation. The net charge must remain within the range $-\tau_A \leq q_A \leq \tau_A$, τ_A itself must always equal $1 - n_{A^0}$ in (3.65), regardless of the value of q_A.

Ionicity is much less well understood than net charge. However, it is evident that another variable is required to maintain correspondence with the underlying occupation numbers. Indeed, one can observe this need all the way back in the original paper where Mulliken proposes his definition of electronegativity [20]. That definition is based exactly on the same three states that motivated the FH model, $|+-\rangle$, $|00\rangle$, and $|-+\rangle$. The relationship of the FH model to the Mulliken charge will be discussed in Application II below.

Moreover, while the definition of ionicity is no more unique than the definition of net charge, it is at least consistent with the qualitative notion of ionicity given by others [57], especially Phillips. That is, ionicity is some fractional character of the bonding environment, associated with the *total* charge fluctuations. The correspondence between our definition and Phillips' use of the term is not perfect. He develops the notion based on a two-state model, whereas we are developing the notion from a three-state model. For better or for worse, the distinction matters.

In the present development of our three-state FH model, the net charge appears very naturally. It will be seen to play the same role that it does in other models. The role of ionicity is yet to be defined and has not previously appeared in atomistic models. It will be the variable associated with the EAM background density $\bar{\rho}$. Taking this step has important implications for how time-honored concepts of electronegativity and hardness appear and operate in chemically-attuned atomistic models.

3.4.3 Three-State Fragment Energies

Before proceeding with some applications, we need to combine the results from the previous two subsections, and derive other essential quantities that will be used for the remainder of the chapter. The Mulliken state is applied exclusively here. Starting then with $\mathbf{H}_A$ from (3.58), the energy of the neutral state of A is separated from the other energies to yield

$$\mathbf{H}_A = E_{A^0} + \begin{pmatrix} \mathcal{I} & H^{+0} & H^{+-} \\ H^{0+} & 0 & H^{0-} \\ H^{-+} & H^{-0} & -\mathcal{E} \end{pmatrix}_{\boldsymbol{A}} \tag{3.68}$$

where $E_{A^0} = E_A(\zeta_A = 0)$. The negative of $\mathcal{E}_A$ appears in $\mathbf{H}_A$ because $E_A(\zeta_A = 0)$ is subtracted from the diagonal matrix elements, and the definition of $\mathcal{E}_A$ is designed to yield an electron affinity that is positive for isolated atoms.

Both of these energies, ionization and electron affinity, are generalizations of the conventional quantities found in textbooks and in the literature. See the definition of these matrix elements in (3.51), for $\zeta' = \zeta$. One can recover additional detail within the FH definitions of ionization potential and electron affinity by substituting the appropriate matrix elements into the definitions and expanding the fragment wave functions in the original valence basis states.

Moreover, the present definitions apply equally to situations where an atom is interacting with a whole environment of say water, dissolved salts, and proteins. Thus, the number of electrons entering into the valence states can be vast, while the number of electrons on a fragment operator will typically be very limited to a range of a few electrons on either side of the nuclear charge.

Next we transform the total energy for $\bar{E}_A$ from (3.50), using Γ_A from (3.67) and $\mathbf{H}_A$ from (3.68) into a form that looks more like a typical atomistic model. In a first step, recognize that the reference energy E_{A^0} multiplies an identity matrix $\mathbf{I}_A$ that has the same dimensions as the density matrix Γ_A In essence, $\mathbf{H}_A$ has been decomposed into

$$\mathbf{H}_A = E_{A^0}\mathbf{I}_A + \Delta\mathbf{H}_A \tag{3.69}$$

where $\Delta\mathbf{H}_A$ is the difference between $\mathbf{H}_A$ and $E_{A^0}\mathbf{I}_A$. The density matrix is normalized by construction so that $\mathrm{tr}\boldsymbol{\Gamma}_A = 1$ and $\mathbf{I}_A\boldsymbol{\Gamma}_A = \boldsymbol{\Gamma}_A$.

The second step is to calculate the trace of the matrix product $\Delta\mathbf{H}_A\boldsymbol{\Gamma}_A$. The fragment energy is rendered as

$$\begin{aligned}\bar{E}_A(q_A,\tau_A) \equiv E_{A^0} + \mathcal{I}_A\frac{(\tau_A+q_A)}{2} - \mathcal{E}_A\frac{(\tau_A-q_A)}{2} + H_A^{-+}\sqrt{\tau_A^2-q_A^2} \\ + 2\left[H_A^{0+}\sqrt{(1-\tau_A)(\tau_A+q_A)/2} + H_A^{-0}\sqrt{(1-\tau_A)(\tau_A-q_A)/2}\right]\end{aligned} \tag{3.70}$$

This expression is central to the remainder of the chapter.

A particular implementation of the three-state FH model pertains to any two-fragment system *AB*. The system can be any number of structures, such as two atoms, two molecules, two crystals, an impurity in contact with a reservoir, or an inclusion in a matrix. With just two fragments, the density matrix for *A* and the density matrix for *B* are completely correlated, because one always has $q_B = -q_A$ and $\tau_B = \tau_A$ when system-wide charge neutrality is enforced. For the same reason of system neutrality, only the neutral state of the *AB* is occupied in this example. Under this condition, only one occupation density matrix is required. The density matrix Γ_A can be applied since

$$\Gamma_A(q_A,\tau_A)=\Gamma_A(-q_A,\tau_A)=\Gamma_{(AB)^0}(q_A,\tau_A) \tag{3.71}$$

The last equality comes from the fact that the system is neutral.

From the correlation, the total energy for the two-fragment system $\bar{E}_{AB}$ can be expressed as the sum

$$\begin{aligned}\bar{E}_{AB}&=\bar{E}_A+\bar{E}_B+\bar{V}_{AB}\\&=tr\Big(\mathbf{H}_A\Gamma_A(q_A,\tau_A)+\mathbf{H}_B\Gamma_B(-q_A,\tau_A)+\mathbf{V}_{(AB)^0}\Gamma_{(AB)^0}(q_A,\,\tau_A)\Big)\end{aligned} \tag{3.72}$$

where $\bar{E}_A$ and $\bar{E}_B$ are the fragment energies from (3.70) and $\bar{V}_{AB}$ is the interaction energy between A and B.

The interaction matrix can be brought to the same form as $\bar{E}_A$. That is,

$$\mathbf{V}_{(AB)^0}=V_{A^0B^0}+\begin{pmatrix}V_{A^+B^-}-V_{A^0B^0} & V_{A_0^+B_0^-} & V_{A_-^+B_+^-}\\ V_{A_0^+B_0^-} & 0 & V_{A_0^+B_0^-}\\ V_{A_-^+B_+^-} & V_{A_0^-B_0^+} & V_{A^-B^+}-V_{A^0B^0}\end{pmatrix}_A \tag{3.73}$$

In this expression $V_{A^0B^0}$ is the pair interaction for the case where both A and B are individually neutral. The functions $V_{A^+B^-}$ and $V_{A^-B^+}$ are also diagonal contributions, whereas the functions $V_{A_0^+B_0^-}$, $V_{A_0^-B_0^+}$, and $V_{A_-^+B_+^-}$ are transitional interactions among various combinations of charge states.

Making the appropriate substitutions into (3.72), the total energy of AB may be expressed as the sum

$$\begin{aligned}\bar{E}_{AB}(q_A,\tau_A)&=E_{A^0}+E_{B^0}+V_{(AB)^0}\\&+(\mathcal{I}_A-\mathcal{E}_B+V_{A^+B^-}-V_{A^0B^0})\frac{(\tau_A+q_A)}{2}\\&+(\mathcal{I}_B-\mathcal{E}_A+V_{A^-B^+}-V_{A^0B^0})\frac{(\tau_A-q_A)}{2}\\&+\Big(H_A^{-+}+H_B^{+-}+V_{A_-^+B_+^-}\Big)\sqrt{\tau_A^2-q_A^2}\\&+2\Big(H_A^{0+}+H_B^{-0}-V_{A_0^+B_0^-}\Big)\sqrt{\frac{(1-\tau_A)(\tau_A+q_A)}{2}}\\&+2\Big(H_A^{-0}+H_B^{0+}-V_{A_0^-B_0^+}\Big)\sqrt{(1-\tau_A)(\tau_A-q_A)/2}\end{aligned} \tag{3.74}$$

Note that the pair interactions are modifying the values of the ionization and electron affinity energy scales, yielding effective values for these quantities. These effective energies are defined as

$$\mathcal{I}_A^*\equiv\mathcal{I}_A+\frac{V_{A^+B^-}-V_{A^0B^0}}{2} \tag{3.75}$$

and

$$\mathcal{E}_A^* \equiv \mathcal{E}_A - \frac{V_{A^-B^+} - V_{A^0B^0}}{2} \tag{3.76}$$

First note the change of sign in the definition of the effective electron affinity. Second note the factor of ½ in both definitions. This factor of ½ is the same as the one appearing in (3.31) and comes from dividing the pair contributions evenly between the two fragments. They will be utilized in the coming applications.

Consequently, the basic results of the embedding energy for the fragments also apply to the energy for pairs of fragments. For instance, as noted above, Γ_A can be employed for the composite pair interactions as well. Beyond that though, one can also say what the difference in the pair interactions is for different values of ionicity. In turn, this distinction advances a qualitative difference between metallic and covalent bonding (see Fig. 3.4). As τ_A gets larger, a greater fraction of the atom-ion type interactions (equal proportions of the resonances $|A^0B^-\rangle + |A^-B^0\rangle$ and $|A^0B^+\rangle + |A^+B^0\rangle$) enter the pair energy compared to the resonance $|A^0B^0\rangle$ fraction. This assertion applies to each pair contributing to $\bar{V}_A^*$. Importantly, A and B being neural on average does *not* mean that only the $|(AB)^0\rangle$ resonance contributes to the pair interaction.

The many-body effects captured in $\Gamma_A(0, \tau_A)$ assign an occupancy emanating from the other charge states of each AB pair, available from the neighboring atoms. Collecting terms yields

$$h_A^* \equiv W_A^{\text{le}} + W_\bullet^{\text{le}} - \left(V_{A_0^+} + V_{A_0^-}\right) \tag{3.77}$$

and

$$\bar{\eta}_A^* \equiv U_A^{eff} + U_\bullet^{eff} + \frac{1}{2}(V_{A^+} + V_{A^-}) - V_{A^0} + V_{A_+^-} \tag{3.78}$$

Thus, with interacting fragments, there are many variations on the basic definitions of electronegativity and charge-transfer gap.

3.4.4 Applications

In the following, five applications are explained to illustrate the different features of the three-state FH model. These applications cover the H_2 molecule, the generalizations of the Mulliken electronegativity and Parr-Pearson hardness, the concept of "metallic character" in this FH model, charge-flow regulation in a two-fragment, two-state model, and chemical potential equalization.

3.4.4.1 Application I: H_2 Molecule

In this application we want to bring together the three basic ideas introduced in the immediately previous sections, (1) Hamiltonian decomposition, (2) specific valence states, and (3) charge fluctuations on the Hamiltonian fragments. For this purpose, consider yet again the neutral H_2 molecule.

The specific behavior that we want to illustrate here is how the FH approach intertwines the different distributions of electrons on the Hamiltonian fragments with the valence states. The electron distributions appear in (3.25) and (3.27). The corresponding valence states appear in (3.32) and (3.33). When $\widehat{H}^{(f)}(H_2)$ operates on the covalent state $|00\rangle$, the Hamiltonian decomposition corresponds to that of (3.15), whereas when $\widehat{H}^{(f)}(H_2)$ operates on the anion-cation state $|-+\rangle$, the Hamiltonian decomposition corresponds to that of (3.25). Operation of $\widehat{H}^{(f)}(H_2)$ on the cation-anion state $|+-\rangle$ simply interchanges the values of N_A and N_B from 2 and 0, to 0 and 2.

In the present example, the matrix elements come from a two-electron wave function for the *molecule* operating a 0-, 1-, or 2-electron operator. These wave functions need not, and in general will not, be the eigenfunctions for the isolated atoms or ions. These generalizations will be expanded upon in the next Application.

The net charge q_A is also assumed to be zero, as it must for the ground state of the H_2 molecule (This distribution is not necessary for charge-polarized excited states. See for instance [58]). Applying $q_A = 0$ to (3.70), along with collecting terms, leads to the expression

$$\bar{E}_A(0,\tau_A) = E_A(0) + \frac{(\mathcal{I}_A - \mathcal{E}_A)\tau_A}{2} + 2\sqrt{(1-\tau_A)\tau_A/2}\left(H_A^{0+} + H_A^{0-}\right) \tag{3.79}$$

It is now apparent that the leading terms constitute a *generalized*, Fragment Hamiltonian definition of hardness that is analogous to Parr-Pearson chemical hardness [26]. Borrowing the standard notation for this quantity, even though there are differences in the detailed meanings, define the fragment hardness of A as

$$\eta_A = (\mathcal{I}_A - \mathcal{E}_A)/2 \tag{3.80}$$

Substituting gives the final expression for $\bar{E}_A$ as

$$\bar{E}_A = E_A(0) + \eta_A\tau_A + 2\sqrt{(1-\tau_A)\tau_A/2}\left(H_A^{0+} + H_A^{0-}\right) \tag{3.81}$$

Specific values for the neutral-atom energy, fragment hardness, and hopping energies, as functions of bond length, d, in a system of interest yields $\bar{E}_A$ as a function bond length, for selected values of ionicity τ_A. In terms of established literature, see the classic results of Coulson and Fischer where their value of mixing coefficient can be converted into an optimum value of τ_A using (3.65) with $\eta_{A^+} = \eta_{A^-}$.

There are several important features to notice in the last equation. One is that the variable multiplying η_A is the ionicity [56], not net charge [55]. The other is that there are new contributions that do not appear in any other atomistic model. These terms correspond to various kinds of one- and two-electron hopping processes. The hops are of a charge-transfer variety. They do not entirely comply with the hopping processes described in single-band tight-binding models. Finally, the two-electron hopping contribution is rarely discussed outside of superconductivity theory. Each of these features arises naturally from the FH approach and will be considered in more depth below.

3.4.4.2 Application II: Mulliken Electronegativity, Parr-Pearson Hardness, and Their FH Counterparts

The whole analysis in the previous example applies to any two-fragment (or two-component system). It is only when actual numerical values are needed that the true identity of the molecule is required. In this next application, we further generalize the analysis of (3.70) to a generic diatomic molecule *AB*. FH counterparts to Mulliken electronegativity and Parr-Pearson hardness come to the fore.

For this applications, recall that the Mulliken electronegativity [20] begins with a mixture of the states, $|+-\rangle$, $|00\rangle$, and $|-+\rangle$, for *AB* (3.47). With three states, there must be three energies and two independent coefficients. The third coefficient can be eliminated by application of the normalization condition on the total wave function. One is a zero of energy chosen for convenience. We have chosen $E_A(0)$ for that purpose and there is no separate variable associated with it.

A second energy level can be related closely to the Mulliken electronegativity, χ. It is half of the difference between energies of the cation and anion states of an atom,

$$\chi_A \equiv (E_A(\zeta = +1) - E_A(\zeta = -1))/2 \tag{3.82}$$

The contributing energy levels are expectation values of fragment Hamiltonian operators over wave functions for all electrons in the system. See (3.51). To be strictly identical to the Mulliken definition, the wave functions would have to be applied to isolated atoms, as we did in (3.34), and the Coulombic interaction operators between fragments would need to be negligible. In contrast, the FH definition retains the same fragment Hamiltonian operators, but averages over more general states of the interacting system. The factor of two comes about because electronegativity is also an average of the ionization potential and electron affinity,

$$\chi_A \equiv (\mathcal{I}_A + \mathcal{E}_A)/2 \tag{3.83}$$

Both definitions can be obtained from (3.70) by collecting the terms linear in q_A. In this context, recall that net charge is a proxy for particle number. Thus, this energy level is directly related to local chemical potential. This association is

explored more fully when the differential treatment of electronegativity is introduced shortly.

A third energy needs to be identified and one more variable accounted for. Collecting terms linear in τ_A, (3.70) we find an energy scale that is a generalization of the finite-difference version of the Parr-Pearson hardness η, or a charge-transfer gap in other disciplines, defined for A as

$$\eta_A \equiv E_A(\zeta = +1) + E_A(\zeta = -1) - 2E_A(\zeta = 0)/2 \tag{3.84}$$

Mulliken never discussed this energy scale. It was only years later that Pearson [59] and then Parr and Pearson [26] provided a physical basis and identification for the scale. In the FH approach, though, the variable associated with (or conjugate to) hardness is the ionicity τ_A, rather than the square of an electron number as in the I-M model.

The same generalization for hardness appears here as in the generalization of electronegativity above. The energy expectation values involved employ averages of states for the total system over fragment Hamiltonian operators, rather than states appropriate to the isolated fragments.

As has long been recognized, the Parr-Pearson hardness (finite-difference form) is closely related to the Hubbard U parameter ($U = 2\eta$). These ideas can be used interchangeably if one applies the concept of the Hubbard U in the sense of its qualitative, but intended, definition. The description of the Hubbard U as "on-site electron-electron repulsion" is an approximation based on Hartree-Fock theory that Hubbard found necessary to apply at the time of his ground-breaking work on the subject.

With these definitions it is natural to rearrange (3.70) to the form

$$\begin{aligned}\bar{E}_A(q_A, \tau_A) \equiv\ & E_A(0) + \chi_A q_A + \eta_A \tau_A + H_A^{-+}\sqrt{\tau_A^2 - q_A^2} \\ & + 2\left[H_A^{0+}\sqrt{(1-\tau_A)(\tau_A + q_A)/2} + H_A^{-0}\sqrt{(1-\tau_A)(\tau_A - q_A)/2}\right]\end{aligned} \tag{3.85}$$

Herein both the generalized electronegativity and hardness appear, multiplied by net charge and ionicity, respectively. One expects these quantities to persist in the next step of the analysis.

To determine or estimate ground-state energies, our original goal in the pursuit of interatomic potentials, it will be necessary to use (3.85) in an optimization procedure. Practically speaking, this means differentiating this equation with respect to both variables. Since both χ_A and η_A appear in (3.85), they will also appear in any differentiated expressions. For this reason, it will be natural to refer to the quantities in (3.83) and (3.84) as "static" electronegativity and hardness, and the differential analogs, to be defined presently, as "dynamic" electronegativity and hardness. This nomenclature is chosen in analogy to the terminology of "static" and "dynamic" Born charges [60].

The differential, or dynamic, form for fragment electronegativity χ_A^* brings in an influence from hopping processes. One can anticipate this influence within the structure of $\bar{E}_A$. Taking the derivative of $\bar{E}_A$ in (3.85) with respect to q_A at constant τ_A yields

$$\chi_A^* \equiv \left(\frac{\partial \bar{E}_A}{\partial q_A}\right)_{\tau_A} = \mathrm{tr}\left[\mathbf{H}_A\left(\frac{\partial \Gamma_A}{\partial q_A}\right)_{\tau_A}\right]$$
$$= \chi_A - H_A^{-+}\left(\frac{q_A}{\sqrt{\tau_A^2 - q_A^2}}\right) + \sqrt{\frac{1-\tau_A}{2}}\left[\frac{H_A^{0+}}{\sqrt{\tau_A + q_A}} - \frac{H_A^{-0}}{\sqrt{\tau_A - q_A}}\right] \tag{3.86}$$

Moreover the static electronegativity χ_A appears as an explicit contribution to χ_A^* as anticipated. One can describe the relationship between these two energy levels as a modulation of χ_A. Besides modulation by one-electron hopping processes, χ_A is also modulated by a contribution from two-electron hopping processes. Very importantly, dynamic electronegativity is *conjugate* to net charge. When fragment A is neutral, $q_A = 0$, the local Fermi level is not shifted from its static value by two-electron hopping processes. The shift of the Fermi level from one-electron hopping processes will generally be muted because the cation and anion hopping integrals appear with opposite signs. See Fig. 3.6 (left).

Of course, either the static or dynamic fragment electronegativity must be related to a local chemical potential or Fermi level. Specifically the relationship is

$$\mu_A^* = -\chi_A^* = \frac{\partial \bar{E}_A}{\partial N_A} = -\frac{\partial \bar{E}_A}{\partial q_A} \tag{3.87}$$

for the dynamic quantities. The relationships among the PPLB, I-M, and FH models for μ_A^* are illustrated in Fig. 3.5 in the dissociation limit.

By analogy with Parr-Pearson hardness, one can also define a differential or "dynamic" hardness. This time, taking the derivative of $\bar{E}_A$ with respect to τ_A at constant q_A yields the dynamic hardness η_A^*

$$\eta_A^* \equiv \left(\frac{\partial \bar{E}_A}{\partial \tau_A}\right)_{q_A} = \mathrm{tr}\mathbf{H}_A\left(\frac{\partial \Gamma_A}{\partial \tau_A}\right)_{q_A}$$
$$= \eta_A + H_A^{-+}\left(\frac{\tau_A}{\sqrt{\tau_A^2 - q_A^2}}\right) + \frac{1}{\sqrt{2(1-\tau_A)}}\left[H_A^{0+}\left(\frac{1-2\tau_A+q_A}{\sqrt{\tau_A+q_A}}\right) + H_A^{-0}\left(\frac{1-2\tau_A-q_A}{\sqrt{\tau_A-q_A}}\right)\right] \tag{3.88}$$

Our definition of τ_A requires that ionicity always be positive. Consequently two-electron hopping can either narrow ($H_A^{-+} < 0$) or widen ($H_A^{-+} > 0$) the gap embodied within the static hardness. The one-electron hopping contributions, H_A^{0+} and H_A^{-0}, generally decrease the gap. This topic of gap behavior will be explored more fully in Application III below.

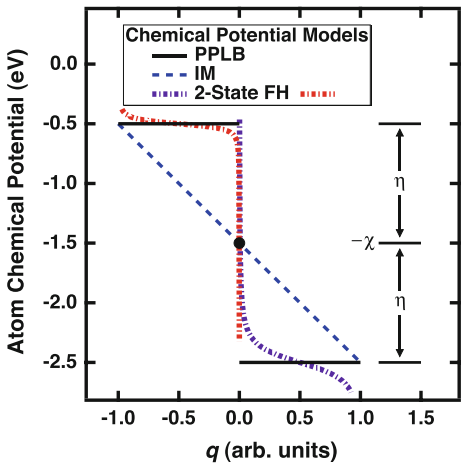

Fig. 3.5 Three-state models of the atom chemical potential (3.85) underlying I-M (*blue*, *dash*), PPLB (*black*, *solid*) models of chemical potential, and separate two-state FH models for the cation (*purple*, *dot-dash*) and anion (*red*, *dot-dash*) states [18]. One parameter for the two-state model is given an approximate value to aid in visualizing its differences with the PPLB model. Fragment hardness appears as components of μ_A in both the I-M and PPLB models

Parr and Pearson thought that they were making an approximation by using the static hardness as a proxy for the dynamic hardness [26]. From the FH perspective though, this is not the case. The static contributions for both the fragment electronegativity and hardness appear within the corresponding dynamic definitions.

Finally, note that there is no dependence of net charge and ionicity in the fragment energy matrix. This is because the discussion in this Application is predicated on a two-fragment system. In the more general application of an atomic model operating under nonequilibrium conditions, there are other contributions to the dynamical definitions of both fragment electronegativity and hardness. These are not discussed in this volume.

For a homo-nuclear diatomic molecule, where $q_A = 0$ for the ground state, both dynamic quantities simplify considerable and are therefore more amenable to analysis. The dynamic electronegativity reduces to

$$\chi_A^* = \chi_A + \left[H_A^{0+} - H_A^{0-}\right]\sqrt{\frac{1-\tau_A}{2\tau_A}} \tag{3.89}$$

or

$$\frac{\chi_A^*}{\chi_A} = 1 + \left[\frac{H_A^{0+} - H_A^{0-}}{\chi_A}\right]\sqrt{\frac{1-\tau_A}{2\tau_A}} \tag{3.90}$$

for a neutral fragment. Whenever the energies for cation and anion hopping are equal, then $\chi_A^* = \chi_A$.

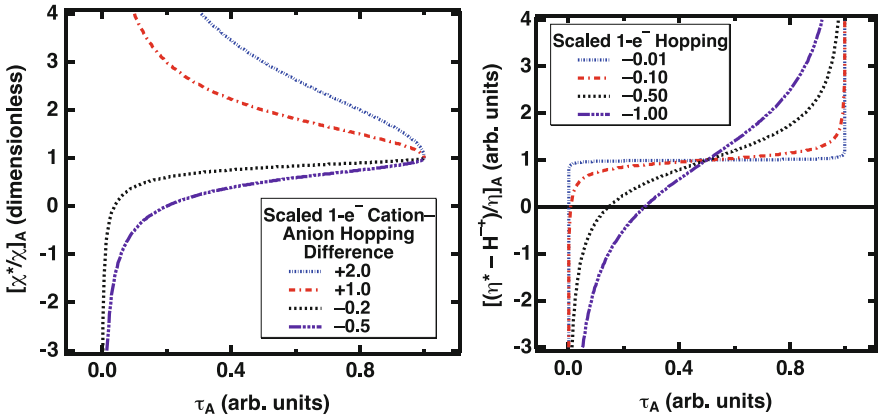

Fig. 3.6 Illustrations of (*scaled*) differential fragment electronegativity (*left*) and hardness (*right*) versus ionicity, for various values of scaled one-electron hopping. The scaling factor for fragment electronegativity is the static electronegativity and the one-electron hopping pertains to difference between cation and anion hopping. The scaling factor for the fragment hardness is the static hardness. One proposed definition of absolute hardness as a function of N is most closely approximated, in this illustration, by the curve labeled "−0.01"

One the other hand, the differential hardness η_A^* reduces to

$$\eta_A^* = \eta_A + H_A^{-+} + \left[\frac{H_A^{0+} - H_A^{0-}}{\sqrt{2}}\right] \frac{1 - 2\tau_A}{\sqrt{(1-\tau_A)\tau_A}} \tag{3.91}$$

or

$$\frac{\eta_A^* - H_A^{-+}}{\eta_A} = 1 + \left[\frac{H_A^{0+} - H_A^{0-}}{\sqrt{2}\eta_A}\right] \frac{1 - 2\tau_A}{\sqrt{(1-\tau_A)\tau_A}} \tag{3.92}$$

The two-electron hopping integral remains an influence. The scaled form is shown in the right-hand panel of Fig. 3.6.

3.4.4.3 Application III: Metallic Character in an Atomistic Model

Even though the FH approach is an atomistic one, a notion long associated with metallic character has been retained. That is, a differential charge-excitation gap for each fragment η_A^* is defined in (3.91) and evidently changes its values as the result of three distinguishable effects. For this reason, FH ionicity is put forth as a measure of *metallic character*. A first effect comes from η in (3.84) being defined by expectation values for wave functions of the *total*, interacting system, rather than for changes in the number of electrons of an isolated *fragment*. A second effect comes from the differential charge-excitation gap that brings in electron hopping

influences. The third and final effect is new compared to what has been presented to this point. That effect originates from the fragment-fragment interactions. It is emphasized here as a first example of metallic character in any interatomic potential.

By the term "metallic character" we mean a qualitative concept in analogy with the term "metallicity" defined in molecular orbital theory [61]. This term needs to be distinguished from the term "metallic behavior" in the sense of electron conduction. While measures of metallic character or metallicity are helpful, they are not rigorous proof of actual conductivity.

To illustrate the qualitative relationship between ionicity and metallic character, we solve a simple impurity model. The system is divided into two parts, a central fragment A or impurity of interest and 'a reservoir' denoted by a bullet •, and constituting the remainder of the system. Obviously the impurity-reservoir problem fits the paradigm of all the other two-fragment systems considered here already. We opt for the notational change to demarcate the difference between this example and the prior, molecular examples.

Charge neutrality of A is assumed and its ionicity τ_A is to be optimized. The solution for the optimum ionicity of A is

$$\tau_A = \frac{1}{2}\left[1 \pm \frac{\bar{\eta}_A^*/\eta_A}{\sqrt{\left(\eta_A^*/\eta_A\right)^2 + \left(h_A^*/\eta_A\right)^2}}\right] \tag{3.93}$$

where

$$h_A^* \equiv W_A^{ie} + W_\bullet^{ie} - \left(V_{A_0^+} + V_{A_0^-}\right) \tag{3.94}$$

and

$$\bar{\eta}_A^* \equiv U_A^{eff} + U_\bullet^{eff} + \frac{1}{2}\left(V_{A^+} + V_{A^-}\right) - V_{A^0} + V_{A_+^-} \tag{3.95}$$

The energies appearing in these definitions have been given above. The ground state ionicity corresponds to the negative root if $\bar{\eta}_A^*$ is positive. The total charge fluctuations are thus minimized, as one might expect. In contrast, the excited state maximizes the total charge fluctuations, driving A toward a plasma state. As the one-electron hopping energies become larger, both the ground- and excited-state ionicities converge to a value of 1/2, consistent with the qualitative notion of delocalization. See Fig. 3.7.

If, on the other hand, $\bar{\eta}_A^*$ becomes negative, the optimized value of τ_A can become greasster than 1/2. However, even if $\bar{\eta}_A^*$ is negative, as h_A^* increases, the optimized τ_A still tends to 1/2. Thus, the behavior of the optimized ionicities in this more general case follows the same pattern as in the specific case discussed in [56].

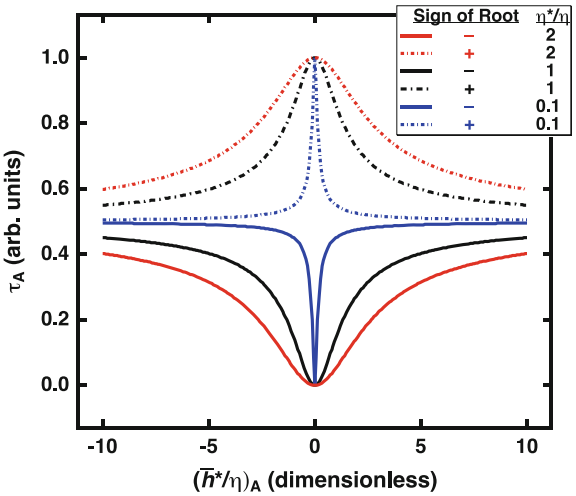

Fig. 3.7 Optimized ionicities from (3.93) versus scaled effective hopping. The hopping integral from (3.94) is scaled by the fragment hardness η_A, for different values of the hardness ratio $\bar{\eta}_A^*/\eta_A$, with $\bar{\eta}_A^*$ defined in (3.95). The negative roots approaches 1/2 from below, at very large, effective charge-transfer hopping, while the positive roots approach 1/2 from above

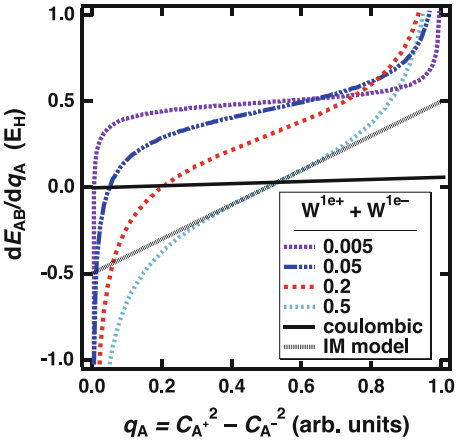

Fig. 3.8 Characteristic behavior of CT model with comparison to I-M model. *Slope* of *black line* corresponds to electrostatic interaction at 10 Å. The points where the variously *colored curves* cross the *black line* correspond to the optimized charges. Characteristic CT behavior of I-M model is the *dashed*, *gray line*

It must be made clear here that the present result, being more general, supports the contention that the FH, when applied to metals, can be interpreted as estimating the degree of metallic character (Fig. 3.8).

3.4.4.4 Application IV: Charge-Flow Regulation in a Two-State, Two-Fragment Model

For this model of the FH (3.70) for a diatomic molecule is again the simplest to consider. This example is intended to aid in understanding chemical potential equalization [18, 46, 62]. When A and B differ significantly from each other in electronegativity, the diatomic molecule AB may be described by just two states $|0\rangle = |A^0B^0\rangle$ and $|+\rangle = |A^+B^-\rangle$. Here we are assuming that A is going to play the role of the cation.

Choosing $C_{A^-} = 0$, one arrives at the relations $C_{A^+} = \sqrt{q}$ and $C_{A^0} = \sqrt{1 - C_{A^+}^2} = \sqrt{1-q}$. The iconicity τ equals q in this example. See (3.60) and (3.61). The coefficients for B have the same values as for A in this case, which is to say that the states of A and B completely correlated with each other. The subscript A is dropped from q since only one charge is needed to define the system. Thus its density matrix, as a function of q, reduces to

$$\Gamma_A = \begin{pmatrix} q & \sqrt{q(1-q)} \\ \sqrt{q(1-q)} & q \end{pmatrix}_A \tag{3.96}$$

Clearly, since there are only two states, the resulting model and the I-M model cannot be compatible. Neither electronegativity nor hardness as defined by the finite difference relations associated with (3.87) can play a role in this simplest application of the model's FH. On the other hand, new second- and higher-order contributions in the charge-dependence play a vital role in regulating charge flow.

The essential charge dependence of $E_A(q) = \mathrm{tr}(\mathbf{H}_A\Gamma)$ differs considerably from the I-M model. These differences can be brought to light by expanding the fragment energy $E_A(q)$,

$$E_A(q) = H_{00}^A + \left(H_{++}^A - H_{00}^A\right)q + 2H_{0+}^A\sqrt{q(1-q)} \tag{3.97}$$

where the energy matrix elements are defined in (3.58). The asterisk indicates that these energies come from the system states, rather than the properties of the isolated atoms. One term that is linear in q appears with an energy scale determined by a generalized ionization potential $\mathcal{I}_A^* \equiv H_{++}^A - H_{00}^A$. The associated chemical potential μ_A^* may be identified as

$$\mu_A^*(q) = \mathcal{I}_A^* + H_{0+}^A\frac{1-2q}{\sqrt{q(1-q)}} \tag{3.98}$$

It is evident from this expression that a Taylor series representation of $\mu_A^*(q)$ cannot be centered on either $q = 0$ or $q = 1$. Furthermore, neither the Mulliken electronegativity nor the chemical hardness appears in this model. Instead, first- and higher-order terms appear that have nothing to do with the finite-difference

definition of chemical hardness. Whether or not a constant term appears, depends on the expansion point used to determine this contribution. At $q = 1/2$, no constant is definable to combine with the ionization potential. At any other permissible expansion point, a nonzero constant does appear.

Thus, the square-root dependence from the transition matrix elements constitutes a unique facet of the model [39, 44, 45]. These dependencies correspond to rigorous density-functional [38] and perturbative [63] analyses of nonlinear behavior in the fragment chemical potentials. This model of transitional or hopping contributions to the fragment energies provides a physical basis for the nonlinearities in recently proposed models of chemical potential [49, 54].

3.4.4.5 Application V: Chemical Potential Equalization

Given the environmental dependencies of these fragment energy matrix elements, chemical potential equilibration takes on a new dimension. Equilibration is achieved by optimizing (3.55), subject to the constraint of charge neutrality, $0 = \sum_A q_A$. We use μ as the Lagrange multiplier associated with enforcing that neutrality and optimize with respect to each charge, so that $\mu = -\frac{\partial \bar{E}}{\partial q_K}$. Expanding $\frac{\partial \bar{E}}{\partial q_K}$, one obtains

$$\frac{\partial \bar{E}}{\partial q_K} = \mathrm{tr} \sum\nolimits_A \left(\mathrm{H}_A \frac{\partial \Gamma_A}{\partial q_K} + \frac{\partial \mathrm{H}_A}{\partial q_K} \Gamma_A + \frac{1}{2} \sum\nolimits_{B \neq A} \frac{\partial (\mathrm{V}_{AB} \Gamma_{AB})}{\partial q_K} \right) \tag{3.99}$$

In defining $\frac{\partial \bar{E}}{\partial q_K}$, one has two choices. In one choice, not only are the other net charges held constant, but also the total ionic character variables of each fragment are held constant. Then, a separate optimization process is required to determine those variables. In the other choice, one could assume that the τ are functions of their respective q. The first term is conventional as far as its physical content goes, being directly analogous to the I-M model. For either definition of $\frac{\partial \bar{E}}{\partial q_K}$, each fragment density matrix Γ_A will depend only on its q_A. Consequently, the K term is the sole survivor in the sum over the fragment energies.

The second term, $\sum_A \mathrm{tr} \frac{\partial \mathrm{H}_A}{\partial q_K} \Gamma_A$, is unprecedented. Ordinarily, it is simply set to zero, for lack of a more complete model. Physically $\frac{\partial \mathrm{H}_A}{\partial q_K}$ describes how excitations of A influence the chemical potential of fragment K, an effect originating from the many-body effects introduced here. These and other effects described in this section are unknown and unavailable in the I-M model. The present approach both describes a physical origin to new effects and allows for systematically incorporating them as needed for a particular application. One is then able to select the appropriate level of approximation for a particular problem. Extensions to more complex systems proceed along parallel lines of reasoning.

3.5 Informing the Embedded-Atom Method from FH: Two-Variable Embedding Energy

In this section and the next, the FH is cast into forms recognizable as atomistic models for metals and for ceramics as special cases. After these two cases are explored, a number of the existing atomistic models are derived starting from the FH point of view [64].

3.5.1 General Form of the Model Potential from FH

To recap, the general form of the fragment, or embedding, energy $\bar{E}_A$ in the three-state model (3.85) may be expressed as a function of q_A and τ_A through the density matrix $\Gamma_A(q_A, \tau_A)$ for each fragment and energies more familiar from EAM models. We have recognized that H_A^{+0} and H_A^{0-} represent one-electron, charge-excitation hopping integrals [65], associated with transitions between cation and neutral states of the embedded atom, respectively. These charge-excitation hops occur between anion and neutral states of a fragment. The two-electron hopping is a unique feature of the FH approach. The two-electron hopping energy might also be thought of as an electron-hole-pair hopping energy. In the FH view, electrons are changing which atoms they are associated with, leading to changes in the charge-states of the atoms. The situation here differs from that of the established tight-binding view, where electrons are changing the local atomic orbitals that they occupy. However, as the crystal dissociates, FH hopping energies all tend to zero, the energy assigned to the ground state of each atom.

Previously we defined ionization potential $\mathcal{I}_A$, electron affinity $\mathcal{E}_A$, electronegativity χ_A, and charge-transfer gap U_A for each fragment in a system. In addition, we adopt the notation $W^{1e+} = H^{+0}$, $W^{1e-} = H^{0-}$, and $W^{2e} = H^{+-}$ to make the physical significance of these quantities more explicit. Applying these definitions to (3.85), $\bar{E}_A(q_A, \tau_A)$ is rewritten as the sum of three embedding functions,

$$\bar{E}_A(q_A, \tau_A) = E_A^{gap}(q_A, \tau_A) - 2\left[W^{1e+} f^{1e+}(q_A, \tau_A) + W^{1e-} f^{1e-}(q_A, \tau_A)\right] \quad (3.100)$$

where the two one-electron-hopping embedding functions are

$$f^{1e\pm}(q, \tau) \equiv \sqrt{(1 - \tau)(\tau \pm q)/2} \quad (3.101)$$

and

$$E^{gap}(q, \tau) \equiv E_0^* + \chi^* q + \frac{U^* \tau}{2} - W^{2e} f^{2e}(q, \tau) \quad (3.102)$$

with the two-electron-hopping embedding function f^{2e} as

$$f^{2e}(q,\tau) \equiv \sqrt{\tau^2 - q^2}/2 \tag{3.103}$$

This is the fundamental form of the embedding energy that we will examine in the remainder of the chapter.

For a normal elemental, all of the net charges are zero for the ground state. As a result, the energies Δf_A^{gap} and χ_A^* are eliminated. If one wanted to determine the effective chemical potential, then these terms would have to be included in the optimization process. The optimization of the net charge must yield charges of zero. This fuller generality would also be required in more general materials modeling situations, but is not investigated further here.

3.5.2 Many-Body Interactions

The proposition in this model of the FH approach is that the state of each atom can be described sufficiently by three charge states. The process of contracting the many-electron, valence wave functions into fragment wave functions yields substructure in the fragment energy matrix elements. When the matrix elements in (3.51) are expanded, one finds

$$H_A^{\zeta\zeta'} = \frac{\sum_{ij} \delta_{\zeta\zeta_i^A} \delta_{\zeta'\zeta_j^A} c_i^* c_j \left\langle i \middle| \widehat{H}_A^{\zeta\zeta'} \middle| j \right\rangle}{\sqrt{\left(\sum_i \delta_{\zeta\zeta_i^A} c_i^* c_i\right)\left(\sum_j \delta_{\zeta'\zeta_j^A} c_j^* c_j\right)}} \tag{3.104}$$

with $\zeta = \pm 1$ or 0. Some substructure results from the number of states contributing to any single charge state of a fragment. Other substructure comes simply from decomposing the system into more than two fragments. In either case, substructure should be understood as originating from many-body effects. Note though that, even when the number of fragments exceeds two, the fragment energy has the same structure as (3.55). The energy matrix elements therein are functions of expansion coefficients. Specifically, how the effects of multiple neighbors can be cast as functions of charge and ionicity of those neighbors is the subject of this subsection.

One insightful prototype is a three-fragment, linear chain *B*–*A*–*B*′, where *B* and *B*′ need not be equivalent atoms. If all three states for three atoms are included, there are a total of seven states. In this prototype, charge exchange between the end atoms is neglected. The approximation is analogous to neglecting neighbor shells beyond the first in an ordered crystal. Under this approximation, the unnormalized, trial wave function of *BAB*′ involves the five valence states, $|000\rangle$, $|+-0\rangle$, $|-+0\rangle$, $|0+-\rangle$, and $|0-+\rangle$. The coefficients are labeled similarly.

The total Hamiltonian operator expressed in fragment form for BAB' is

$$\widehat{H}_{BAB'} \approx \widehat{H}_B + \widehat{H}_A + \widehat{H}_{B'} + \widehat{V}_{AB} + \widehat{V}_{AB'} \tag{3.105}$$

where $\widehat{H}_\alpha$ is the Hamiltonian fragment for each $\alpha \in \{B, A, B'\}$, and $\widehat{V}_{\alpha\beta}$ is a Coulombic interaction operator for the pair αβ. The B-B' interaction is neglected in keeping with our "nearest-neighbors only" approximation. For A in its many-body environment, it has two neighbors. The neutral state of A has only one component $\left|A^0\right\rangle = |000\rangle$. The cation and anion states each have two contributing states, namely $|A^+\rangle = \{|-+0\rangle, |0+-\rangle\}$ and $|A^-\rangle = \{|+-0\rangle, |0-+\rangle\}$. For these states, and taking H_A^{0-} as the primary example, (3.104) reduces to

$$H_A^{0-} = \frac{c^*_{+-0}c_{000}\left\langle +-0\middle|\widehat{H}_A^{0-}\middle|000\right\rangle + c^*_{0-+}c_{000}\left\langle 0-+\middle|\widehat{H}_A^{0-}\middle|000\right\rangle}{\sqrt{\left(c^*_{+-0}c_{+-0} + c^*_{0-+}c_{0-+}\right)c^*_{000}c_{000}}} \tag{3.106}$$

Through two factorizations, this expression reduces to

$$H_A^{0-} = \frac{c^*_{+-0}\left\langle +-0\middle|\widehat{H}_A^{0-}\middle|000\right\rangle + c^*_{0-+}\left\langle 0-+\middle|\widehat{H}_A^{0-}\middle|000\right\rangle}{\sqrt{c^*_{+-0}c_{+-0} + c^*_{0-+}c_{0-+}}} \tag{3.107}$$

Now we translate for the expansion coefficients into ionicity and charge. Because the two charge-fluctuation resonances between B and B' have been neglected, it is possible to find a mapping between the expansion coefficients and fragment occupation numbers. Each charge state of each end fragment has only one resonance, and so its occupation number maps to a single expansion coefficient. For this reason, it is simplest to use this mapping: $n_{+-0} = n_{B^+}$, $n_{-+0} = n_{B^-}$, $n_{0-+} = n_{B'^+}$, and $n_{0+-} = n_{B'^-}$.

From (3.66), one finds that

$$n_{B^-} = (\tau_B - q_B)/2 \tag{3.108}$$

and similarly for $n_{B'^+}$. Applied to (3.105), one finds

$$H_A^{0-} = \frac{\sqrt{(\tau_B - q_B)/2}\left\langle +-0\middle|\widehat{H}_A^{0-}\middle|000\right\rangle + \sqrt{(\tau_{B'} - q_{B'})/2}\left\langle 0-+\middle|\widehat{H}_A^{0-}\middle|000\right\rangle}{\sqrt{(\tau_B - q_B)/2 + (\tau_{B'} - q_{B'})/2}} \tag{3.109}$$

Similar relationships are found for the other energy matrix elements in the embedding energies, including the diagonal energies H_A^{++} and H_A^{--}. That is, both the states of A and transitions among them depend on the state of its environment, defined by the end fragments in this case. However, we were able to arrive at these

relationships because of the great simplicity of the three-fragment chain. In the general case needed for atomistic simulations, the relationships are much more complex and will need to be approximated in some way. This is because there are ordinarily far more expansion coefficients than fragments and charge-states of fragments.

Next consider the embedding energy of one of the end atoms, say B. The neutral state of B is composed of three contributions, $|B^0\rangle = \{|000\rangle, |0-+\rangle, |0+-\rangle\}$. Each of its ionic states involves only one resonance. Recall that, in contrast, the middle fragment A has only one resonance contributing to its neutral state, while the other four states contribute to its total ionic character. Thus, in terms of the number of contributions, the neutral state of B^0 is being excited by the other two resonances that come from polarization of the AB' pair. For this reason, we expect the FH in this model system to differentiate atoms on a surface from those in the bulk, as well as from differences in numbers of neighbors.

The pairs of fragments are just other kinds of fragments. However, only the Coulombic interactions between them are needed instead of the full diatomic Hamiltonian. For the AB pair interaction, again consider the three-fragment, linear chain B–A–B', with the same five states as adopted previously.

To understand the structure of the interaction, we only need to investigate $\widehat{V}_{AB}$ as a representative case. Because of the chosen prototype, the pair interaction may be expressed as

$$V_{BA} = \mathrm{tr}\mathbf{V}_{BA}\Gamma_{BA} \tag{3.110}$$

just as in (3.54), but for a three-fragment system. To understand the components of V_{BA}, recognize that the charge states of BA are the same as those of B'. Consequently, the matrix is $\mathbf{V}_{BA}$ is 3 × 3, the two ionic states only have a single contribution, but three states contribute to $(BA)^0$.

Using the $(BA)^0$ matrix element as the main example, it can itself be represented as the trace over a product

$$V_{(BA)^0} = \mathrm{tr}\mathbf{V}_{(BA)^0}\Gamma_{(BA)^0} \tag{3.111}$$

For the AB neutral pair, its density matrix has the environmental dependencies

$$\Gamma_{(AB)^0} = \begin{pmatrix} n_{B^+} & \sqrt{n_{A^0}n_{B^+}} & \sqrt{n_{B^+}n_{B^-}} \\ \sqrt{n_{A^0}n_{B^+}} & n_{A^0} & \sqrt{n_{A^0}n_{B^-}} \\ \sqrt{n_{B^+}n_{B^-}} & \sqrt{n_{A^0}n_{B^-}} & n_{B^-} \end{pmatrix} \tag{3.112}$$

However, n_{A^0} is a function of all of the ionic-state occupation numbers for *both* end fragments. The occupation number representation can be transformed into charge and ionicity variables as well. The occupancy density matrix and its corresponding fragment energy matrix act substantially as the embedding function for the AB pair.

3.5.3 Structure of the FH Embedding Energy for Metals

From the general definition of embedding energy, setting $q_A = 0$ reduces $\bar{E}_A$ in (3.100) to a one-variable function $F_A(\tau_A) = E_A(q_A = 0, \tau_A)$ with the explicit form

$$F_A(\tau_A) = \bar{E}_A(q_A = 0, \tau_A) \tag{3.113}$$

where $W_A^{1e} \equiv W_A^{1e+} + W_A^{1e-}$ and $F_A^{gap}(\tau_A) \equiv E_{A^0}^* + \frac{1}{2} U_A^{eff} \tau_A$. The hopping functions in F_A^{1e} reduce to $f^{1e+} = f^{1e-} = \sqrt{(1-\tau_A)\tau_A/2}$, and $f_A^{2e} = \tau_A/2$. For purposes of fitting below, W_A^{2e} has been grouped into the effective, local gap $U_A^{eff} = U_A^* - W_A^{2e}$ since both are linear in τ_A when local charge neutrality is assumed. The fragment electronegativity χ_A^* also falls out of the model because of the assumed charge neutrality of each atom (i.e. $q_A = 0$ in (3.100)). This outcome is to be expected because the average number of electrons on each atom is not allowed to vary. Consequently, one is not able to determine the fragment electronegativity, or equivalently, a local Fermi level.

The FH-impurity pair potential is then identified as V_{A^0}, the component of the potential interaction that is independent of ionicity. From (3.31),

$$V_{A^0} = \frac{1}{2} \sum\nolimits_B V_{A^0 B^0} \tag{3.114}$$

Therefore, we view it as the analog of the sum of pair interactions in (3.100)

$$F_A(\tau_A) = E_{A^0}^* + U_A^{eff} \tau_A - 2\left(W_A^{1e} - V_{A_0^+} + V_{A_0^-}\right) f^{1e}(\tau_A) \tag{3.115}$$

A longstanding problem in EAM-related models is that there is no direct functional relationship between the background density $\bar{\rho}_A$ and iconicity τ_A. The actual physical identity of $\bar{\rho}$ has never been established in the existing EAM-based models [66–69], although an identity has been posited from an analogy between N-body potentials and the second-moment approximation to tight-binding models [70]. To proceed from here requires that such a relationship. By necessity, a relationship will have to be postulated, and then justified a posteriori.

To this end, note that the ionicity, as defined above, is a sum of occupation numbers, and therefore must not be larger than one. Background densities, on the other hand, are unbounded. One possibility that suggests itself from these two observations can be obtained from the basic expressions for a fragment wave function. In terms of unnormalized mixing parameters, suppose that $|A\rangle$, for the special case of equal occupancy of the cation and anion states, can be expanded as

$$|A\rangle = \frac{|A^0\rangle + \lambda_A(|A^+\rangle + |A^-\rangle)/\sqrt{2}}{\sqrt{1+\lambda_A^2}} \tag{3.116}$$

The corresponding ionicity is just

$$\tau_A = \frac{\lambda_A^2}{1+\lambda_A^2} \tag{3.117}$$

This argument suggests that we relate the background density to the unnormalized mixing coefficients, rather than to τ_A itself. That is, assume a functional relationship

$$\lambda_A = \lambda(\bar{\rho}_A) \tag{3.118}$$

Performing the appropriate substitutions transforms the two embedding functions to

$$F_A^{1e}(\bar{\rho}_A) = \frac{1}{\sqrt{2}} \frac{\lambda(\bar{\rho}_A)}{1+\lambda^2(\bar{\rho}_A)} \tag{3.119}$$

and

$$F_A^{(gap)}(\bar{\rho}_A) = \frac{1}{\sqrt{2}} \frac{\lambda^2(\bar{\rho}_A)}{1+\lambda^2(\bar{\rho}_A)} \tag{3.120}$$

For the present purposes, we accept the approximation $\lambda(\bar{\rho}_A) \approx \bar{\rho}_A$. Combining these contributions, we obtain the embedding energy

$$F_A^{FH}(\bar{\rho}_A) \approx \frac{\left(\frac{1}{2}U_A^* - W_A^{2e}\right)\bar{\rho}_A^2 - \sqrt{2}W_A^{1e}\bar{\rho}_A}{1+\bar{\rho}_A^2} \tag{3.121}$$

It is interesting to note the qualitative similarity of (3.121) to the Mishin embedding energy in (3.115), which is also a rational function. In addition, $E_{A^0}^*$ has been absorbed into the pair potential, or may be regarded as being approximated by $E_{A^0}^* \approx E_{A^0} = 0$, the value of energy for any isolated, neutral atom A. For high-symmetry, single component structures, in the limit of first-nearest-neighbor interactions, the total cohesive energy per atom reduces to the usual EAM expression

$$\bar{E}_A(\bar{\rho}_A; R_{AA}) = F_A^{FH}(\bar{\rho}_A) + L_A\phi_{AA}(R_{AA})/2 \tag{3.122}$$

In this expression, L_A is the number of first-nearest neighbors, $\bar{\rho}_A$ itself depends on R_{AA}, the first-nearest neighbor distance.

3.5.4 Coordination Dependencies of Energy Scales

Next we deduce the coordination dependencies of the fragment energy matrix elements. This dependence pertains only to single-component materials. Once again, the letter B will label neighboring lattice sites of A, but are of the same atom type as A. Only charge fluctuations between A and its first nearest neighbors are presently accounted for, but each atom is still constrained to be neutral. The one-electron hopping integral was defined above to be the expectation value

$$W_A^{1e\zeta} = -\left\langle A^0 \middle| \widehat{H}_A^{0\zeta} \middle| A^\zeta \right\rangle \tag{3.123}$$

For each nearest-neighbor interaction in perfect lattice structures, the individual matrix elements $\left\langle i \middle| \widehat{H}_A^{0\zeta} \middle| j \right\rangle$ are equal. From the nearest-neighbor count, one determines that

$$W_A^{1e\zeta} \sim \frac{L_{A^\zeta}}{\sqrt{L_{A^\zeta}}} W_A^{1e\zeta} = \sqrt{L_{A^\zeta}} W_A^{1e\zeta} \tag{3.124}$$

where $W_A^{1e\zeta}$ is a representative value of $\left\langle i \middle| \widehat{H}_A^{0\zeta} \middle| j \right\rangle$ and L_{A^ζ} is the number of resonances comprising $\left| A^\zeta \right\rangle$. In the example of Fig. 3.9, $L_{A^+} = 4$. Clearly, in the present context, there are an equal number of resonances associated with A^- as with A^+, so that $L_{A^-} = L_{A^+}$. There may be charge fluctuations around A that do not involve the charge state of A directly. In Fig. 3.9, such fluctuations have been excluded. This is the rationale for attributing only one structure to A^0. Observe that the $\sqrt{L}$-dependence in

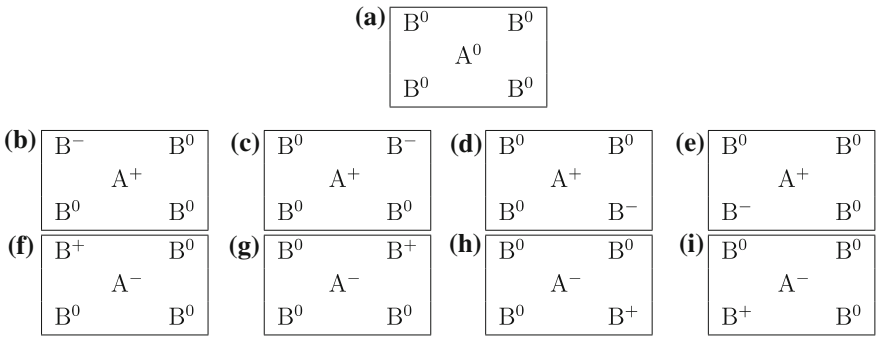

Fig. 3.9 2D square lattice around atom A with first-nearest neighbors marked as B. Fluctuations involving single charges are accounted for here between A and its first-nearest neighbors, for neutral, cationic and anionic states of A. The valence states shown are: **a** $|00000\rangle$, **b** $|+-000\rangle$, **c** $|+0-00\rangle$, **d** $|+00-0\rangle$, and **e** $|+000-\rangle$, **f** $|-+000\rangle$, **g** $|-0+00\rangle$, **h** $|-00+0\rangle$, and **i** $|-000+\rangle$. From [64]

the one-electron hopping energy also appears in a second-moment treatment of tight-binding theory of embedded-atom methods [71].

Concerning the effective fragment gap U_A^{eff}, it is simpler to analyze its individual components, U_A^* and W_A^{2e}. The structures of the energy expectation values associated with both U_A^* and W_A^{2e} are more complex than for the one-electron hopping energies. Because there are two electrons involved in the hopping processes, triplets of atoms enter into their counting arguments. These triplets are comprised of the embedded atom and any two of its first-nearest neighbors. The effect is to multiply the anion and cation coordinations together to produce a value of $W_A^{2e} = L_A w_A^{2e}$, where L_A is the number of neighbors of A. This also holds for U_A^* itself. The coordination dependencies have two- and three-body components. The two-body, two-electron hopping events involve only one nearest neighbor instead of two distinct neighbors. The neighbor-count dependence scales out of these energies. The three-body contributions require that the two neighbors be distinct from one another. Thus, we may decompose U_A^* into two- and three-body contributions with differing neighbor-count arguments: $U_A^* = U_A^{*d} + (L_A - 1)U_A^{*t}$.

3.5.5 Background Densities

Before proceeding with more general expressions for the energy scales, it is necessary to make a choice about the relationships among expansion coefficients and the background density. Our choice needs to be guided by two requirements. One is that consistency is required by the general definition of the fragment occupancies and expansion coefficients. For fragment A, the occupancies of its various charge states are given by (3.41). That is, whatever choice is made has to be applicable to coefficients appearing in definition of the background density of A and the coefficients that are to appear in the energy scales.

The second requirement is the assumption that the state of each fragment in a material be determined by its charge and ionicity. It is required because of our desire to describe each fragment as a function of just two variables, net charge and ionicity. Contracting the description of a fragment in this way assumes that the model remains viable in a physical sense. The viability has not yet been assessed.

In the present model, the background densities are not determined through an optimization process. The optimum values for these coefficients that are available through an explicit quantum mechanical treatment acquire a global character in the process. Specifically, we adopt an association between valence expansion coefficients and neighboring background densities that reflects the consistency requirement.

With these two requirements in mind, we make the choice that background densities are estimated from the background densities (or ionicities) of neighboring atoms. For some coefficient $\lambda_i = c_i/c_0$ in the neighbor list of A,

$$\lambda_i \approx \frac{\lambda_B}{\sqrt{L_A^{(ref)}}} = \frac{\bar{\rho}_B}{\sqrt{L_A^{(ref)}}} \tag{3.125}$$

for neighbor B of A in an appropriate charge-state given the charge-state of A, with the other neighbors neutral, and $L_A^{(ref)}$ as the first-neighbor count for the reference state of A. The situation for A in its anion state is completely analogous, except for the reversal of the charges on A and one of its neighbors. Combining these contributions, we obtain

$$\bar{\rho}_A = \sqrt{\sum\nolimits_{B \in L_A} \bar{\rho}_B^2} \tag{3.126}$$

The expression for $\bar{\rho}_A$ is a requirement for self-consistency.

3.5.5.1 Generalized Energy Scales

The general form of environmental dependencies for the key energy scales of this model of the FH are now developed. The generalizations require us to consider the background density dependencies at the same time. To this end, consider a three-atom, linear-chain prototype B–A–B', where B and B' need not be equivalent [35]. The charge-state occupancies for the central atom A may be expressed in terms of the occupancies of the neighbors. Thus, the anionic state of A is

$$|A^-\rangle = \left(\sqrt{n_{+-0}}|+\,-\,0\rangle + \sqrt{n_{0-+}}|0-\,+\rangle\right)/\sqrt{n_{+-0}+n_{0-+}} \tag{3.127}$$

while the cationic state is

$$|A^+\rangle = \left(\sqrt{n_{-+0}}|-\,+0\rangle + \sqrt{n_{0+-}}|0+-\rangle\right)/\sqrt{n_{-+0}+n_{0+-}} \tag{3.128}$$

These occupancies may be related to the ionicities of the neighbors of A through the expressions $n_{+-0} = n_{-+0} = n_{B^\pm} = \tau_B/2$ and $n_{0+-} = n_{0-+} = n_{B'^\pm} = \tau_{B'}/2$. The factor of two in these expressions come from the equality of valence occupancies and there being two neighbors. As above, the ionicities themselves should be expressed as functions of their corresponding $\bar{\rho}$, given in (3.117).

The general form of environmental dependence for the key energy scales of the FH model is developed next. As a preliminary step, consider the known, three-atom, linear-chain prototype B–A–B', where B and B' again need not be equivalent, the anion state occupancies for the central atom A may be expressed in terms of the occupancies of the neighbors. The anion state of A is

$$|A^-\rangle = (c_{+-0}|+\,-\,0\rangle + c_{0-+}|0-\,+\rangle)/\sqrt{c_{+-0}^2 + c_{0-+}^2} \tag{3.129}$$

while the cation state is

$$|A^+\rangle = (c_{-+0}|-+0\rangle + c_{0+-}|0+-\rangle)/\sqrt{c_{-+0}^2 + c_{0+-}^2} \tag{3.130}$$

These occupancies may be related to the unscaled ionicities (λ's) of the neighbors of A through the expressions $\lambda_{+-0} = \lambda_{-+0} = \lambda_B/\sqrt{2}$ and $\lambda_{0+-} = \lambda_{0-+} = \lambda_{B'}/\sqrt{2}$. The factor of the square root of two in these expressions come from the equality of valence occupancies.

This prototype suggests fashioning a similar approximation that accommodates the multiplicity of neighbors normally contributing to the ionic states of *A*. The states of interest are determined purely by the nearest neighbors of the atom or fragment of interest. The quantities that we wish to apply are the background densities. The approximation adopted here is that each site is characterized only by its background density. The background density contribution of the neighbors is scaled by a number of first-neighbors that appear in a reference state used to determine the properties of *A*. The neighbor count scaling comes from the fact that each valence state coefficient is shared among multiple atoms. In the example above, the count is two for each atom.

While for the single-component structures, the neighbor scaling does factor out of the energy expressions, it will not do so in more general circumstances. For this reason, in more general configurations, the neighbor-count scaling dependence of (3.124) will have to be retained. For elemental materials of interest here, applying these choices for either the $\zeta = \pm 1$ values of $W_A^{1e\zeta}$ yields

$$W_A^{1e\zeta}(R_{AB}) \approx \frac{\sum_{B\in L_{A^\zeta}} W_{AB}^{1e\zeta}(R_{AB})\bar{\rho}_B(R_{AB})}{\sqrt{\sum_{B\in L_{A^\zeta}} \bar{\rho}_B^2(R_{AB})}} \tag{3.131}$$

Because metals have special properties, the total, one-electron hopping energy combines W_A^{1e+} and W_A^{1e-}. During the fitting exercise to follow, we will keep both energies.

Under the special circumstances where the two sums of the cation and anion neighbors have comparable background densities, the combination lends itself to defining the average energy

$$\frac{\left(W_A^{1e+} + W_A^{1e-}\right)}{\sqrt{2}} = W_{AB}^{1e} \tag{3.132}$$

for pairs of neighbors.

Both W_{AB}^{1e} and $\bar{\rho}_B$ are volume-dependent by construction. This property is consistent with standard EAM-based models, where the R_{AB} distance is employed to capture the local volume dependence. If the $\bar{\rho}_B$ were all positive, then one could interpret (3.131) as a weighted average. However, there is no such restriction on $\bar{\rho}_B$, as these quantities are related to wave function expansion parameters rather than to

occupation numbers. On the other hand, it is essential that background densities appearing in $W_{AB}^{1e\pm}$, as well as for other energy matrix elements, be consistent with the definition of background density for A itself. That is, the formalism requires that the denominator of (3.131) reduce to $\bar{\rho}_A$.

The two-electron hopping energies have a similar form, but are complicated by the presence of both ionic states for each fragment. Recall that both cationic and anionic states are composed of contributions from multiple charge fluctuations with the neighbors of a fragment. Retaining all of the same assumptions for W_A^{2e}, about how the fragment ionicities enter these states, as for the one-electron hopping energies, one obtains the expression

$$W_A^{2e}(R_{AB}) \approx \frac{\sum_{B\in L_{A^\zeta}} \sum_{B'\in L_{A^\zeta}} W_{BAB'}^{2e}(R_{AB})\bar{\rho}_B(R_{AB})\bar{\rho}_{B'}(R_{AB'})}{\sqrt{\left(\sum_{B\in L_{A^\zeta}} \bar{\rho}_B^2(R_{AB})\right)\left(\sum_{B'\in L_{A^\zeta}} \bar{\rho}_{B'}^2(R_{AB'})\right)}} \tag{3.133}$$

The coordinate dependence of $W_{BAB'}^{2e}$ is more complex and is not stated explicitly here. The neighbors B and B' have been arbitrarily assigned coordinations for opposite charge states. For metals, this assignment is of no consequence, as noted above. When all of the $W_{AB}^{1e\pm}$ are equal among themselves, the diagonal two-electron hopping integrals $W_{BAB}^{2e}(B' = B)$ are equal, and the two-electron transition energies $W_{BAB'}^{2e}(B' \neq B)$ are equal, then the simpler relations derived above are recovered. Note also that, as discussed above, the diagonal energy contributions U_A^{*d} are independent of the neighbor scaling, while the off-diagonal contributions U_A^{*t} scale as $L_A - 1$.

The fragment charge state energies all comply with a single form. Then, just as for the expressions in (3.133) for two-electron hopping energy, one are obtains for either the $\zeta = \pm 1$ state of A the expressions

$$H_A^{\zeta\zeta} \approx \frac{\sum_{B\in L_{A^\zeta}} \sum_{B'\in L_{A^\zeta}} h_{BAB'}^{\zeta} \bar{\rho}_B \bar{\rho}_{B'}}{\sum_{B\in L_{A^\zeta}} \bar{\rho}_B^2} \tag{3.134}$$

The energies $h_{BAB'}^{\zeta} = \left\langle i \middle| \widehat{H}_A^{\zeta\zeta} \middle| i' \right\rangle$ are the appropriate energy expectation values for the substates of A^ζ, and for particular resonances contributing to these charge states of A. If more than one state contributes to A^0, then H_A^{00} would take on an analogous substructure.

The behavior of the effective fundamental gap is somewhat more complex though, than H_A^{++} and H_A^{--} themselves, because of the aforementioned assumption that $\left|A^0\right\rangle$ is comprised of only one resonance. Thus, the generalization for U_A^* becomes, after some rearrangement,

$$U_A^* = \frac{\sum_{B\in L_A}\sum_{B'\in L_A}\left(h^+_{BAB'} + h^-_{BAB'} - 2h^0_A\delta_{BB'}\right)\bar{\rho}_B\bar{\rho}_{B'}}{\sum_{B\in L_A}\bar{\rho}_B^2} \tag{3.135}$$

The neutral species energy h_A^0 is defined implicitly for the correct neighbor scaling so that $H_A^{00} = \sqrt{L_{A^+}L_{A^-}}h_A^0$ is assumed to hold. We have included only one resonance for the covalent state. That is the reason behind the $\delta_{BB'}$ factor multiplying h_A^0. The signs on the charge states of A have been dropped because of the equivalence of the neighbor lists in this application. If this equivalence is not assumed, then (3.88) would need to be generalized further. Obviously this definition depends on the specific resonances allowed in each component of the effective fundamental gap. A further generalization including two-electron hopping is also possible, under the same assumption that the cation and anion neighbor lists are equivalent. In that variant, $h^+_{BAB'} + h^-_{BAB'}$ is simply replaced by $h^+_{BAB'} + h^-_{BAB'} - W^{2e}_{BAB'}$.

Note that the BAB' indices on the component energies in U_A^{eff} allows for angular dependencies to appear naturally in the formulation. Angular dependencies of this kind are reminiscent of those in MEAM [66, 67] and A-EAM [68]. However, in practice, the A-EAM angular dependencies really come into refinements to H_A^{00} that arise from orbital hybridization.

3.5.6 FH Model for Elemental Ni

The FH embedding energy is applied to the Ni system in various crystalline structures and over a range of lattice constants. To construct the model we rely first and foremost on an equation of state (EOS) for a "reference structure". A reference structure is a lattice or molecule whose cold equation of state is known, and is used to calibrate the model for the material of interest. For instance, we will us the fcc structure of Ni as a reference structure from which we calculate properties of all other structures of Ni.

The EOS is used to determine an effective pair interaction for a material. To construct an EOS for fcc Ni, we determine the energy-volume curve from DFT GGA calculations with the PW91 functional [72]. For convenience, these data are fit to a third-order, Rose EOS plus an energy offset,

$$E_{Rose}(a^*) = -E_0\left(1 + a^* + \delta\frac{(a^*)^3}{\alpha + a^*}\exp(-a^*)\right) + E_{off} \tag{3.136}$$

with the scaled radial strain variable a^* defined as

$$a^* = \alpha\epsilon(R) \tag{3.137}$$

where α as an exponential decay factor. The volume dependence of this model is carried through a radial strain variable

$$\epsilon(R) = \frac{R}{R^{(ref)}} - 1 \tag{3.138}$$

With R being the distance between the embedded atom of interest and one of its neighbors. The scaling distance $R^{(ref)}$ is usually the equilibrium first-nearest neighbor distance for the reference structure associated with $L^{(ref)}$.

Other parameters include E_0 the cohesive energy and δ an expansion coefficient for the third-order term. Values for each of these parameters are deduced from fits to electronic structure calculations, are reported in Table 3.1. The offset E_{off} is provided to account for a nonzero energy of the isolated atoms in the electronic structure calculations. In a molecular dynamics simulation, this offset is discarded.

When other structures are added to the database, E_{off} is allowed to vary from one lattice structure to another because of (1) small differences in the energies of the dissociated atoms reported from the electronic structure calculations for each lattice structure, and (2) differences in the range of lattice constants included in these fits. The differences in structures and dimensionality exercise the two most fundamental features of the model developed in the formalism in the previous section, volume dependence and coordination.

In addition to the EOS, other properties are required to generate a modest database for the purpose of constructing the model using the VASP code [73, 74]. The database consists of energy-nearest neighbor distance curves for several crystal structures of different spatial dimensions.

For the database, we calculate the cold curves for the face-centered cubic (fcc), simple cubic (sc), 2d-square, and 1d-linear chain structures, as functions of first nearest neighbor distances. The fcc and sc cold curves are used in the fit. The cold curves 2d-square and 1d-chain are used to assess the quality of the fit.

Commensurate with normal practice on constructing MEAM potentials, the pair interaction is determined from a reference structure, chosen by convenience and

Table 3.1 Rose equation of state for lattice structures used to construct FH model of Ni

Lattice	Face-centered cubic	Simple cubic	2d square	1d linear chain
E_0(eV)	4.9165	4.2942	3.4962	2.0865
α	4.8858	4.8365	4.5289	4.2220
δ	0.22119	0.4041	0.3150	0.0874
R_0(Å)	2.4953	2.3292	2.2591	2.1793
E_{off}(eV)	−0.53803	−0.49792	−0.46425	−0.59421

The equation of state used (3.136) has an offset energy E_{off} to accommodate small differences in the dissociation limits for the various lattices. From [64]

efficacy. For this purpose, we choose the fcc lattice as a function of volume. From (3.122), the pair potential is given by

$$\varphi = 2(E_{\rm fcc}(R) - F(\bar{\rho}_{\rm fcc}(R)))/L_{\rm fcc} \tag{3.139}$$

where $E_{\rm fcc}(R)$ is the cold curve for the fcc structure (for Ni in this case), $L_{\rm fcc} = 12$ is the number of first nearest neighbors.

Substituting the fcc cold curve into (3.122) yields, for structure α,

$$\Delta E_\alpha(R; L_\alpha) = E_\alpha(R; L_\alpha) - \frac{L_\alpha}{L_{\rm fcc}} E_{\rm fcc}(R; L_{\rm fcc}) \tag{3.140}$$

These differences are shown in Fig. 3.10. It is evident that the simple cubic (sc) and 2d-square lattices provide the same information in the region of the fcc equilibrium 1NN distances. Consequently only one of these structures can be used for fitting. The sc structure will be used for this purpose.

The model of $\bar{\rho}$ adopted for the present purposes consists of a single exponential decay,

$$\bar{\rho}(R; L) = \sqrt{L/L^{(ref)}}\rho^{(0)} \exp - \left(\beta^{(0)}\epsilon(R)\right) \tag{3.141}$$

for an embedded atom with L first-nearest neighbors, assuming that they are all the same kind of atom. It depends parametrically on $L^{(ref)}$ first-nearest neighbors in a reference state, a constant prefactor $\rho^{(0)}$, and decay constant $\beta^{(0)}$.

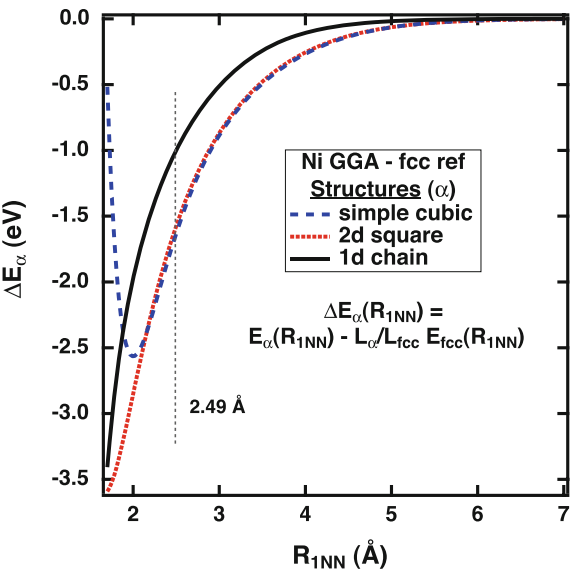

Fig. 3.10 Fitting and evaluation curves for elemental Ni, with limited coordination effects accounted for, using the fcc structure as the reference state. Values of the constants are given in the text. The 2d-square lattice and 1d-linear chain are used for evaluation purposes. From [64]

Table 3.2 Primary FH parameters for elemental Ni

$\rho^{(0)}$	$\beta^{(0)}$	W^{1e} (eV)	γ^{1e}	U^* (eV)†
2.50	1.55	2.625	2.6	6.48

$W^{2e} = \gamma^{2e} = 0$ has been assumed in these fits. From [64]
†From [26]

For Ni, the strain variable is then defined as $\epsilon_{\text{Ni}}(R) = \frac{R}{R^{(eq)}_{\text{Ni-fcc}}} - 1$, with $R^{(eq)}_{\text{Ni-fcc}} = 2.4953$ Å as the equilibrium first-nearest neighbor distance for the *fcc* reference structure. The prefactor $\rho^{(0)}_{\text{Ni}} = 1.30943$ and the exponent, $\beta^{(0)}_{\text{Ni}} = 3.0460$, are both dimensionless. More elaborate models of $\bar{\rho}$ may be adopted as needed. We chose one of the simplest for purposes of illustration.

The key pair-wise hopping function W^{1e}_{ij} is modeled as a single, exponentially decaying function,

$$W^{1e}(R) = W^{1e}\exp\left(-\gamma^{1e}\epsilon(R)\right) \tag{3.142}$$

where W^{1e} and γ^{1e} are constants. For Ni, $U^{eff}_{\text{Ni-Ni}} = 3.24$ eV is assigned the same value as the chemical hardness for atomic Ni [26]. No volume dependence was applied to this function. W^{2e} has been set to zero, as no dependence could be detected for these structures. For this work we accept coordination dependencies in the hopping and gap energies from (3.121) (Table 3.2).

3.5.7 Metallic Character of Ni Structures

With this result at hand, consider the ionicities for various lattice structures of Ni and, by extension, the metallic character of those structures, calculated in the model for Ni. Ionicities as defined in (3.117) can be assessed from the background densities for the different structures in the Ni database. Figure 3.11 shows that anticipated trends and analysis of (3.64). When the 1NN distance of a given structure is increased to large values the ionicity drops as expected. For the fcc structure, the ionicity is above 1/2 up to 5 Å. The *sc* and 2*d* structures possess smaller ranges of ionicity above 1/2, while the 1*d* chain possesses no range above that value. This trend is also consistent with the qualitative notion that structures with increased coordination should tend toward higher ionicities because of increased electron hopping that is one essential component for metallic behavior. Similarly, as the structures are compressed, ionicities tend to increase and many real materials undergo a metal-insulator transition of some variety.

For comparison with these general notions, the electronic densities of states are plotted for these structures different lattice constants in Fig. 3.12.

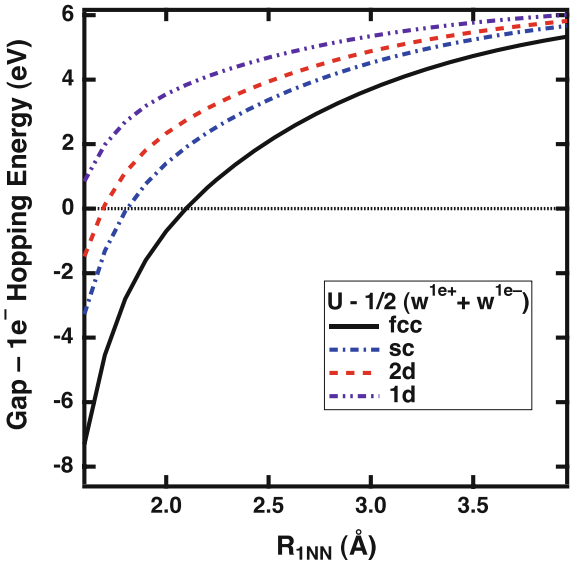

Fig. 3.11 For Ni lattices, general trend normally associated with metallic behavior is observed by differences in the point at which the difference between the chemical hardness and charge-transfer hopping determined by fitting. From [64]

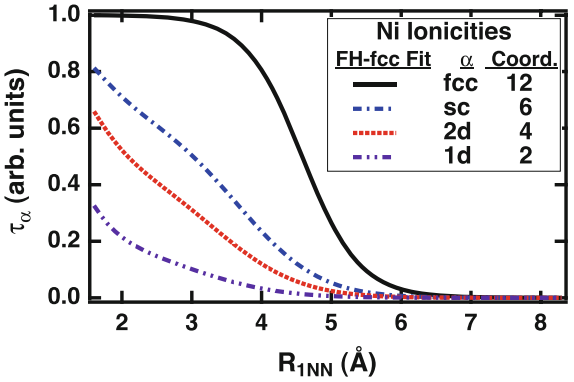

Fig. 3.12 For Ni lattices, general trend in ionicity with lattice dimensionality and nearest neighbor distance tracks with that of the difference in gap and hopping, as illustrated in Fig. 3.11. From [64]

As for establishing actual conductivity in an atomistic-scale model, much more is required that the optimized ionicity. The concepts of charge conduction, gaps, bandwidths, and gap closure are all required. Nevertheless, this model for the FH represents one step in that direction.

3.6 FH Model as a Variable Charge Model

The concept of embedding a fragment in a larger system or reservoir permeates everyday scientific thinking. The most common scenario consists of a single fragment embedded in a reservoir, that is an immutable portion of the system. More generally a system may be viewed as being composed of multiple, interacting, mutable fragments. Examples cover such diverse situations as atoms in crystalline or amorphous materials, functional groups constituting a polymeric material, a molecular wire between a source and sink, or proteins in aqueous solution. At the nanoscale, the fragment interactions may be described by a potential energy surface. Virtually every atomistic potential energy surface is built upon some concept of embedding. Here we construct a general framework of embedding, based on very few assumptions about the electronic structure of the system, and examine the nature of the charge flow that it implies.

Governance of charge transfer among fragments is embodied in the concept of the chemical potential, the change in energy of a fragment with the number of electrons attached to it. The charge, or the time-averaged number of electrons, on a fragment is determined by the condition of chemical potential equalization among fragments [20, 37, 75, 76]. When interactions among fragments are strong, the equilibrated charges on the fragments are typically fractional. Fractional charges may be thought of as originating from a superposition of states representing the different integer charge states that the fragments can assume. However, when fragment interactions are weak, the charges on the fragments are most often integer. This is always true for isolated atoms [28].

When a fragment is in contact with a reservoir of electrons, transfer of electrons between them can be described by an open statistical ensemble. If the fragment is an atom, then it is energetically advantageous for it to acquire as many electrons from the reservoir as it can bind. At the electronic structure level, for atoms at least, Perdew et al. [38] also showed that the chemical potential is constant over intervals of integer numbers of electrons. The chemical potential jumps discontinuously at the integer electron numbers. This model for the limiting case of extremely weak interactions between fragments, or between fragment and reservoir, is highly nonlinear.

At the atomistic level at least, it has been customary, although not exclusively so, to use a model of chemical potential that is linear in the amount of charge transferred. One prominent model is that of Iczkowski and Margrave (I-M) [55] which has been adopted in many molecular and materials models [41, 46, 47, 53, 77, 78]. For concreteness, consider fragment A with chemical potential μ_A, and hardness η_A [26]. Define the I-M fragment energy as

$$E_A^{\mathrm{I-M}}(q) = E_A^0 - \mu_A q + \frac{1}{2}\eta_A q^2 \tag{3.143}$$

with q being the fractional charge state of A and E_A^0 its energy in the neutral state. Then

$$\mu^{\mathrm{I-M}} = -\frac{\partial E^{\mathrm{I-M}}(q)}{\partial q} = \mu_A - \eta_A q \tag{3.144}$$

is manifestly linear in q.

The physical motivation behind (3.143) is often associated with the energies of the neutral atoms, and its cation and anion states, formed by adding or removing one electron, respectively. A three-state model is necessary [20]. The energy difference between the cation and the neutral is the ionization energy $\mathcal{I}$, while the energy difference between the neutral and the anion is the electron affinity $\mathcal{E}$. With these three points, one can fit a quadratic. When the coefficients in (3.143) are determined in this way, one finds that the chemical potential is the negative of the average of the ionization potential and electron affinity $\mu_A = -(I_A + \mathcal{E}_A)/2$ and $\eta_A = (I_A - \mathcal{E}_A)/2$.

Because the basic inputs in the I-M model are properties of isolated atoms, neither μ_A nor η_A can be zero. However, if a linear model of chemical potential with these sorts of inputs is applied to weakly interacting fragments or a fragment in contact with a reservoir, unphysical transfer of charge occurs, a problematic situation. Consequently, more physically-sound models of chemical potential are needed for general atomistic modeling purposes.

Numerous efforts and strategies have been put forth to overcome the deficiencies of the linear model [79–84]. Nevertheless most modifications stay within a linear chemical potential framework, modified to produce desired behaviors. Typically the μ_A are forced unphysically to zero at dissociation, so that optimized charges go to zero in that limit. From Perdew et al. (PPLB) [38], we know that this is incorrect: The quadratic term in $E^{\mathrm{I-M}}$ must go to zero and the chemical potential becomes piecewise constant at dissociation. In this limit,

$$E^{\mathrm{PPLB}}(q) = E_{A^0} + \mathcal{I}_A q \tag{3.145}$$

and

$$\mu^{\mathrm{PPLB}}(q) = -\mathcal{I}_A \tag{3.146}$$

for all q between zero and one, where $\mathcal{I}_A$ the ionization potential of A is $E_{A^+} - E_{A^0}$. Notably, several models recognize the importance of nonlinearity in the chemical potential [49, 54, 85, 86].

A key hurdle in all of this lies in defining fragment charges and energies. There is always an element of arbitrariness in each of these definitions. However, without these definitions, it is difficult to establish the underlying, fundamental behavior of the chemical potential that one wishes to retain at the atomistic level. In the case of charge, we will assume that the practitioner has chosen a definition of fragment charge and that it is applied consistently. Substantially all of the results presented here are independent of the details of the charge definition. The fragment energy is

a different matter. Its definition and how it relates to a chosen charge definition is central to the discussion and results.

Here one possible framework for the definition of fragment energies as functions of charges is put forth that respects known physical limits, connects to earlier models, and naturally regulates charge transfer at the atomistic level. The framework was inspired by the work of Morales and Martînez [44, 45], where charges are directly associated with expansion parameters in a wave function. Most critically for the present model, the focus is on decomposing of a system Hamiltonian into fragment contributions and its variational energy, rather than on decomposition of a ground state energy [48, 82, 87–89]. In particular, the number of electrons on the fragment Hamiltonians *must* be allowed to fluctuate. In the end, the framework falls squarely into the realm of model Hamiltonians. The framework requires a combination of three essential elements: Wave function characteristics, Hamiltonian decomposition, and charge variables.

3.6.1 Charge-Dependent Functional Forms as Embedding Energies

Potential energy surfaces with explicit charge dependence are in widespread use in atomistic simulations of systems with ionic character, including oxide ceramics. Considerable success has been achieved for structural properties and point defects in ionic materials in single phases. More sophisticated applications that require charge transfer in chemical, defect, or phase reactions pose additional difficulties for potential energy surfaces. For instance, one may be interested in a ceramic material where the cation is polyvalent. Characteristically, these processes require a physical model where the charges vary with local composition and/or defect structure.

The empirical valence bond approach (EVB) is one of the more successful methods in this regard [90], although the charge dependencies are implicit and developed only for molecular systems. Notable progress has been made through successful, mixed methodologies such as the QM/MM (quantum mechanical/molecular mechanical) methods [91], linearized tight-binding [92], Fragment Molecular Orbital method, Fragment Potential Method, and other methods that calculate forces directly through ab initio estimates.

The most common charge dependent models consistent of the fixed charge variety. A typical form in wide use for oxide ceramics is the pairwise Buckingham potential. For atomic sites designated by A and B, the total binding energy of the system is

$$E = \frac{1}{2}\sum_{A}\sum_{B\neq A}\varphi_{AB}(R_{AB}) \tag{3.147}$$

where

$$\varphi_{AB}(R_{AB}) = \frac{q_A q_B}{R_{AB}} + \mathrm{A}_{AB}\exp\left(\frac{-R_{AB}}{\rho_{AB}}\right) - C_{AB}^{(6)}/R_{AB}^6 \tag{3.148}$$

In this expression, R_{AB} is the atomic distance between A and B, A_{AB} is a scaling constant with dimensions of energy, ρ_{AB} is a length scale, $C_{AB}^{(6)}$ is the dispersion energy constant, and q_A and q_B are the respective charges on the atoms. These models have been very valuable for structural studies and explorations of simple defects. The limitations of the fixed charge approximation led to the advent of variable charge models.

For variable charge models, we begin with what we will refer to as the Iczkowski-Margrave model [55]. This model says that the atomic contributions to the energy $E_A^{\mathrm{I-M}}$ as a function of charge q are quadratic in the charge. For each atom A with chemical potential μ_A, and hardness η_A,

$$E_A^{\mathrm{I-M}}(q) = E_A^{(0)} - \mu_A q_A + \frac{1}{2}\eta_A q_A^2 \tag{3.149}$$

where the sign of q_A is the negative of the difference between the number of electrons on A in isolation and in the dimer. These atomic energies are added to a pairwise coulombic interaction.

Thus, for just two, well-separated atoms A and B of opposite charge q, there would only be the coulombic interaction $-q^2/R_{AB}$. That is,

$$E_{AB}(R_{AB}, q) = E_A^{\mathrm{I-M}}(q) + E_B^{\mathrm{I-M}}(-q) - q^2/R_{AB} \tag{3.150}$$

If one optimizes q from $\frac{dE_{AB}}{dq} = 0$, one obtains the unphysical result for atoms in the gas phase of

$$q = \frac{\mu_A - \mu_B}{\eta_A + \eta_B - 2/R_{AB}} \tag{3.151}$$

The properties μ and η belong to the isolated atoms, and so obviously do not disappear at dissociation. According to this model the AB dimer does not dissociate to neutrals as it should. The equilibration process is represented graphically in Fig. 3.8, where the gray, short-dashed line is the charge-derivative of the I-M terms and the solid black line is the negative of the charge-derivative of the coulombic term. PPLB explained the problems with this model for weakly interacting system [38]. Although the I-M model is fairly reasonable for through-bond interactions, it is not appropriate for more distant interactions (Fig. 3.13).

More advanced models have overcome some of the limitations of the original I-M model [55]. These approaches attain dissociation of atoms in the gas phase to neutral states [38] by either turning off the charge dependence in (3.149) or eliminating the linear term at dissociation. The latter route thus achieves charge neutrality, but in fact

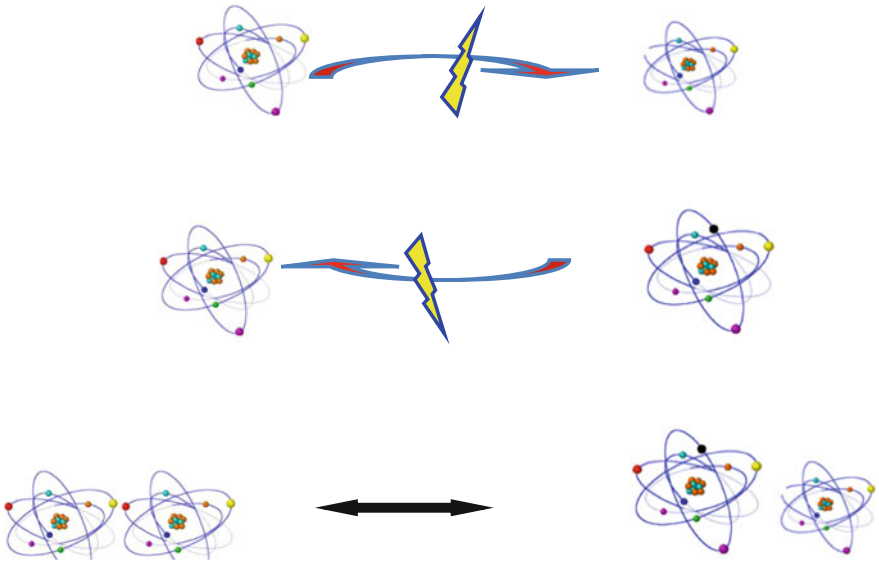

Fig. 3.13 *Top* Isolated atom releasing an electron to the "environment" (*lightning bolt*) or a reservoir to become a cation, a process requiring expenditure of the first ionization energy. *Middle* Isolated atom receiving an electron from the environment corresponding to the first electron affinity energy. *Bottom* A pair of atoms reacting to form a cation-anion pair. The processes is identified in different disciplines as a charge-separation reaction in some areas of materials science and in chemistry or electron-hole production in other areas of materials science and in condensed-matter physics

does so by eliminating the very term, the linear one, that should survive according to PPLB. Moreover, many models of the charge dependence cannot actually allow the valence state of the ions to change dynamically with the environment.

For these reasons, we have been developing a class of potentials that are founded on a combination of atom-in-molecule representations of the electronic Hamiltonian and the variational energies that come from those representations [39, 40, 93, 94]. The initial impetus for this approach was the work of Morales and Martìnez [44, 45], where charges were directly related VB expansion coefficients. These surfaces can be designed to retain multiple charge states of the constituent atoms. At the level of a 2-state model, a pair potential can be expressed in the form of a spectral model

$$E(R,q) = E_0(R) + \omega(R,q)(E_1(R) - E_0(R)) \tag{3.152}$$

where $E_0(R)$ and $E_1(R)$ are the ground- and ionic excited-state energies as functions of R, and

$$\omega(R,q) = \left(\sqrt{(1-q_0(R))q} - \sqrt{q_0(R)(1-q)}\right)^2 \tag{3.153}$$

is the occupation number for the excitation. It depends explicitly on $q_0(R)$, the ground-state charge, as a function of R. The essential functional dependence on charge differs considerably from the electrostatic form, but is consistent with more rigorous quantum mechanical reasoning [38] and numerical results [63]. Moreover the origin of the classical-electrostatic $1/R$ dependence can be identified from $E_1(R)$, but its limits of applicability, namely large R, are clearly defined. When the electron distributions begin to overlap, the natural core-core repulsion takes over. At the same time, $E_0(R)$ can be thought of as being analogous to the short-range and dispersion terms of the Buckingham potential.

The present model is consistent with PPLB [38], and its extensions to subsystem interactions of finite strength [39]. Consequently, differentiating ω is fraught with difficulties. For instance, in the context of (3.145), $q_0 \sim 0$, corresponding to a neutral H atom as the reference. At $q_0 = 0$, (3.153) collapses to $\omega(R, q) = q$, for $q > 0$. If one simply takes $\omega = q$, then $\omega' \equiv \frac{\partial \omega}{\partial q} = 1$, but one does not get the right minimum energy E_0. Instead one gets $E_1(\infty) - E_0(\infty) = 0$, an unphysical result in the present study.

The issue can be addressed in the following way. Fully expanded,

$$\omega' = 1 - 2q_0 - (1 - 2q)\sqrt{\frac{q_0(1 - q_0)}{q(1 - q)}} \tag{3.154}$$

and covers the full range from $\pm\infty$, with a plateau in the middle region, if q_0 is small. See the curves labeled with different values of q_0 in Fig. 3.14 for an illustration. In this figure, ω' was scaled by a nominal gap energy of 0.5 E_H. If (3.154) is

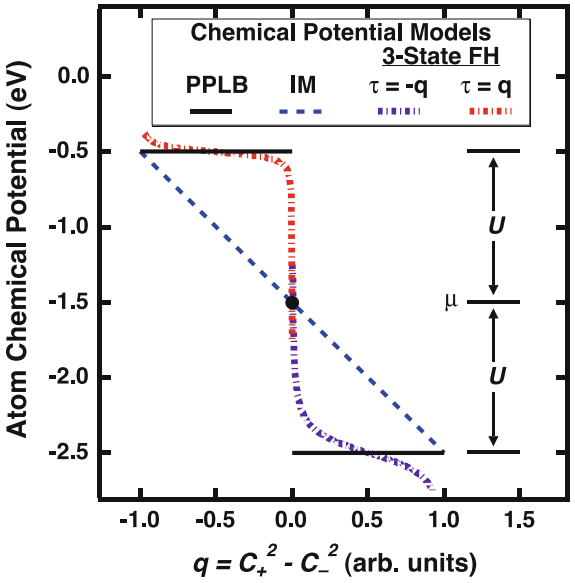

Fig. 3.14 Piecewise linear and quadratic appropriate at different material densities [38]

evaluated at $q_0 = 0$, one recovers the prior, incorrect result $\omega' = 1$ for all q. If instead ω' is evaluated at $q = q_0$, one achieves the correct result $\omega' = 0$, for all q. Thus the order of operations is important in evaluating these quantities.

In fact, the sensitivity of the charge equilibration to the details of the model is in evidence here. The present quantum-based models shed considerable clarity on how this transition should occur. Expanding (3.153) one finds that

$$\omega(R, q) = q_0(R) + (1 - 2q_0(R))q - 2\sqrt{q_0(R)(1 - q_0(R))}\sqrt{q(1 - q)} \quad (3.155)$$

The surviving linear contribution which equals the chemical potential contributions in the I-M model behaves exactly the same. However, the gap term which equals hardness terms in (3.149) are scaled by $\sqrt{q_0(R)(1 - q_0(R))}$. As $R \to \infty, q_0 \to 0$ as the square root of q_0. Thus, transitional behavior of the model between linear and locally quadratic is controlled in a well-defined way by the reference charge.

3.6.2 *Embedded Atom (EAM), Modified Embedded Atom (MEAM), and N-Body Methods*

Not all metals are equally "metallic" in character. This well-known observation applies to different phases of a single-component material such as nickel. At the atomistic modeling level, it is difficult to distinguish differences in metallic quality, which must, by extension, influence material properties of the different phases, beyond basic symmetry or geometric considerations.

Certainly several forms of atomistic models have been developed that are fairly successful at modeling at least some phases of some metals. Among these are the embedded atom method EAM [66, 67], modified embedded atom method (MEAM) [66, 67], Finnis-Sinclair (FS) [70], bond-order [95–97], A-EAM [68], and related potentials. Several of these share certain characteristics, and can even be traced in origin to one-electron theories, such as tight-binding or density functional theory. However, once an atomistic model is derived from a one-electron theory, one inherits whatever limitations are applicable to that theory.

In the EAM, FS, tight-binding (TB), and many other approaches, the total energy of a metal is partitioned between site energies and interactions between pairs of sites. The sum of the site energies is referred to as an *embedding energy*. The energy of each site is thought of as an atom interacting with an electron density provided by the remaining atoms in the system. The pair interactions occur between sites. With the sites labeled by Roman letters, the total system energy is expressed, in analogy with (3.55), as

$$\bar{E} = \sum\nolimits_A \left[F_A(\bar{\rho}_A) + \frac{1}{2}\sum\nolimits_{B \neq A} \phi_{AB}(R_{AB}) \right] \quad (3.156)$$

where F_A is the embedding, or site, energy for atom A, φ_{AB} is the pair interaction between atoms A and B, and R_{AB} is the corresponding distance between them. In methods such as EAM and FS, the embedding energy is a function of a quantity called the background electron density $\bar{\rho}_A$. It may be approximated by, for instance, a superposition of atomic electron densities from the neighboring atoms

$$\bar{\rho}_A = \sum\nolimits_{B \neq A} \rho_B^a(R_{AB}) \tag{3.157}$$

where ρ_B^a is an atomic electron density being contributed by neighboring atom B. For the embedding energy, EAM often encodes this quantity in tabular form.

Finnis and Sinclair [70] proposed a similar approach to EAM, the "N-body Potential", where the embedding energy for a chosen atom is given by the analytical form

$$F^{\mathrm{FS}}(\bar{\rho}) = -A^{\mathrm{FS}}\sqrt{\bar{\rho}} \tag{3.158}$$

This form is motivated by arguments from a one-band tight-binding model. The constant A^{FS} can depend on pairs of elements, as opposed to depending only on the element A. MEAM takes F as an analytical expression as well,

$$F^{\mathrm{MEAM}}(\bar{\rho}) = A^{\mathrm{MEAM}}\bar{\rho}\ln\bar{\rho} \tag{3.159}$$

based on coordination-dependency arguments [66, 67]. By assumption, A^{MEAM} depends only on the type of the embedded atom. One EAM model for Cu is cast in the form of the rational function [69]

$$F^{\mathrm{Mishin}}(\bar{\rho}) = \frac{\sum_j T_j(\bar{\rho}-1)^j}{\sum_j Q_j(\bar{\rho}-1)^j} \tag{3.160}$$

where the range of the sums are left undefined here, and the constants T_j and Q_j are material-dependent expansion coefficients. None of the energies scales implied in (3.160) have been identified in terms of established physical concepts. Finally, the embedding energy for angular EAM (A-EAM) [66, 67] is equivalent to

$$F^{\mathrm{A-EAM}}(\bar{\rho}) = A^{\mathrm{A-EAM}}\bar{\rho}^2 \tag{3.161}$$

In each instance, $A^{\mathrm{A-EAM}}$ is an energy scale and assumes a different value for the embedding energy of each element.

The FS embedding energy follows most directly from a more fundamental theory, tight-binding [98]. Generally, speaking though, according to the foundation for the FS N-body potential, it is accepted that the embedding energy represents electron hopping [99]. This interpretation is consistent with the relative success of these methods for good metals. The correct physical identity of the background

density remains nebulous though. In the FS interpretation, the background density is related to the hopping energies themselves, which are, in turn, related to a second-moment, density-of-states estimate of some bandwidth for a material. In MEAM, on the other hand, the background density is thought of as an actual electron density [66, 67], or electron density differences [100].

For both the background density $\bar{\rho}$ and the embedding energy, the lack of more definitive physical identification hampers further development of these concepts. Furthermore, the lack of deeper connections to first-principles concepts limits how many energy scales within the embedding energy can be given some physical identity. For instance, one might wonder if and how a "Hubbard-U" parameter [101] enters atomistic models for metals. We note at this point that this kind of parameter does appear in atomistic models for ceramics in the form of the Parr-Pearson hardness [26, 42, 43, 47–49, 52–54, 102].

The embedding energies for elemental materials, where typically all atoms are neutral, are the focus of this investigation. These embedding energies follow directly from the general FH embedding energy of (3.100), by setting the net charge of each atom or fragment equal to zero. The environmental dependencies for the different energies appearing in (3.100) are developed in progressive stages of complexity, ranging from high crystal-symmetry configurations to general configurations with no particular symmetry.

Graphically one can depict the embedding energies in the EAM and I-M models as limiting cases of the three-state FH embedding energy in (3.85). The two cuts are the $\tau = 0$ and the $\tau = q$ curves in Fig. 3.15. A final model for comparison combines the EAM and I-M models into a single model. The Streitz-Mintmire model [77, 78] approximates the combination by a direct sum of the EAM and I-M embedding energies, as also noted in Fig. 3.15. The FH model provides for an entire surface for the embedding energy as a function of both charge and ionicity.

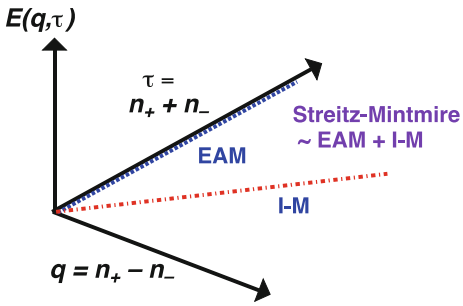

Fig. 3.15 Three-state fragment energy schematic shown as a function of charge and ionicity or occupation numbers. Elemental models fall on the $n_+ = n_-$ cut, so that EAM appears along the $q = 0$ axis. Ionic models appear along the $\tau = q$ cuts and are labeled as "I-M". The Streitz-Mintmire model [77, 78] corresponds to approximating the entire surface by the sum of the EAM cut plus I-M cut

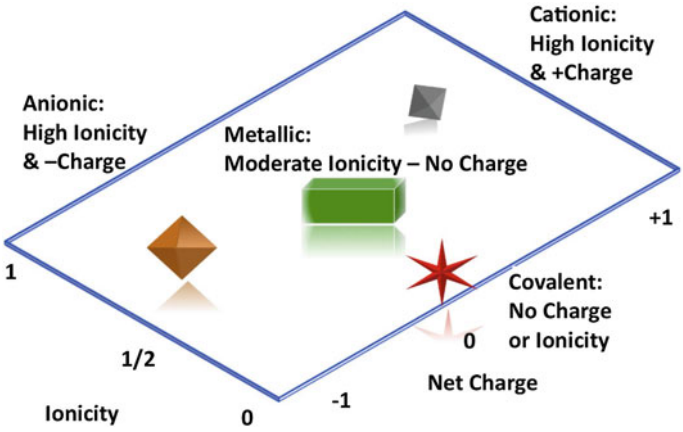

Fig. 3.16 Schematic of a three-state model indicating regions of different kinds of bonding according to different mixtures of net charge and ionicity

3.7 FH View of Bonding

The general outcome of the Fragment Hamiltonian view of materials is the possibility of capturing all three major categories of bonding, covalent, metallic and ionic, in a single model. These categories are shown in Fig. 3.16 as they relate to net charge and ionicity. Broadly speaking, covalent materials occupy the neutral state, and so have both low net charge and ionicity. Ionic materials occupy states where the ionicity is high, but occupancy of the cation or anion states dominates, leading high net charge as well. Metallic materials occupy a low net charge state, but still possess high ionicity because of the charge fluctuations in the system. In this effort, metallic bonding within the context of embedded atom methods is the subject of primary interest. With the foundation of the Fragment Hamiltonian approach, different components of the embedding energy can be derived and examined.

3.8 Environment-Dependent Dynamic Charge (EDD-Q) Model Potentials

As we have previously noted, if we wish to capture the insights gained from the FH approach in MD simulations some interatomic potential must be provided. Recall that in our examination of the FH approach we have restricted ourselves to examination of the so-called Ehrenfest potential of (3.7a). In the development of interatomic potentials, we will continue to neglect the non-B.-O. terms in the electron-nuclear potential while concentrating on the electronic Hamiltonian contribution to the nuclear forces for a given nuclear configuration. Recalling from (3.21) that a system can be broken into an arbitrary number of fragments each associated

with some number of electrons, we are lead to conceive of a model interatomic potential that has a varying charge. When this insight from the FH is combined with an AIM perspective, it becomes clear that the varying charge must be affected by the local bonding networks in which a given atom engages. We think of this bonding network as the local 'environment' for an atom. Again, in the FH framework, an atom itself, whether in a molecule or not, can be considered as a 'fragment'.

However, for an interatomic potential it is not practical to consider multiple charge states and potentials for each, but rather the charge on a given atom is better described by a partial charge that reflects the molecular, or periodic, environment in which the atom is situated. It is also necessary to choose some fiduciary quantity that will represent to atomic charge on the atom and some functional form, or forms, that will be used to capture the behavior of the atomic charge as a function of its environment. It is this set of choices that define the environment-dependent dynamic charge (EDD-Q) potential models that will be explored in this section.

The general form of an EDD-Q potential between two atoms A and B is:

$$V_{AB} = \frac{q_A\left(\{\vec{R}_N\}\right) q_B\left(\{\vec{R}_N\}\right)}{\left|\vec{R}_{AB}\right|} + \Phi\left(\vec{R}_{AB}\right) \tag{3.162}$$

where the partial charges q are explicit functions of position and the Coulomb interaction is coupled with a supplementary potential (Φ). That supplementary potential can take on a number of forms as can the functions that capture the environment dependence of the charge. Further, an embedding term defined in terms of the environmental dependent charge is included in the EDD-Q functional form thereby ensuring consistency with the metallicity and ionicity requirements of a FH-AIM approach towards potential parameterization. Thus, within the EDD-Q framework and by analogy with (3.159), the energy of an atom A is given by

$$\bar{E} = F(q_A) + \frac{1}{2}\sum\nolimits_{B\neq A}^{N} \Phi(R_{AB}) + \frac{1}{2}\sum\nolimits_{B\neq A}^{N} \frac{q_A q_B}{R_{AB}} \tag{3.163}$$

where q_A is the charge of atom A, and is an explicit function of the chemical environment (coordination, bond distances, bond angles) of A. The ability to accurately incorporate chemical dependent variation in partial atomic charges is a fundamental requirement for the parameterization of the EDD-Q potential. In this regard, Mulliken or Lowdin populations or Bader charges that are obtained from quantum chemical calculations can serve as worthy proxies for atomic charges, provided these are obtained from appropriate underlying training-sets that allow for a significant sampling of different chemical environments that the constituent atoms will experience when in a gas-phase cluster or in a liquid or solid-state. Equally importantly and in addition to parameterizing the atomic-charges as a function of chemical environment, the potential must reflect an underlying energy landscape that is characteristic of the phase-space that will be sampled by the atoms/molecules at different thermodynamic conditions. Towards this end, the following strategy for developing a framework for EDD-Q parameterization is provided below:

1. Selection of suitable gas-phase molecules/clusters that represent the simplest building blocks of the systems to be modeled.
2. A Hartree-Fock/DFT based investigation of the normal modes of vibrations of the chosen molecules.
3. Deformation of the molecules along their respective normal modes and cataloging the change in partial atomic charge of the constituent atoms as a function of chemical environment. Alternately, depending on the size of the molecule/cluster, the charge-catalog can be generated based on affine transformations of the molecule.
4. Simultaneous catalog of the change in energy of the molecules as a function of respective deformation.
5. Parameterization of the atomic charges based on the charge-catalog; concurrently, using the obtained energies (in step 4), parameterize the embedding function as well as the supplementary potential term (Φ).

In the EDD-Q formulation, as noted earlier, the charge on an atom is a function of its coordination-i.e. the number of bonds it forms. In particular, the total atomic-charge is decomposed into 'bond' charges and the net charge on an atom is given by the sum over all bond-charges. The bond-charge is a function of the chemical environment of both participating atoms. The bond cutoff distance is obtained based on the distance at which the atom-charge saturates. For a bond formed between atoms A and B, the bond-charge (q_{AB}^{bond}) is a function of the number of nearest neighbors of $A(L_A)$ and $B(L_B)$, the interatomic distance (R_{AB}), and the bond-angles that arise due to the remaining nearest neighbors of A and $B(\theta_{BAC}$ and $\theta_{ABC'})$, where C and C' represent the neighbors of A and B respectively. The embedding term is defined in terms of atomic charge rather than electronic density, though the corresponding functional form is similar to the original EAM description. The functional forms of the charge variation and the Hamiltonian describing the interatomic interactions are given in the Appendix.

The variations in atomic-charges are fitted to a functional form that is expressed in terms of bond-distances and bond-angles. The other potential parameters are obtained with respect to the variation in the energies of the different molecules. In particular, both charge as well as the potential parameters represents best-fit values that can faithfully reproduce the underlying variations for all considered molecules within 5 % error.

The above procedure represents a hierarchical approach to developing predictive interatomic potentials, where first-principles information derived from the smallest structural units that are representative of the modeled system is only used for parameterization. To illustrate the EDD-Q potential development, we focus on two distinct materials-systems namely water and silica (SiO_2). In Table 3.3, we present a summary of the parameterization procedure adopted for each of the above systems.

First we start off by examining the development of EDD-Q for water, and then follow with silica.

Table 3.3 Summary of the parameterization training sets and the level of theory in developing EDD-Q for water and silica

	Training set	Deformation procedure	Level of theory	Basis set	Population analysis	Cutoff (Å)
Water	$(H_2O)_k$; k = 1, 2	Normal modes	Unrestricted Hartree Fock (UHF)	6-31G**	Lowdin	1.5
Silica	$H_2Si_2O_7$, $(SiO_2)_k$; k = 2, 4	Affine transformation, normal modes	DFT-B3LYP	6-31G**	Mulliken	3.2

3.8.1 EDD-Q Potential for Water

In this study we present the implementation of the EDD-Q formulation for water. Specifically, we apply it to the structure and energetics of neutral water clusters $(H_2O)k$, k = 3, −20. Importantly, we focus on water as a paradigmatic small molecule system of immense practical applications to biomolecular and materials applications. Water also represents an extremely challenging test system for any interatomic potential due to its strong polar effects arising from underlying charge transfer and charge polarization and associated many body effects. Many 'water potentials' are available in literature, of which the best known empirical potentials include SPC [103, 104], TIP3P [105, 106], TIP5P [107] while potentials such as MCDHO [108], NCC [109], MCY [110], NEMO [111] and POL5 [112] rely on first-principles data for developing parameters. In almost all the potentials, the parameterizations were carried out with the implicit assumption that the basic structural unit consists of the water monomer/molecule (notable exceptions being Halley et al. [113], Corrales et al. [114], Voth et al. [115, 116]). Typically, the molecule is represented by a collection of point charges to yield the right dipole moment and minimum energy structures. Simpler potentials hold the geometry of the water molecules rigid [103–107], while more realistic potentials allow for OH bond flexibility [41, 117–120] modeled as harmonic and/or anharmonic oscillators. Further, polarization effects are also introduced via a polarizable term as incorporated in the diffuse charge potential model [121, 122].

In contrast to the above described potentials, using water monomers and dimers (Fig. 3.17) as the training set molecules, the EDD-Q potential is parameterized with respect to Lowdin charges (as obtained from ab initio calculations) that vary as a function of the atoms' chemical environment. Figure 3.18a shows a comparison between UHF obtained Lowdin charges and the EDD-Q charges for a representative atom within the water monomer and the dimer when deformed along respective normal modes. Note the EDD-Q fits emulate the UHF data very well. The charge and potential parameters are provided in the appendix. Table 3.4 provides a comparison in the EDD-Q predicted properties of the monomer and dimer with the corresponding UHF values, while Fig. 3.18b represents the variation in the energy

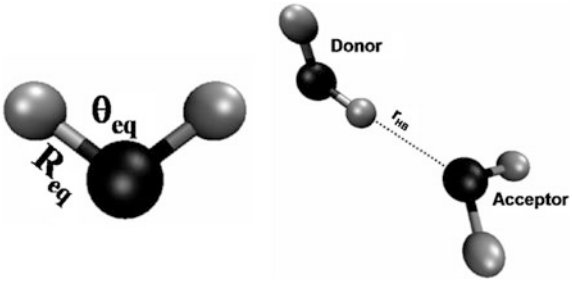

Fig. 3.17 Illustration of a water monomer (*left*) and a dimer (*right*). For the monomer R_{eq}, θ_{eq} represent the equilibrium bond distance (0.945 Å) and bond angle (105.5°) respectively. For the dimer, r_{HB} represents the hydrogen bond between the 'donor' and the 'acceptor' molecules that constitute the dimer

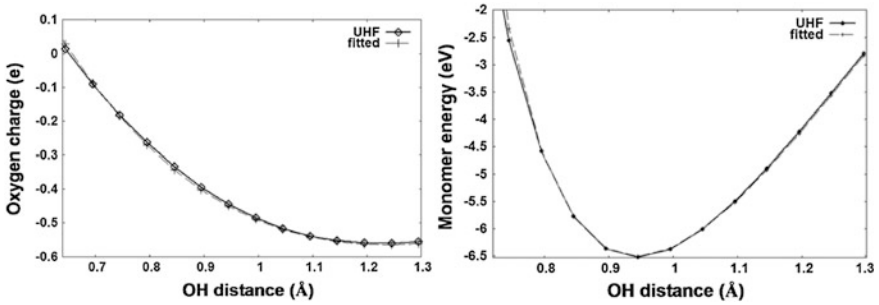

Fig. 3.18 *Left* Oxygen charge as a function of OH distance for a symmetric variation at the equilibrium angle (θ_{eq}) = 105.5° of the water monomer. *Right* UHF and predicted energies for a water monomer for the symmetric mode at the equilibrium angle. In the above figures 'fitted' refers to the EDD-Q data

Table 3.4 Equilibrium properties of the monomer and dimer

Monomer		
	Predicted (EDD-Q)	UHF
R_{eq} (Å)	0.943	0.943
θ_{eq} (deg)	105.5	106.0
E_{eq} (eV)	−6.51	−6.52
Dimer		
r_{HB} (Å)	1.94	2.04
r_{OO} (Å)	2.89	2.98
$\mathcal{U}$ (kcal/mol)	−5.86	−5.50

$\mathcal{U}$ refers to the interaction energy of the dimer, which is obtained by subtracting the two monomer equilibrium energies from the total energy of the dimer

Table 3.5 Equilibrium properties of water clusters (k = 3, −5). Θ is the angle among three water molecules

	EDD-Q	TIP5P	TIP4P/FQ	UHF
Trimer-cyclic				
$\mathcal{U}$	−13.74	−14.99	−12.58	−17.10
$<r_{OO}>$	2.72	2.77	2.91	2.78
$<\Theta_{HOH}>$	105.3			105
$<r_{OH}>$	0.95			0.95
Tetramer-cyclic				
$\mathcal{U}$	−27.31	−28.43	−23.64	−29.1
$<r_{OO}>$	2.83	2.67	2.76	2.79
$<\Theta_{HOH}>$	105.3			105.3
$<r_{OH}>$	0.95			0.95
Pentamer-cyclic				
$\mathcal{U}$	−36.03	−38.12	−32.96	−37.7
$<r_{OO}>$	2.85	2.66	2.77	2.87
$<\Theta_{HOH}>$	105.3			105
$<r_{OH}>$	0.95			0.96

All distances in Å, angles in degrees, energy ($\mathcal{U}$) in kcal/mol

of the monomer for the symmetric mode deformation as obtained from EDD-Q and the underlying UHF calculation.

Having demonstrated the ability to accurately emulate the properties of the training sets, we turn our attention to larger water clusters. It has been shown that the most stable structures of small water clusters correspond to 2-D cyclic structures, where each molecule serves as both an acceptor and a donor, while 3-D structures are energetically favorable for larger clusters. This crossover is seen for the water hexamer and larger clusters. In this regard, in the following section, we present EDD-Q predictions of the energetics and geometry of water clusters and compare them with the underlying UHF calculations. Note that for locating the energy minimum structures, the modified BFGS routine was used [123].

Trimer-Pentamer Table 3.5 provides the properties of the three clusters as predicted by EDD-Q; EDD-Q predicts the cyclic structure to be the most stable structure for all three clusters (i.e. trimer, tetramer and pentamer). For comparison we provide relevant data corresponding to other potentials as well as first-principles/ab initio data. The notation <> in Table 3.5 reflects the fact that all quoted distances are averaged over the cluster structure. Figure 3.19 depicts the energy minimized structures of the trimer, tetramer and the pentamer molecules. Note that the trimer chain structure as shown in Fig. 3.19b is a metastable structure. Also, the reported interaction energy $\mathcal{U}$for all molecules in Table 3.4 (and following tables) is obtained by the difference between the total energy of the molecule and the sum of the energy of the monomers that constitute the given molecule.

Water Hexmer Ab initio calculations [124] have established that the (i) 3-D structures are energetically more stable than 2-D structures and (ii) the 3-D cage, prism and book structures (Fig. 3.20) corresponding to the water hexamer are

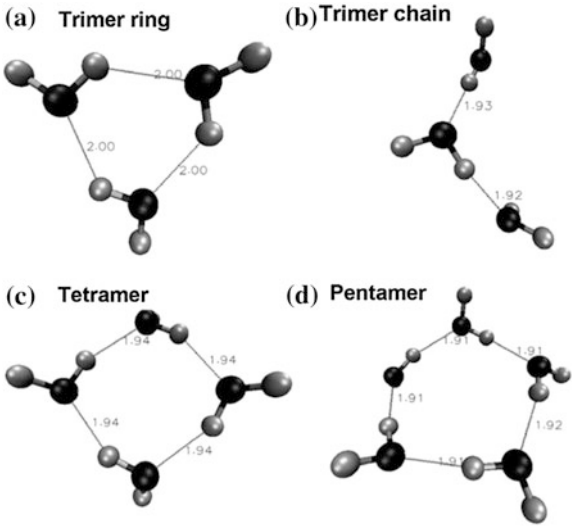

Fig. 3.19 Equilibrium geometries of the different energy minimized $(H_2O)_k$ molecules when k = 3, 4, 5. Interatomic distances in Å

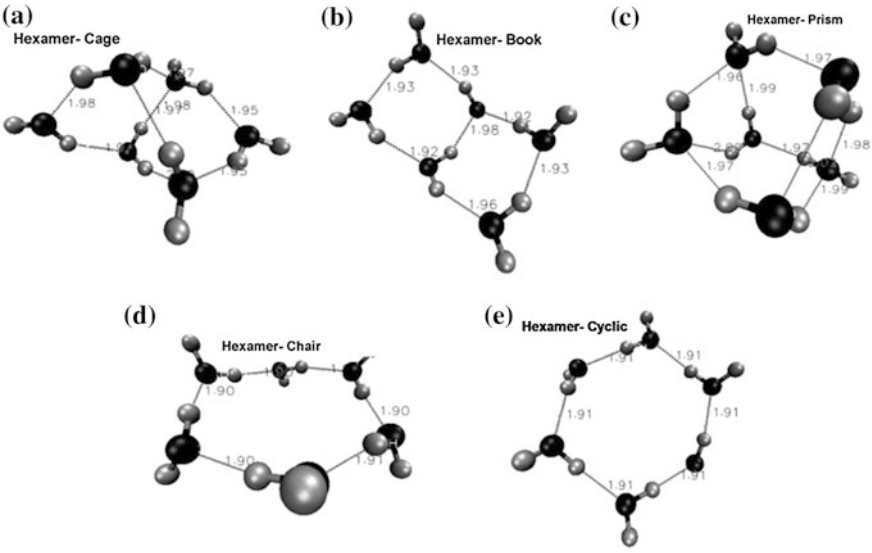

Fig. 3.20 Equilibrium geometries of the different water hexamers. Interatomic distances in Å

energetically degenerate. EDD-Q successfully predicts both trends and further also predicts the structural properties (bond length, bond angles) accurately. This ability to correctly predict the crossover from 2-D to 3-D structures for the hexamer

Table 3.6 Interaction energy (kcal/mol) of the different hexamer polymorphs

	EDD-Q	TIP5P	TIP4P/FQ	UHF
Cage	−46.50	−45.39	−45.38	−48.60
Book	−46.49	−46.68	−40.15	−48.51
Prism	−46.47	−45.59	−39.30	−48.56
Chair	−44.07			−47.93
Cyclic	−43.92	−47.31	−41.37	−47.34

Table 3.7 Interaction energy (kcal/mol) of water clusters (k = 7, −9)

	EDD-Q	TIP5P	TIP4P	UHF
Heptamer (a)	−58.26	−57.91	−58.27	−60.53
Octamer (D2d)	−74.32	−72.54	−73.09	−76.01
Nanomer (a)	−84.12	−83.62	−82.40	−85.05

distinguishes EDD-Q from other interatomic potentials available for water. A comparison of EDD-Q energetics and relevant data from previous investigations is given in Table 3.6.

Beyond the Hexamer Table 3.7 provides equilibrium properties of water clusters for k = 7–9 as obtained by using EDD-Q. The most stable energy minimum heptamer ($(H_2O)_{k=7}$) structures are shown in Fig. 3.21. The EDD-Q obtained energetics for these two structures agree with ab initio data [124]; specifically, EDD-Q predicts heptamer-a (see Fig. 3.21) to be more stable than heptamer-b by 1 kcal/mol, consistent with UHF predictions.

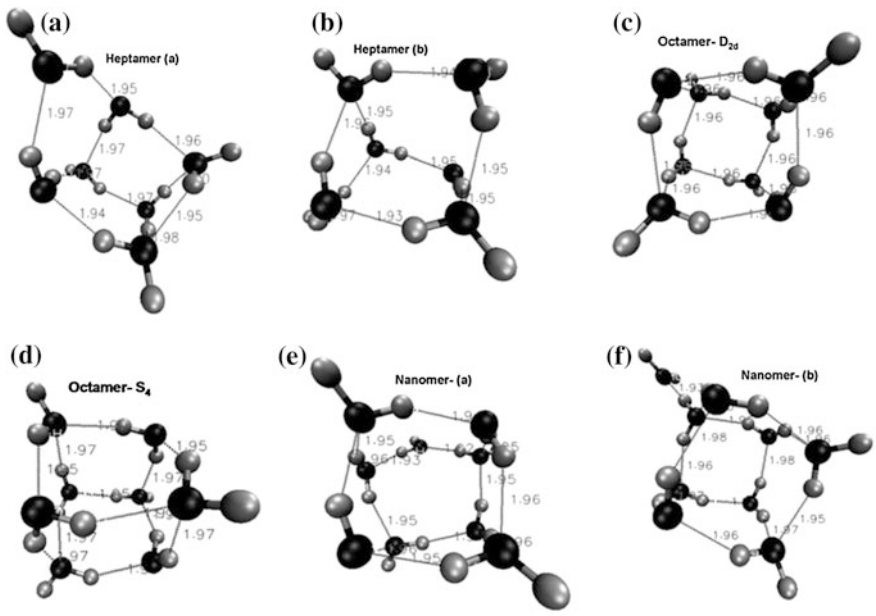

Fig. 3.21 Equilibrium geometries of the different energy minimized $(H_2O)_k$ molecules when k = 7, −9. Interatomic distances in Å

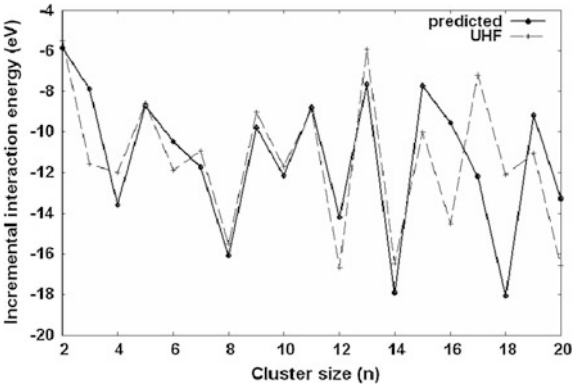

Fig. 3.22 Incremental interaction energies of water clusters ($\Delta E = E_{k+1} - E_k - E_1$) as a function of cluster size k

The two most stable octamer structures in our work correspond to a cubic structure with D2d symmetry and a cubic structure with S4 symmetry. This agrees well with other ab initio predictions [124] as well as UHF data.

The global minimum water nanometer structure can be described in terms of a pentamer and a tetramer ring connected by hydrogen bonds (nanomer-a in Fig. 3.21) [124]. EDD-Q predicts this structure to be the most energetically favorable too. Another structure, namely nanomer-b is also shown to be energetically stable, but metastable with respect to nanomer-a.

For locating the energy minimum structures for larger clusters ($k > 9$), we use geometries predicted by ab initio calculations [124] or the classical potential TIP5P [107] as starting configurations. The energies of the resulting energy-minimum structures for k = 10–20 agrees well with ab initio data, unequivocally demonstrating the efficacy of the EDD-Q parameterization method. A summary comparison of EDD-Q and ab initio data [124] is given in Fig. 3.22. As evident from the figure, the alternation in stability of the cluster depending on n is captured by EDD-Q, consistent with ab initio predictions. In particular, the enhanced stability of even n-mers relative to odd n-mers is described well by EDD-Q.

3.8.2 EDD-Q Potential for Silica

In this section, we present a potential for silica that is developed in a fashion similar to the EDD-Q potential for water. We assess its capabilities, by examining its ability to describe silica nanoclusters with respect to DFT calculations. Specifically we study global-energy minimum structures of small silica nanoclusters ($(SiO_2)_k$, k = 1, −27) as identified by [125]. In addition, we also compare EDD-Q to select potentials that were also developed for silica. These potentials include (i) the BKS potential [81], (ii) the TTAM potential [126] and the (iii) FB potential [127], all of which are two-body potentials (with fixed atomic charges) of the Buckingham form,

$$V_{AB} = \mathcal{A}_{AB} \exp(-b_{AB} R_{AB}) - \frac{\mathcal{C}_{AB}}{R_{AB}^6} + \frac{q_A q_B}{r_{AB}} \tag{3.164}$$

The BKS and TTAM potentials were primarily parameterized with respect to bulk silica polymorphs while FB parameters were chosen to yield the energetics and geometries of small silica nanoclusters as predicted by DFT with the B3LYP functional and a 6-31G* basis set.

In order to obtain 'environment-dependent' charges of the constituent atoms, we use pyrosilicic acid ($H_6Si_2O_7$) as our training system. Pyrosilicic acid (Fig. 3.23) consists of a central-bridging oxygen, two 4-coordinated silicon atoms, and six end oxygen atoms that are hydrogen terminated. One could think of the central oxygen and the two silicon atoms as being representative of 'bulk' atoms in their normal coordinated states. Using the Mulliken population as a surrogate for atomic charges, we obtained the charge on the central oxygen atom (O_c) for some select geometries at the restricted Hartree-Fock (RHF) 6-31G** level of theory. Specifically, fixing one half of the molecule with respect to O_c, we varied the other Si_1–O_c distance between 1.25–2.4 Å and the O_c–Si_1–O(1,2,3) angles from 60° to 150°. Next, we defined a Si–O bond-charge, which was a function of its bond distance and the bond angles with respect to the other oxygen neighbors of the silicon atom. In our case, the neighbor cut-off distance was defined to be r_c = 2.4 Å. More information on the charge functional forms is given in the appendix. Figure 3.24 depicts the Mulliken population on the central oxygen atom obtained as a function of Si–O distance when the O–Si–O angles equal the equilibrium angle of 106° and clearly indicates the accuracy of our fitting procedure.

In order to obtain the potential parameters, we use the energetics and structures of small SiO_2 clusters ($(SiO_2)_k$, k = 2–4) clusters (Fig. 3.25) as our training sets. Specifically, we use DFT with the B3LYP exchange-correlation functional and the 6-31G* basis set, same as that of FB, to calculate the ground-state structure and binding energy of the above clusters, as well as the energies for isotropic and uniaxial (along the primary axis) deformations of the above clusters (90–120 % of the energy-minimized structure). The binding energy was calculated with respect to the triplet-state isolated oxygen and silicon atoms.

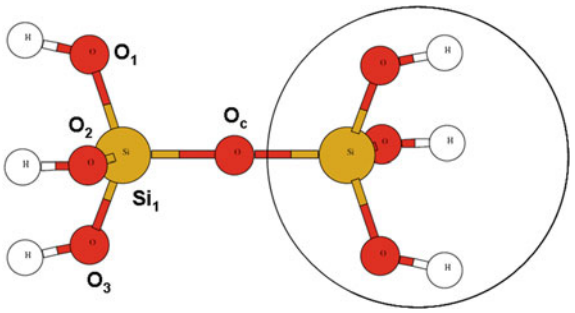

Fig. 3.23 Geometry of the pyrosilicic acid molecule

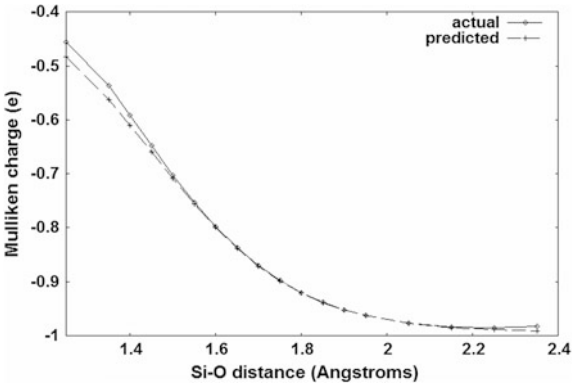

Fig. 3.24 Variation of the Mulliken charge on the central oxygen atom as a function of Si–O distance

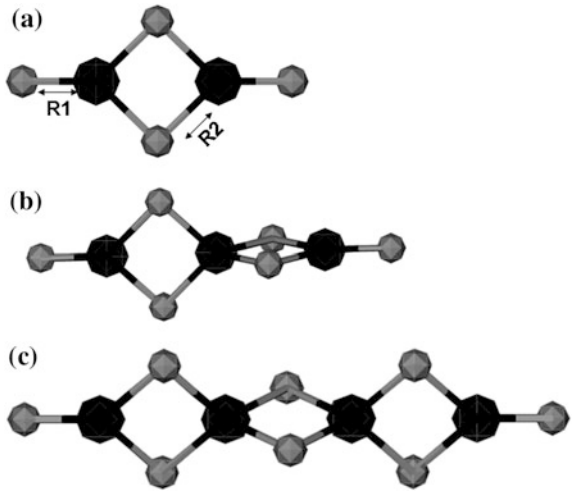

Fig. 3.25 Geometries of **a** $(SiO_2)_2$, **b** $(SiO_2)_3$, **c** $(SiO_2)_4$ molecules. *Black* and *grey spheres* represent silicon and oxygen atoms respectively. R1 and R2 represent the corresponding Si–O distances respectively

We need to point out that the reason for using two different training sets for our charge and potential parameterizations was merely to combine two different sets of information. In other words, we use the charge fluctuations for atoms that were representative of the bulk, while the energy parameterizations are carried out on small clusters. Also, we use RHF level of theory to calculate the charges and DFT to calculate the energetics, to maintain a certain level of consistency with the FB approach, where the (fixed) charges that they used were obtained from a RHF-level calculation performed on the geometry optimized of the pyrosilicic acid, while using DFT for energy parameterization.

Table 3.8 lists the ground state energy and structural properties of the three model clusters as obtained by DFT and EDD-Q. In our work, we use the modified BFGS procedure to obtain energy-minimum structures. As is obvious from the above table, EDD-Q does a good job of matching both the DFT binding energies as

Table 3.8 Energies and structural properties of the training sets when energy-minimized using DFT and the EDD-Q potential

	Binding energy (eV/SiO_2)		R_1 distance (Å)		R_2 distance (Å)	
	DFT	EDDQ	DFT	EDDQ	DFT	EDDQ
$(SiO_2)_2$	−13.97	−14.00	1.51	1.50	1.68	1.69
$(SiO_2)_3$	−14.82	−14.81	1.51	1.50	1.68	1.69
$(SiO_2)_4$	−15.23	−15.18	1.51	1.50	1.68	1.69

well as structural quantities R1 and R2 as defined in Fig. 3.25 (within 1 %). Note that the potential parameters are given in the appendix.

Next, we examine larger silica clusters. Bromley and co-workers [125, 127, 128] have systematically examined the energy-minimum structures of silica nanoclusters via painstaking scanning of the potential energy surface (PES) using DFT. In addition, a wide variety of nanostructures including nanorings, nanochains and nanotubes were shown to be energetically stable structures, though not necessarily the ground-state structures. In this regard, as a first step towards validating our procedure, we check the ability of EDD-Q to match DFT results in obtaining the energetics and structural properties of the reported ground-state structures as well as some select nanostructures, rather than attempting to do a full-blown computationally intensive global optimization search for the energy-minimum structures of silica. For comparison, we later present the energetics and structural properties of the ground-state structures when optimized with other silica potentials (BKS, TTAM and FB).

For very small cluster sizes ($k \leq 6$ SiO_2 units), it has been shown that the nanochains are the energetically preferred structures, while for slightly larger clusters ($7 \leq n \leq 12$), the ground-state structures do not follow any consistent growth pattern as discussed by Flikkema and Bromley [128]. For $k > 13$, a distinct pattern emerges, such that for $13 < k < 23$, columnar structures are preferred, while for $23 \leq k < 27$, the ground-state structures are disk-like and much more compact in comparison. The columnar structures are characterized by four-membered Si_4O_4 rings (Fig. 3.26a), while the disk-like structures (Fig. 3.26b) are characterized by five-membered Si_5O_5 ring cages, and all the above structures characterized by the presence of some singly-coordinated oxygen and tri-coordinated silicon defect states (for more structural details refer Bromley [125]). Further, an odd-even alternation in the relative energy stability is seen for the columnar structures as reflected in the sequential energetic change $\Delta E = BE_k - BE_{k-1}$, (BE is the binding energy of the given cluster expressed in eV/SiO_2) in adding a SiO_2 unit to each cluster. Greater increments in binding energy are realized by adding a unit to odd numbered structures than to even numbered structures (see Fig. 3.27a). In contrast, for the disk-like structures, (i) alternation is insignificant, (ii) for $k = 26$, ΔE is higher than the corresponding odd-structures.

Using the above ground-state structures as our initial starting structures, we energy-minimized them using the EDD-Q potential; the variation in binding energy as well as ΔE with n is given in Fig. 3.27b. As evident from Fig. 3.26b, the

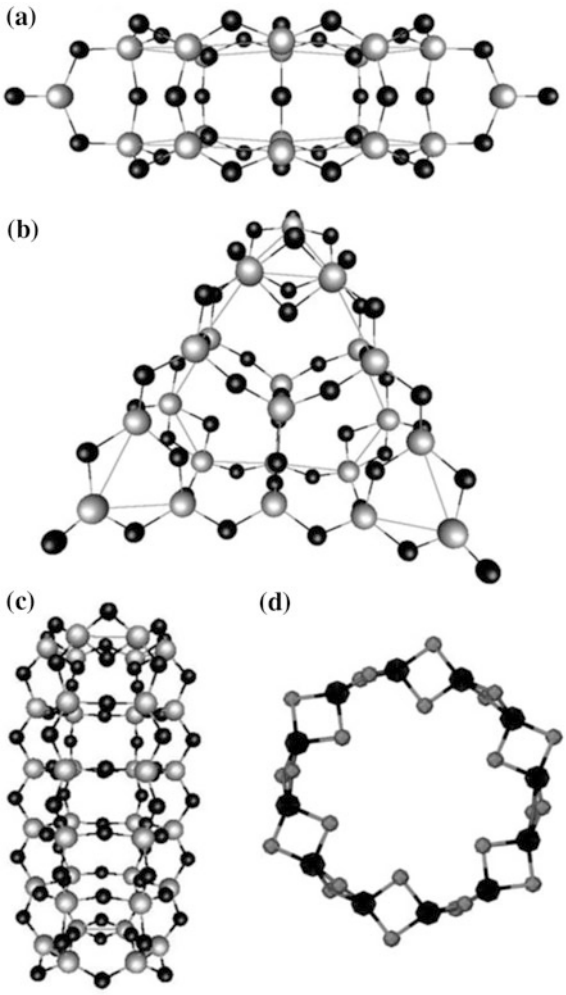

Fig. 3.26 Geometries of **a** columnar structure (k = 18), **b** disk-like structure (k = 23), **c** nanotube (k = 36), **d** nanorings (k = 12). *Black* and *grey spheres* represent silicon and oxygen atoms respectively

odd-even alternation in ΔE of the columnar structures is also observed; it has to be pointed out that with EDD-Q, the columnar and disk-like structures are isoenergetic for k = 24 and 25 (differences in binding energies <0.01 eV/SiO_2), and the disk-like structures are more stable when k > 25. Though DFT results show that the structural transition occurs at k = 23, our results are still consistent with the discussions in [115], which point out that structural transition from an elongated to a more compact form in silica as well as silicon clusters occur in the range k = 23–25.

Bromley et al. had fitted the variations in binding energies using a polynomial of the form ($E(\mathrm{k}) = E_{res} + a\mathrm{k}^{-b}$) for the smaller (k < 13) and larger (k ≥ 14) clusters separately and noticed that for the larger clusters, b equaled 0.333, while it was larger for the smaller clusters. We noticed a similar trend as given in Fig. 3.27b,

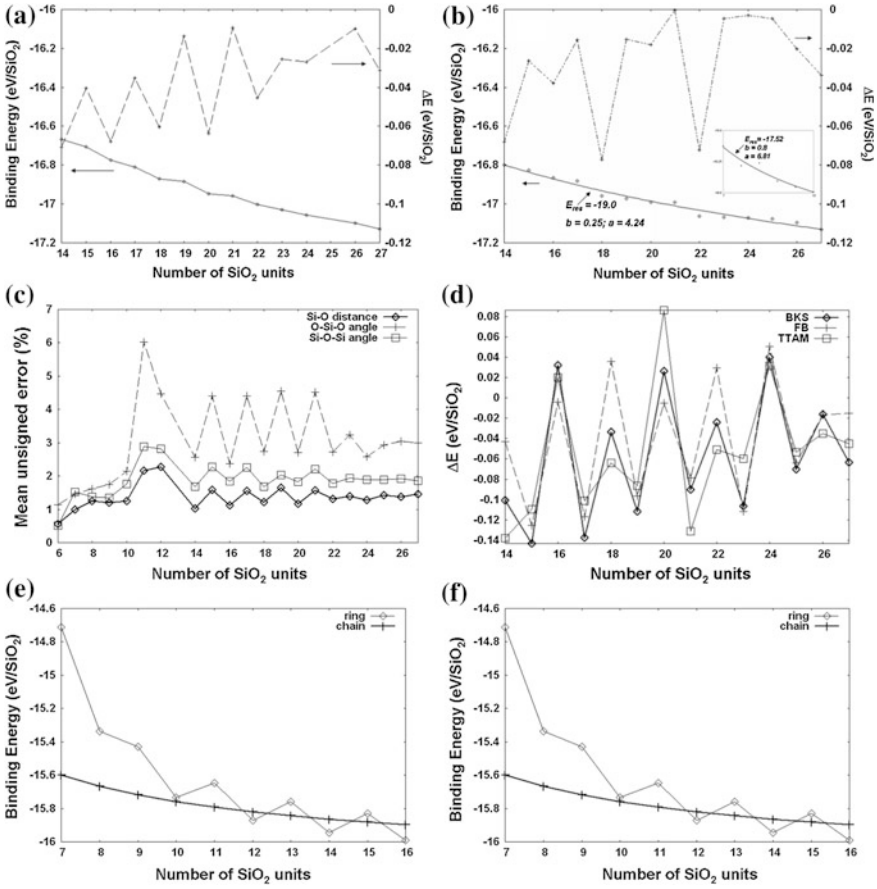

Fig. 3.27 Variation (as a function of n) of **a** DFT BE and ΔE of the ground state structures ($14 \le k \le 27$), **b** EDD-Q BE and ΔE of the ground state structures, **c** structural mean unsigned error (MUE) of the EDD-Q ground state structures, **d** ΔE of the ground state structures ($14 \le 1k \le 27$), **e** EDD-Q BE of nanochains and nanorings, and **f** BE of the nanotubes and select nano-cage structures (given in *inset*)

with the predicted E_{res} for the larger-sized clusters similar to the experimentally measured energy of α-quartz as well as that obtained by Bromley et al. [115].

In order to characterize the structural details of our energy-minimized structures, we compared the (i) Si–O bond distances, (ii) O–Si–O angles and (ii) Si–O–Si angles with the DFT-optimized structures. Specifically, for each cluster size, we tabulated the various bond-distances and bond-angles in each structure and calculated the mean unsigned error (MUE) with respect to the DFT structures for the above quantities (Fig. 3.27c). For all n, the errors in Si–O bond distances and Si–O–

Si bond angles are below 2 and 3 % respectively, while the deviation in O–Si–O bond angles are slightly larger. Also, for the intermediately-sized clusters ($14 \leq k \leq 22$), the errors seem to be larger for the odd-structures, while for the large clusters, the errors relatively diminish and do not seem to fluctuate.

In order to compare the performance of EDD-Q with other silica potentials, we calculated the ground-state energies of the clusters using BKS, FB, and TTAM potentials. For BKS and TTAM, the columnar structures were always more stable than the disk-like structures for $14 \leq k \leq 27$, while FB exhibited the correct structural transition. The odd-even alternation in the columnar structure energies was reversed for all three potentials as shown in Fig. 3.27d.

Next, we looked at some of the interesting meta-stable nanostructures that were previously examined by Bromley and co-workers. Nanochains, which were used as our training set structures, consist of planar two-membered silica rings (Si_2O_2) that are roughly perpendicular to their neighboring rings, with non-bonding terminal oxygens (NBO) at the end of chains, while nanorings (Fig. 3.26c) are fully coordinated structures that can be constructed by joining the NBO ends of the nanochain. Bromley et al. [129] using DFT with B3LYP and the 6-31+G* basis set examined the nanochain and nanorings structures as a function of cluster size ($k < 15$) and noticed that (i) the nanochains were relatively more stable for $k < 12$, (ii) nanorings structures were slightly more stable for $k \geq 12$. EDD-Q yields a similar trend, as shown in Fig. 3.27e, with the crossover occurring at $k = 12$, though odd-membered ring structures seem to be less stable than either their even-membered counterparts or odd-membered chain structures.

Nanotubes are quasi one-dimensional structures that have generated a lot of interest due to their structural similarity to carbon nanotubes [130]. These consisted of a repeated six-membered silica ring structure, with bridging oxygens connecting neighboring rings. The ends are oxygen terminated to ensure the proper coordination of atoms, with the two sets of terminal oxygens anti-aligned to obtain lower energy tubular structures (Fig. 3.26d). Though higher in energy than the ground-state counterparts, given the relative ease with which these structures could be constructed, we examined the minimized energies of these structures as a function of k for $k \leq 216$. Figure 3.27f shows the variation in energy as a function of k. We repeated the fitting procedure previously applied to the ground-state structures (Fig. 3.27c) and found that the fitted E_{res} was comparable to the corresponding value for the small clusters ($k \leq 12$). In addition, we also calculated the ground-state energies of other meta-stable nano-cages for $k = 12$, 18, 24 as discussed by Bromley [131], and given in Fig. 3.27d as an inset. Specifically, we examined the truncated tetrahedron and the hexagonal prism structures for $k = 12$, the fully coordinated polyhedron for $k = 18$, and the fullerene and truncated cube structures for $k = 24$. The calculated energies of the nanotubes, and nano-cages for $k = 12$, and 18 compare well with the reported DFT energies and are within a few hundredths of an eV/SiO_2 of the DFT energies, while for $k = 24$, though the calculated energy of the nanotube is similar to DFT, it is lower than that for the fullerene contrary to the DFT results of Bromley [131].

In summary, the EDD-Q potential successfully matches DFT results in the prediction of energetics and structures of many silica nanoclusters and performs significantly better than some select silica potentials (BKS, FB, and TTAM) in its ability to predict trends observed by DFT results. When viewed in conjunction with the work on water, it is clear that the charge fluctuation and polarization effects are crucial to the success of interatomic potentials, and the EDD-Q parameterization philosophy ensures accuracy and the ability to naturally capture environment dependent fluctuation in atomic charges.

Appendix

The functional forms for the environment dependent charges as well as the EDD-Q potential form for water and silica are given below:

Silica

The charge on a silica atom is obtained as follows:

$$q^{\mathrm{Si}} = q_{res}^{\mathrm{Si}} + \sum_{B=1}^{N_O} \varepsilon_B \, dq_B^{\mathrm{SiO}}, \text{ where } \varepsilon_B = \left\{ \begin{array}{c} 1, R_{\mathrm{SiO}} \leq r_c \\ 0, \text{ otherwise} \end{array} \right\} \tag{3.165}$$

where the bond charge is defined by

$$dq_B^{\mathrm{SiO}} = \Delta q \left[1 - \tan h_3 \left(\frac{R_{AB} - r_{c1}}{r_0} \right) \right] + \sum_{C \neq B}^{L_O} \varepsilon_C dq_{BC}^{\mathrm{OSiO}} \tag{3.166}$$

and

$$\begin{aligned} dq_{BC}^{\mathrm{OSiO}} = & \left[b_1 \sin^2 \theta_{BAC} \exp(-\theta_{BAC}) + b_2 \right] \tan h_4 \left(\frac{R_{AB} - r_{c2}}{r_0} \right) \\ & + \left[p_1 \exp\left(-\theta_{BAC}^2 \sin^2 \theta_{BAC} \right) \right] \tan h_2 \left(\frac{R_{AB} - r_{c3}}{r_0} \right) \end{aligned} \tag{3.167}$$

In the above equations, r_0 equals unity and has dimensions of Å, R_{AB} and R_{AC} are the distances of the Si atom A from oxygens B and C respectively, and θ_{BAC} is the angle formed by those three atoms; the various parameters are given in Table 3.9.

Table 3.9 Charge parameters for the EDD-Q potential

$q_{\mathrm{Si}}^{res}(e)$	$q_{\mathrm{o}}^{res}(e)$	$\Delta q(e)$	$b_1(e)$	$b_2(e)$	$P_1(e)$	$r_{c1}(\mathring{A})$	$r_{c2}(\mathring{A})$	$r_{c3}(\mathring{A})$
2.34	$\frac{-q_{\mathrm{si}}^{res}}{2}$	0.333	0.3	−0.01	0.2	1.46	1.65	1.3

Similarly, the net charge on an oxygen atom having L_{Si} silicon neighbors is given by

$$q^{\mathrm{O}} = q^{\mathrm{O}}_{res} - \sum_{A=1}^{L_{\mathrm{Si}}} \varepsilon_A dq_A^{\mathrm{SiO}}, \tag{3.168}$$

where we have used the definition of bond charge from (3.166). As is obvious from (3.165) and (3.168), a net residual charge q_{res} is associated with isolated silicon or oxygen atoms.

Equations (3.169) and (3.170) was used as the basis for our potential function with the functional form and the parameters of ϕ_{AB} and $\mathcal{A}_A$ chosen to yield the ground state (as well as the deformed) geometry and energetics of the model clusters as predicted by DFT. The effective charges on atoms were scaled by a constant multiplicative factor while evaluating the Coulombic term in the energy expression. In other words, we use a screened value for the charge rather than the Mulliken populations, such that the ratio of the effective atomic charge and the Mulliken population is a constant.

$$\bar{E}_A = F(q_A) + \frac{1}{2}\sum_{B\neq A}^{N} \phi_{AB} + \frac{1}{2}\sum_{B\neq A}^{N} \frac{q_A q_B}{R_{AB}} \tag{3.169}$$

where R_{AB} was the distance of separation between atoms A and B and $F(q_A)$ was expressed as

$$F(q_A) = \mathcal{A}_A q_A \ln q_A^2 \tag{3.170}$$

The functional forms of ϕ_{AB}, $\mathcal{A}_A$, $F(q_A)$ and the total energy expression are given below in (3.171)–(3.176), and Table 3.10 contains the values of the various parameters. ϕ is given by

$$\phi_{\mathrm{SiSi}} = \frac{a_{\mathrm{SiSi}}}{r^{12}_{\mathrm{SiSi}}} \tag{3.171}$$

$$\phi_{\mathrm{OO}} = \frac{a_{\mathrm{OO}}}{r^{12}_{\mathrm{OO}}} \tag{3.172}$$

$$\phi_{\mathrm{SiO}} = \frac{a_{\mathrm{SiO}}}{r^{20}_{\mathrm{SiO}}} - \frac{c_{\mathrm{SiO}}}{r^{6}_{\mathrm{SiO}}} \tag{3.173}$$

and

$$F(q_A) = \left(q^A_{res} - q^A\right)\left[\mathcal{A}^1_A \ln\left(\frac{q^A_{res} - q_A}{d_A}\right)^2\right] \tag{3.174}$$

Table 3.10 Potential parameters for the EDD-Q potential

a_{SiSi} ($eV/Å^{12}$)	a_{OO} ($eV/Å^{12}$)	a_{SiO} ($eV/Å^{20}$)	c_{SiO} ($eV/Å^{6}$)	A^C_{Si} (eV/e)	A^C_O (eV/e)	η	d_{Si} (e)	d_O (e)	f_{Si} (eV/e)	r_{ae} (Å)	s_q
2.133×10^4	2.05×10^4	275.0	12.75	0.465	−0.1642	5.25	2.8		1.63	1.86	$\sqrt{q^{Si}_{res}}$

where for silicon atom A,

$$\mathcal{A}^1_A = \mathcal{A}^C_{Si}\left\{1 + \sum_{B=1}^{L_O} \varepsilon_B f^{Si}\left[1 - \tan h_\eta\left(\frac{r_{AB} - r_{ae}}{r_0}\right)\right]\right\} \tag{3.175}$$

and for oxygen atom B,

$$\mathcal{A}^1_B = \mathcal{A}^C_O. \tag{3.176}$$

$\mathcal{A}^0_{Si}$, $\mathcal{A}^0_O$, $\mathcal{A}^C_{Si}$, and $\mathcal{A}^C_{Si}$ are parameters given in Table 3.10. The potential energy of atom A is

$$\bar{E}_A = F(q_A) + \frac{1}{2}\sum_{B\neq A}^{N} \phi_{AB} + \frac{1}{2}\sum_{B\neq A}^{N} \frac{q^{eff}_A q^{eff}_B}{R_{AB}} \tag{3.177}$$

where q_A is the charge as obtained from (A3.1) to (A3.4) and q^{eff}_A is given by

$$q^{eff}_A = q_A/s_q \tag{3.178}$$

with s_q given in Table 3.10.

Water

Water polymorphs are characterized by the formation of hydrogen bonds between neighboring water molecules. In this regard, in the EDD-Q formulation, the net atomic charge on oxygen and hydrogen atoms constituting a water monomer depends both on the two underlying OH bond distance(s) and HOH bond angle as well as the hydrogen bond distance between neighboring water monomers.

First, we consider the functional form for calculating atomic charges within a monomer (H_1OH_2) that consists of two OH bonds (r_{OH1} and r_{OH2}) and a bond angle (θ) expressed in radians.

The charge (expressed in e) on the hydrogen atom (q_{H1}) is given by

$$\begin{aligned} q_{H1} = {} & [\alpha(\theta)\exp(-2r_{OH1}) + \beta(\theta) r_{OH1}\exp(-r_{OH1}) + c(\theta) r_{OH1}]\sin^2(\theta) \\ & + (r_{OH1} - r_{OH2}) d(\theta)\sin^2(\theta), \end{aligned} \tag{3.179}$$

where α, β, c, d are functions of θ (given below in (3.180)). The charge on the other hydrogen atom within a water monomer can also be obtained in a similar fashion.

$$\begin{aligned}\alpha(\theta) &= \alpha_1 \exp(\theta) + \alpha_2 \exp(\theta^2/4) + \alpha_3\theta \\ \beta(\theta) &= \beta_1\theta \exp(-\theta) + \beta_2\theta \exp(-2\theta) + \beta_3\theta^2 \\ c(\theta) &= c_1\theta \exp(-\theta) + c_2\theta \exp(-2\theta) + c_3 \\ d(\theta) &= d_1(\theta - \theta^2) + d_2\theta^3\end{aligned} \tag{3.180}$$

To prevent energy discontinuities at the neighbor cutoff distance (=1.5 Å), a switching function *S(t)* is used for modulating the calculated hydrogen charge as given below:

$$\begin{aligned}q_{\mathrm{H1}} &= q_{\mathrm{H1}} S(r_{\mathrm{OH1}} - t_{cut}) \\ S(t) &= 0.5(1 - \tan h(t/t_1))\end{aligned} \tag{3.181}$$

The charge on the oxygen atom in the water monomer is given by

$$q_{\mathrm{O}} = -(q_{\mathrm{H1}} + q_{\mathrm{H2}}). \tag{3.182}$$

The corresponding parameters for the water monomer charges are given in Table 3.11.

To account for charge transfer between neighboring monomers, a net intermolecular charge transfer (*dq*) between two monomers is obtained as follows:

$$dq = a \exp(-br_{\mathrm{HB}}), \tag{3.183}$$

where r_{HB} is the hydrogen bond distance between donor hydrogen atom and acceptor oxygen atoms belong to neighboring monomers respectively. The neighbor cutoff distance s_{cut} between two monomers is set to 2.5 Å and thus dq is modulated by the switching function *S* as defined before.

$$dq = dqS(r_{\mathrm{HB}} - r_{im}). \tag{3.184}$$

Further, *dq* is partitioned among the respective monomer atoms as follows:

$$\begin{aligned}dq^{\mathrm{O}_{donor}} &= -0.5dq \\ dq^{\mathrm{O}_{acceptor}} &= 0.75dq \\ dq^{\mathrm{H}_{donor}} &= -0.4dq\end{aligned} \tag{3.185}$$

Table 3.11 Charge parameters for the water monomer

α_1	α_2	α_3	β_1	β_2	β_3	c_1	c_2	c_3	d_1	d_2	t_1	t_{cut}
0.816	−3.22	0.573	−14.01	61.82	1.319	12.235	−15.78	−3.593	0.443	0.115	0.03 (Å)	1.41 (Å)

Table 3.12 Water dimer charge parameters

a	b	r_{im}	s_{cut}
1.581	1.822	2.35 (Å)	2.5 (Å)

In addition, for the acceptor molecule, $dq^{\mathrm{H}_{acceptor}}$ is obtained by partitioning $[dq - dq^{\mathrm{O}_{acceptor}}]$ equally among the two hydrogen atoms, while within the donor molecule the other hydrogen atom acquires an additional charge equaling $[dq^{\mathrm{O}_{donor}} + dq^{\mathrm{H}_{donor}} - dq]$.

Thus the total charge for any atom is given by a sum of its 'monomer' charge plus additional charge transfer that is acquired due to hydrogen bonding with neighboring monomers. Note that an equivalent dq arises for every hydrogen bonded interaction a monomer participates in Table 3.12.

Similar to the Hamiltonian defined for silica, the energy of an atom i is given by

$$\bar{E}_A = F(q_A) + \frac{1}{2}\sum_{B\neq A}\Phi_{AB} + \frac{1}{2}\sum_{B\neq A}\frac{q_A q_B}{R_{AB}} \tag{3.186}$$

Here,

$$F(q_A) = \mathcal{A}_A^m q_A \ln(q_A^2) + \mathcal{A}_A^d dq_A \ln(dq_A^2) \tag{3.187}$$

For an oxygen atom, $\mathcal{A}_A^m$ is defined in terms of the monomer bond angle and bond distances as follows:

$$\mathcal{A}_{\mathrm{O}}^m = -2A_{EO}\sin^2(\theta)\exp\left(-\frac{r_0^2}{2}(r_{\mathrm{OH1}} - r_{\mathrm{OH2}})^2\sin^2(\theta)\right), \tag{3.188}$$

while for a hydrogen atom, $\mathcal{A}_A^m$ depends on the number of hydrogen bonds (L_{OH}) that the hydrogen atom (H) is involved with. Further, the index 's' in (3.189) refers to the oxygen atom in the monomer to which H is associated with. $\theta_{u\mathrm{H}s}$ refers to the angle between the $\mathrm{O_s}$–H bond (within the monomer) and the respective hydrogen bonds that H forms with other $\mathrm{O_u}$ atoms that belong to neighboring monomers.

$$\mathcal{A}_{\mathrm{H}}^m = \mathcal{A}_{E\mathrm{H}}\left(1 + \eta\sum_{u=1}^{L_{\mathrm{OH}}}\right)\left[\left(\exp\left(-2(1+\cos(\theta_{u\mathrm{H}s})\right)^2\right)\right)\left(1 - \tan h\left(\frac{r_{u\mathrm{H}} - r_D}{t_2}\right)\right)\right]. \tag{3.189}$$

Table 3.13 Potential parameters (part I)

a_{OO}	a_{OH}	a_{HH}	b_{OH}	c_{OH}	A_{EO}	$A_O{}^d$	A_{EH}	A_H^d	η
25.0	3.011	25	−2.405	2.5×10^{-6}	−11.429	0.0	4.762	−0.5	0.051

Table 3.14 Potential parameters (part II)

H_{cut}	r_D	t_2	r_{cut}
2.43	2.1	0.1	2.5

Finally, Φ_{ij} is given below:

$$\begin{aligned}
\Phi_{OO} &= a_{OO}\exp(-4r_{OO}) \\
\Phi_{OH} &= 2\left[a_{OH}r_{OH} + b_{OH}\exp(-r_{OH}) + \frac{C_{OH}}{r_{OH}^{24}}\right]S(r_{OH} - r_{cut}) \\
\Phi_{HH} &= 2a_{HH}\exp(-2r_{HH})S(r_{HH} - H_{cut}).
\end{aligned} \tag{3.190}$$

Tables 3.13 and 3.14 provides the corresponding potential parameters.

For the two systems, the charge and the other potential parameters, while parameterized with respect to ab initio data, were selected in an ad hoc fashion. In order to provide a more streamlined approach, a more systematic approach leading to a more intuitive functional form is proposed as noted below. Future development of EDD-Q potentials will be based on these functional forms (3.191–3.199).

As discussed before, for a bond formed between atoms A and B, the bond-charge (q_{AB}^{bond}) is a function of the number of nearest neighbors of $A(L_A)$ and $B(L_B)$, the interatomic distance (R_{AB}), and the bond-angles that arise due to the remaining nearest neighbors of A and B (θ_{BAC} and $\theta_{ABC'}$), where C and C' represent the neighbors of A and B respectively. Note the qualitative similarity of the functional forms for charges (3.191–3.198) to the multipole expansion used commonly to express the electrostatic potentials that arise from a distribution of charges.

$$q_A = \sum_{B=1}^{L_A} q_{AB}^{bond} \tag{3.191}$$

$$q_{AB}^{bond} = [q_{AB} + q_{BA}] \tag{3.192}$$

$$q_{AB} = \frac{1}{2}\left[\frac{1}{L_A}q_{AB}^{S}(R_{AB}) + \sum_{C\neq B}^{L_A}\Delta q(R_{AB},\theta_{BAC})\right] \tag{3.193}$$

$$\text{if } L_A = 1,\ q_{AB} = q_{BA}$$

$$q_{AB}^{S}(R_{AB}) = \sum_{p} \frac{\mathcal{A}_p}{R_{AB}^{p}}; p > 2 \tag{3.194}$$

$$\Delta q(R_{AB}, \theta_{BAC}) = f^{A}(\theta_{BAC}, R_{AB}) + g^{A}(R_{AB}) \tag{3.195}$$

$$f^{A}(\theta, r) = w^{A}(\theta) y(r) \tag{3.196}$$

$$w^{A}(\theta) = \left[u_A^f + \sum_{p} C_p^A \cos^p(p\theta); p > 0 \right] \tag{3.197}$$

$$y(R) = \left[u_A^e + \sum_{p} \frac{B_p}{R_{AB}^{p}}; p > 2 \right]$$

$$g^{A}(r) = u_A^g + \sum_{p} \frac{D_p^A}{R_{AB}^{p}}; p > 2 \tag{3.198}$$

The EDD-Q Hamiltonian which consists of a Coulombic term, a 2-body term (ϕ) and an embedding term ($F(q)$) are given below. $F(q)$ consists of three contributions; the most important of the three terms is the second term, which can be correlated to the 'energy-cost' for embedding a bond that is formed between pairs of atoms.

$$\begin{aligned}
\bar{E} &= \sum_{A} \bar{E}_A \\
\bar{E}_A &= \frac{1}{2} \sum_{B \neq A} \frac{q_A q_B}{R_{AB}} + F(q_A) + \frac{1}{2} \sum_{B \neq A} \phi_{AB}(R_{AB}) \\
F(q_A) &= M_1 q_A^2 \ln(q_A^2) + M_2 \sum_{B \neq A} q_{AB} \ln(q_{AB}^2) + M_3 \sum_{B \neq A} h_{AB} q_{AB}^2 \ln(q_{AB}^2) \\
h_{AB} &= \sum_{C \neq B}^{L_A} \left[1 - \left(h^{(1)} \cos(\theta_{BAC}) + h^{(2)} \cos(2\theta_{BAC}) + h^{(3)} \cos(3\theta_{BAC}) \right) \right] \\
\phi_{AB} &= \frac{T_{AB}}{R_{AB}^{12}} - \frac{S_{AB}}{R_{AB}^{6}}
\end{aligned} \tag{3.199}$$

References

1. J.A. Aguiar, P.P. Dholabhai, Z. Bi, Q. Jia, E.G. Fu, Y.Q. Wang, T. Aoki, J. Zhu, A. Misra, B. P. Uberuaga, Adv. Mater. Interf. **1**, 201300142 (2014)
2. G. Hunter, Int. J. Quant. Chem. **9**, 237 (1975)
3. E. Deumens, A. Diz, H. Taylor, Y. Ohrn, J. Chem. Phys. **96**, 6820 (1992)
4. D.A. Micha, K. Runge, Phys. Rev. A **50**, 322 (1994)
5. A. Abedi, N.T. Maitra, E.K.U. Gross, Phys. Rev. Lett. **105**, 123002 (2010)

6. L.S. Cederbaum, J. Chem. Phys. **138**, 224110 (2013)
7. F.L. Hirshfeld, Theor. Chim. Acta **44**, 129 (1977)
8. R.F.W. Bader, J. Chem. Phys. **73**, 2871 (1980)
9. R.F.W. Bader, Atoms in molecules, in *Encyclopedia of Computational Chemistry* (Wiley, 2002)
10. A. Cedillo, P.K. Chattaraj, R.G. Parr, Intern. J. Quant. Chem. **77**, 403 (2000)
11. R.F. Nalewajski, R.G. Parr, J. Phys. Chem. A **105**, 7391 (2001)
12. R.F. Nalewajski, J. Phys. Chem. A **107**, 3792 (2003)
13. W. Moffitt, Proc. Royal Soc. London. Series A. Math. Phys. Sci. **210**, 245 (1951)
14. F.O. Ellison, J. Am. Chem. Soc. **85**, 3540 (1963)
15. J.C. Tully, J. Chem. Phys. **58**, 1396 (1973)
16. J.C. Tully, C.M. Truesdale, J. Chem. Phys. **65**, 1002 (1976)
17. J.A. Olson, B.J. Garrison, J. Chem. Phys. **83**, 1392 (1985)
18. S.M. Valone, J. Chem. Theor. Comput. **7**, 2253 (2011)
19. J. Rychlewski, R.G. Parr, J. Chem. Phys. **84**, 1696 (1986)
20. R.S. Mulliken, J. Chem. Phys. **2**, 782 (1934)
21. R. Jastrow, Phys. Rev. **98**, 1479 (1955)
22. W. Cencek, W. Kutzelnigg, J. Chem. Phys. **105**, 5878 (1996)
23. R. Colle, O. Salvetti, Theor. Chim. Acta **53**, 55 (1979)
24. R. Colle, O. Salvetti, J. Chem. Phys. **79**, 1404 (1983)
25. C. Lee, W. Yang, R.G. Parr, Phys. Rev. B **37**, 785 (1988)
26. R.G. Parr, R.G. Pearson, J. Am. Chem. Soc. **105**, 7512 (1983)
27. E. Cancès, C. Le Bris, P.-L. Lions, Nonlinearity **21**, T165 (2008)
28. E.P. Gyftopoulos, G.N. Hatsopoulos, Proc. Natl. Acad. Sci. **60**, 786 (1965)
29. R.P. Feynman, in *Statistical Mechanics: A Set of Lectures; Notes Taken by R. Kikuchi and H. A. Feiveson*, ed. by J. Shaham (W.A. Benjamin, Reading, 1972), Book 2, pp. 70–80
30. J.H. van Lenthe, G.G. Balint-Kurti, J. Chem. Phys. **78**, 5699 (1983)
31. D. Lauvergnat, P.C. Hiberty, D. Danovich, S. Shaik, J. Phys. Chem. **100**, 5715 (1996)
32. G. Gallup, *Valence Bond Methods: Theory and Applications* (Cambridge University Press, 2002)
33. A. Shurki, H.A. Crown, J. Phys. Chem. B **109**, 23638 (2005)
34. L. Song, Y. Mo, Q. Zhang, W. Wu, J. Comput. Chem. **26**, 514 (2005)
35. A. Shurki, Theor. Chem. Acc.: Theor. Comput. Model. (Theoretica Chimica Acta) **116**, 253 (2006)
36. A. Sharir-Ivry, H.A. Crown, W. Wu, A. Shurki, J. Phys. Chem. A **112**, 2489 (2008)
37. H.O. Pritchard, H.A. Skinner, Chem. Rev. **55**, 745 (1955)
38. J.P. Perdew, R.G. Parr, M. Levy, J.L. Balduz, Phys. Rev. Lett. **49**, 1691 (1982)
39. S.M. Valone, S.R. Atlas, J. Chem. Phys. **120**, 7262 (2004)
40. S.M. Valone, S.R. Atlas, Phys. Rev. Lett. **97**, 256402 (2006)
41. S.W. Rick, S.J. Stuart, B.J. Berne, J. Chem. Phys. **101**, 6141 (1994)
42. X.W. Zhou, H.N.G. Wadley, J.-S. Filhol, M.N. Neurock, Phys. Rev. B **69**, 035402 (2004)
43. J. Chen, T.J. Martinez, Chem. Phys. Lett. **438**, 315 (2007)
44. J. Morales, T.J. Martìnez, J. Phys. Chem. A **105**, 2842 (2001)
45. J. Morales, T.J. Martìnez, J. Phys. Chem. A **108**, 3076 (2004)
46. W.J. Mortier, S.K. Ghosh, S. Shankar, J. Am. Chem. Soc. **108**, 4315 (1986)
47. A.K. Rappé, W.A. Goddard, J. Phys. Chem. **95**, 3358 (1991)
48. A.C.T. van Duin, S. Dasgupta, F. Lorant, W.A. Goddard, J. Phys. Chem. A **105**, 9396 (2001)
49. J. Yu, S.B. Sinnott, S.R. Phillpot, Phys. Rev. B **75**, 085311 (2007)
50. R.A. Nistor, J.G. Polihronov, M.H. Müser, N.J. Mosey, J. Chem. Phys. **125**, 094108 (2006)
51. R.A. Nistor, M.H. Müser, Phys. Rev. B **79**, 104303 (2009)
52. D. Mathieu, J. Chem. Phys. **127**, 224103 (2007)
53. T. Verstraelen, V.V. Speybroeck, M. Waroquier, J. Chem. Phys. **131**, 044127 (2009)
54. P.T. Mikulski, M.T. Knippenberg, J.A. Harrison, J. Chem. Phys. **131**, 241105 (2009)
55. R.P. Iczkowski, J.L. Margrave, J. Am. Chem. Soc. **83**, 3547 (1961)

56. S.M. Valone, J. Phys. Chem. Lett. **2**, 2618 (2011)
57. J.C. Phillips, Rev. Mod. Phys. **42**, 317 (1970)
58. R.J. Spindler Jr, J. Quant. Spectry. Radiative Trans. **9**, 597 (1969)
59. R.G. Pearson, J. Am. Chem. Soc. **85**, 3533 (1963)
60. P. Ghosez, J.-P. Michenaud, X. Gonze, Phys. Rev. B **58**, 6224 (1998)
61. W. Harrison, *Elementary Electronic Structure of Materials* (W. A. Benjamin, Reading, 1996)
62. M. Cohen, A. Wasserman, J. Stat. Phys. **125**, 1121 (2006)
63. J. Cioslowski, B.B. Stefanov, J. Chem. Phys. **99**, 5151 (1993)
64. S.M. Valone, S.R. Atlas, M.I. Baskes, Model. Simul. Mater. Sci. Eng. **22**, 045013 (2014)
65. J.E. Hirsch, Phys. Rev. B **48**, 3327 (1993)
66. M.S. Daw, M.I. Baskes, Phys. Rev. Lett. **50**, 1285 (1983)
67. M.I. Baskes, Phys. Rev. B **46**, 2727 (1992)
68. A.M. Dongare, M. Neurock, L.V. Zhigilei, Phys. Rev. B **80**, 184106 (2009)
69. Y. Mishin, M.J. Mehl, D.A. Papaconstantopoulos, A.F. Voter, J.D. Kress, Phys. Rev. B **63**, 224106 (2001)
70. M.W. Finnis, J.E. Sinclair, Philos. Mag. A **50**, 45 (1984)
71. A.P. Sutton, *Electronic Structure of Materials* (Oxford Science Publications, Clarendon Press, 1993), Chap. 12
72. J.P. Perdew, J.A. Chevary, S.H. Vosko, K.A. Jackson, M.R. Pederson, D.J. Singh, C. Fiolhais, Phys. Rev. B **46**, 6671 (1992). References 26 and 27 therein
73. G. Kresse, J. Furthmüller, Phys. Rev. B **54**, 11169 (1996)
74. G. Kresse, D. Joubert, Phys. Rev. B **59**, 1758 (1999)
75. R.T. Sanderson, Science **114**, 670 (1951)
76. R.G. Parr, R.A. Donnelly, M. Levy, W.E. Palke, J. Chem. Phys. **68**, 3801 (1978)
77. F.H. Streitz, J.W. Mintmire, Phys. Rev. B **50**, 11996 (1994)
78. F.H. Streitz, J.W. Mintmire, Langmuir **12**, 4605 (1996)
79. A. Alavi, L.J. Alvarez, S.R. Elliott, I.R. McDonald, Philos. Mag. B **65**, 489 (1992)
80. L. Huang, J. Kieffer, J. Chem. Phys. **118**, 1487 (2003)
81. B.W.H. van Beest, G.J. Kramer, R.A. van Santen, Phys. Rev. Lett. **64**, 1955 (1990)
82. M.S. Gordon, M.A. Freitag, P. Bandyopadhyay, J.H. Jensen, V. Kairys, W.J. Stevens, J. Phys. Chem. A **105**, 293 (2001)
83. J. Cioslowski, Phys. Rev. Lett. **62**, 1469 (1989)
84. X.W. Zhou, F.P. Doty, Phys. Rev. B **78**, 224307 (2008)
85. G. Koster, T.H. Geballe, B. Moyzhes, Phys. Rev. B **66**, 085109 (2002)
86. H.S. Smalø, P.-O. Åstrand, L. Jensen, J. Chem. Phys. **131**, 044101 (2009)
87. M.H. Cohen, A. Wasserman, Isr. J. Chem. **43**, 219 (2003)
88. P. Elliott, M.H. Cohen, A. Wasserman, K. Burke, J. Chem. Theor. Comput. **5**, 827 (2009)
89. G. Wu, G. Lu, C.J. García-Cervera, W.E, Phys. Rev. B **79**, 035124 (2009)
90. A. Warshel, *Computer Modeling of Chemical Reactions in Enzymes and Solutions* (John Wiley and Sons Inc, New York, 1991)
91. M.J. Field, P.A. Bash, M. Karplus, J. Comput. Chem. **11**, 700 (1990)
92. F. Bloch, Zeits. Phys. A **52**, 555 (1929)
93. S.M. Valone, S.R. Atlas, Philos. Mag. **86**, 2683 (2006)
94. S.M. Valone, J. Li, S. Jindal, Int. J. Quant. Chem. **108**, 1452 (2008)
95. R. Drautz, D.G. Pettifor, Phys. Rev. B **74**, 174117 (2006)
96. M.W. Finnis, *Interatomic Forces in Condensed Matter* (Oxford University Press, Oxford, 2003)
97. P. Sutton, M.W. Finnis, D.G. Pettifor, Y. Ohta, J. Phys. C: Solid State Phys. **21**, 35 (1988)
98. J.C. Slater, G.F. Koster, Phys. Rev. **94**, 1498 (1954)
99. G.J. Ackland, M.W. Finnis, V. Vitek, J. Phys. F: Metal Phys. **18**, L153 (1988)
100. M.I. Baskes, S.G. Srinivasan, S.M. Valone, R.G. Hoagland, Phys. Rev. B **75**, 94113 (2007)
101. J. Hubbard, Proc. Royal Soc. London. Series A. Math. Phys. Sci. **276**, 238 (1963)
102. A. Elsener, O. Politano, P.M. Derlet, H.V. Swygenhoven, Model. Simul. Mater. Eng. **16**, 025006 (2008)

103. H.J.C. Berendsen, J.P.M. Postma, W.F. van Gunsteren, Hermans, J, in *Intermolecular Forces*, ed. by B. Pullman (Reidel, Dordrecht, 1981), p. 331
104. C.D. Berweger, W.F. van Gunsteren, F. Muller-Plathe, Chem. Phys. Lett. **232**, 429 (1995)
105. W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey, M.L.J. Klein, Chem. Phys. **79**, 926 (1983)
106. E. Neria, S. Fischer, M.J. Karplus, Chem. Phys. **105**, 1902 (1996)
107. M.W. Mahoney, W.L. Jorgensen, J. Chem. Phys. **112**, 8910 (2000)
108. H. Saint-Martin, J. Hernandez-Cobos, M.I. Bernal-Uruchurty, I. Ortega-Blake, H.J.C. Berendsen, J. Chem. Phys. **113**, 10899 (2000)
109. G. Corongiu, E. Clementi, J Chem Phys. **97**, 2030 (1992)
110. O. Matsuoka, E. Clementi, M. Yoshimine, J. Chem. Phys. **64**, 1351 (1976)
111. P.-O. Åstrand, A. Wallqvist, G. Karlström, J. Chem. Phys. 100 (1994)
112. H.A. Stern, B.J. Berne, R.A. Friesner, J. Chem. Phys. **115**, 2237 (2001)
113. J.W. Halley, J.R. Rustad, A. Rahman, J. Chem. Phys. **98**, 4110 (1993)
114. R.L. Corrales, J. Chem. Phys. **110**, 9071 (1999)
115. U.W. Schmitt, G.A. Voth, J. Chem. Phys. **111**, 9361 (1999)
116. T.J.F. Day, A.V. Soudackov, M. Cuma, U.W. Schmitt, G.A. Voth, J. Chem. Phys. **117**, 5839 (2002)
117. S.-B. Zhu, G.W. Robinson, Proc. Int. Conf. Supercomp. **II**, 189 (1989)
118. S.-B. Zhu, S. Singh, G.W. Robinson, J. Chem. Phys. **95**, 2791 (1991)
119. L.X. Dang, T. Cheng, J. Chem. Phys. **106**, 8149 (1997)
120. C. Millot, J.-C. Soetens, M.T.C. Martins Costa, M.P. Hodges, A.J. Stone, J. Phys. Chem. A **102**, 754 (1998)
121. C.J. Burnham, J.C. Li, S.S. Xantheas, M.S. Leslie, J. Chem. Phys. **110**, 4566 (1999)
122. G.C. Groenenboom, E.M. Mas, R. Bukowski, K. Szalewicz, P.E.S. Wormer, A.V. Avoird, Phys. Rev. Lett. **84**, 4072 (2000)
123. R.F. Byrd, P. Lu, J. Nocedal, C. Zhu, **16**, 1190 (1995)
124. S. Maheshwary, N. Patel, N. Sathyamurthy, A.D. Kulkarni, S.R. Gadre, J. Phys. Chem. A **105**, 10525 (2001)
125. S.T. Bromley, E. Flikkema, Phys. Rev. Lett. **95**, 185505 (2005)
126. S. Tsuneyuki, M. Tsukada, H. Aoki, Y. Matsui, Phys. Rev. Lett. **61**, 869 (1988)
127. E. Flikkema, S.T. Bromley, Chem. Phys. Lett. **378**, 622 (2003)
128. E. Flikkema, S.T. Bromley, J. Phys. Chem. B **108**, 9638 (2004)
129. S.T. Bromley, M.A. Zwijnenburg, Th Maschmeyer, Phys. Rev. Lett. **90**, 35502 (2003)
130. N.H. de Leeuw, Z. Du, J. Li, S. Yip, T. Zhu, Nano Lett. **3**, 1347 (2003)
131. S.T. Bromley, Nano Lett. **4**, 1427 (2004)

Chapter 4
Phase Field Methods

Nan Wang and Long-Qing Chen

Abstract The phase-field method to modeling microstructure evolution in materials is introduced in this chapter. It can be employed to model, understand, and predict complex material behaviors at meso-scale (on the order of nano- to micro-meter size), and it covers application-relevant time scales (on the order of seconds to years). Coarse-grained from the underlying atomic-level physics, it is an established approach to fill the length-scale gap between atomic/electronic scale calculations and the macroscopic continuum method, ignoring the underlying material's inhomogeneity. Employing energy functionals and nonlinear partial differential equations, the general formulation of phase-field models is outlined. Several applications of the phase-field model are presented.

4.1 Introduction

The phase-field (PF) method has been applied to various problems in materials science and has received significant attention in the last 20 years [1–7]. It has established itself as one of the classical approaches in multiscale materials modeling in parallel to Density Functional Theory (DFT), Molecular Dynamics (MD), and Monte-Carlo (MC) methods. While the majority of applications focused on understanding the formation and evolution of meso-scale complex microstructures, the phase-field method is very useful in connecting atomic/electronic scale models to large-scale continuum models.

In a phase-field model, phase-field variables are employed to distinguish different spatial domains of the system, and they are continuous across a boundary separating different domains. By tracking the evolution of the phase-field variables,

N. Wang (✉) · L.-Q. Chen
Department of Materials Science and Engineering, The Pennsylvania State University, Pennsylvania, USA
e-mail: nxw13@psu.edu

L.-Q. Chen
e-mail: lqc3@psu.edu

P.A. Deymier et al. (eds.), *Multiscale Paradigms in Integrated Computational Materials Science and Engineering*, Springer Series in Materials Science 226,
DOI 10.1007/978-3-319-24529-4_4

complex interface motions can be captured. There are mainly two categories of PF models. The first one uses a set of physical order parameters as the phase-fields and evolves them by minimizing a free energy functional constructed from the thermodynamics that govern the transition between different phases. The other model uses phase-field as a pure mathematical tool to simplify the otherwise impossible task of tracking complex interface evolution, and asymptotically reduces the model to a sharp-interface theory at the level of differential equations. Recently, a new class of phase-field related models has emerged, namely, the phase-field-crystal (PFC) model and the amplitude equation (AE) approach. Based on a partial differential equation (PDE) that produces spatially periodic peaks, the PFC model reinterprets these peaks as time-averaged atoms and links the interaction of these peaks to the atom pair correlation function. Some interesting diffusion time-scale physics have been captured with atomic-level details using this model. The AE approach is rather similar to the conventional PF model in terms of the governing equations. The idea is to capture the change in the amplitude of different atomic density waves as a system undergoes a transition from one state to another. By deriving AEs from microscopic models (based on classical Density Functional Theory or the Phase-field-crystal model), this approach connects the underlying materials physics directly to a set of continuous PDEs without the need for justification using either the energy argument or deriving the sharp-interface limit.

In this section, we first introduce the common mathematical formulations of conventional PF models and also briefly cover the PFC and the AE approach. Several applications of the conventional PF model are then discussed in the following section. Finally, perspectives of the PF methods are discussed in the final concluding section.

4.2 Methods

4.2.1 Conventional PF Methods

A widely used approach of formulating a PF model starts from a free energy functional which describes the total energy of the system in terms of a volume integral

$$F = \int [K|\nabla\phi|^2 + Hf(\phi, T, c\ldots) + Hf_{others}]dv. \tag{4.1}$$

Here, ϕ is the phase-field variable, K and H are dimensional coefficients, f is a local potential function which couples the phase-field variable to other physical quantities such as temperature T or concentration c, and f_{others} is used to account for elastic, electric or other contributions to the total free energy. In one bulk phase (A phase) the PF variable is $\phi = \phi_A$, while in the other phase (B phase) the PF variable is $\phi = \phi_B$. Since either of these two phases must be independently stable, the

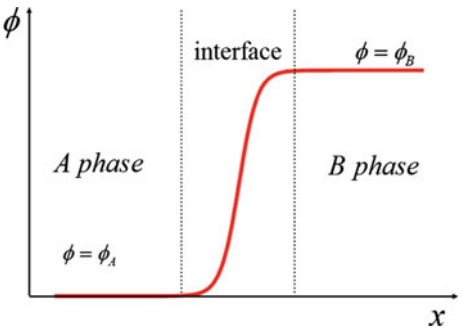

Fig. 4.1 Schematic plot of the PF variable across an A-B interface. x is the surface normal coordinate

potential function $f(\phi, \ldots)$ must have a minimum in each of the two bulk phase states, i.e. $f(\phi, \ldots)$ is minimized at $\phi = \phi_A$ or $\phi = \phi_B$. While this potential term favors a sharp transition from ϕ_A to ϕ_B, the gradient term $|\nabla\phi|^2$ by itself energetically favors a slow varying interface between A phase and B phase. As a result, the PF variable goes smoothly from ϕ_A to ϕ_B (or vice versa) at the phase boundary as shown in Fig. 4.1. The width of this diffused transition region is then decided by the relative magnitude of $|\nabla\phi|^2$ and f. We will see this point again at the differential equation level once the evolution equations of the PF variables are introduced.

To derive evolution equation of the PF variable from the free energy functional F, the variational principle is generally used to guarantee that F decreases monotonically in time. However, the detailed dynamics for energy minimization are not unique. Depending on the nature of the PF variable, there are two types of dynamics commonly used in PF models, conserved and non-conserved. The conserved dynamics model leads to an equation

$$\frac{\partial\phi}{\partial t} = \nabla \cdot M\nabla\frac{\delta F}{\delta\phi}, \tag{4.2}$$

which is better known as the Cahn-Hilliard equation [8], and the non-conserved equation is

$$\frac{\partial\phi}{\partial t} = -L\frac{\delta F}{\delta\phi}, \tag{4.3}$$

which is known as the Allen-Cahn equation [9]. The conserved equation is used for PF variables that can only be rearranged in the energy minimization process (like a density or solute concentration) since it is basically a generalized diffusion equation with flux $J = -M\nabla\delta F/\delta\phi$. The non-conserved equation is used when the PF variable is related to crystal orientation, polarization, or a mere math notion to distinguish two phases. Coefficients M and L, which are commonly referred as mobilities in the PF literature, connect the driving force $\delta F/\delta\phi$ to the time evolution of the PF variables. Using the non-conserved equation as an example, an

evolution equation for the PF variables derived from the free energy functional in (4.1) is

$$\frac{1}{H}\frac{\partial \phi}{\partial t} = L\left(w^2\nabla^2\phi - \frac{\partial f}{\partial \phi} - \frac{\partial f_{others}}{\partial \phi}\right). \tag{4.4}$$

Here, the factor $w = \sqrt{K/H}$ is the relative importance of the potential and the gradient term in the free energy functional, and therefore a measure of the interface width in the model.

The finite width of the PF interface also leads to another important feature in the general PF formulation, the surface energy. Since the gradient term and the potential term are competing at the interface region, the non-zero energy density localized at this region gives the surface energy in the PF model. One can calculate this surface energy γ by integrating the free energy density (both the potential part and the gradient part) across the interface.

To solve the evolution of the PF variables in (4.4), numerical methods are employed. The simplest approach is to discretize the Laplacian operator with a finite difference scheme and apply the forward Euler method for time integration. More sophisticated approaches for solving PF equations can be found in [10, 11]. By tracking the motion of the PF variables, complex interface evolution can be easily captured.

4.2.2 PFC and AE Method

The basic PFC model uses an energy functional [12]

$$F_{pfc} = \int \left\{\frac{\varphi}{2}\left[-\varepsilon + \left(q^2 + \nabla^2\right)^2\right]\varphi + \frac{\varphi^4}{4}\right\}dv, \tag{4.5}$$

with atomic density field φ and model parameter ε. The PFC equation is then derived following the typical conserved PF dynamics

$$\frac{\partial \varphi}{\partial t} = \nabla^2\frac{\delta F_{pfc}}{\delta \varphi}. \tag{4.6}$$

Depending on the choice of ε and the average density of the system $\bar{\varphi}$, this equation (with proper boundary conditions) gives a density field that is either uniform or spatially periodic, which can be then reintepreted as the liquid or the solid phase, respectively. In Fig. 4.2, a typical PFC density field is shown.

The periodic nature in the PFC model is generated by the $\left(q^2 + \nabla^2\right)^2$ term. By looking at this term in the Fourier space, it is easy to identify q as the frequency factor for atom density peaks since it applies an additional energy penalty for any

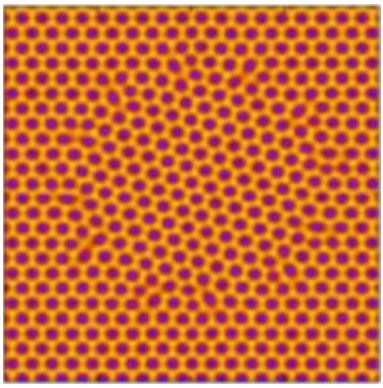

Fig. 4.2 PFC density field showing a circular grain embedded in a matrix. Dark dots in the figure can be seen as "atoms" while a "grain-boundary" can be identified by following the lattice orientation

deviation of the wave number k from the given q. Using a one-mode approximation of the periodic state, one can demonstrate analytically that any deformation to the PFC solid (the spatially periodic state) can be treated as an elastic deformation to the first order. For the basic 2D hexagonal PFC model, the elastic constants are

$$\frac{C_{11}}{3} = C_{12} = C_{44} \sim (q_0 A_0)^2,$$

with atomic spacing a, $q_0 = 2\pi/a$ and density peak amplitude A_0. Dislocations exist naturally in this model at the boundary of periodic states with different lattice orientations; a situation resembling a grain boundary. By calculating the energy cost for forming such a boundary, one can also demonstrate that the boundary energy can be characterized by the well-known Read-Shockley relation at low mis-orientation angles and can be further fitted to grain-boundary energies at high mis-orientaiton angles [12].

A slightly different PFC model can be derived directly from the classical density functional theory (CDFT). The energy for a CDFT system is typically [13]

$$F(n) = \int \{[1+n(r)]ln[1+n(r)] - n(r)\}dr - \int drn(r) \int C_2(r,r')n(r')dr + high\,order,$$

with the scaled density field $n(r) = \rho/\rho_0 - 1$ and pair correlation function $C_2(r, r')$. The first term can be expanded as

$$(1+n)\ln(1+n) - n \sim \frac{n(r)^2}{2} - \frac{n(r)^3}{6} + \frac{n(r)^4}{12} + \ldots, \tag{4.7}$$

while the pair interaction part can be expanded as

$$C_2 \sim C_2^0 - C_2^2\nabla^2 + C_2^4\nabla^4 + \ldots,$$

with expansion coefficients C_2^0, C_2^2 and C_2^4. This expansion of pair correlation function is better explained in Fourier space since

$$\widehat{C}(k) \sim C_2^0 + C_2^2 k^2 + C_2^4 k^4, \tag{4.8}$$

can be used to fit the first peak of $\widehat{C}(k)$. A PFC-type energy functional is then recovered by combining the right-hand side of (4.7) and (4.8).

AE is a classical method that was first developed in the study of hydrodynamic convection patterns in the Navier-Stokes equation. When applying the method to the PFC model, it separates the short wavelength atomic details from the long wavelength component of the density waves. As a result, equations that govern the long-range variation of the density waves in the PFC model can be extracted [14–17].

Since the formulations of the PFC and AE models are rather general, discussion of these approaches are limited here and no examples of application will be given in the next section. For further reading, we refer to the book by Provatas et al. [7] and other reviews [18].

4.3 Applications

4.3.1 Solid-State Phase Transformations

Complex microstructures formed from solid-state phase transformation are commonly observed in alloy materials. Based on the basic physical process involved, there are mainly two categories of solid-state phase transformations, diffusional and diffusionless. The first one is made possible by thermally activated long-range atom transport while the second one is a result of small cooperative motion of local atoms. Using the energy argument, the solid-state transformation can be understood by quantum DFT calculations at the atomic-level. However, at the micron scale, modeling the solid-state transformation and the complex evolution of the resulting microstructures requires a very different approach. By smoothing out the atomic-level details while keeping the essential physics in the solid-state phase transformation, PF models have been widely used to study this phenomenon at the microstructure scale.

To demonstrate the formulation of the PF solid-state phase transformation model, a simple model for diffusionless transformation will be constructed from the very basics. The solid-state phase transformation can be described in Landau's phenomenological theory using long-range order parameter η, which takes one

value in the parent phase and another value in the transformed phase. In this theory, the free energy density f can be written as an even function of the order parameter

$$f = A\eta^2 + B\eta^4 + \ldots, \tag{4.9}$$

with Landau coefficients A and B. If a system is cooled down from a high temperature, several less symmetric variants with equivalent crystal structure form from the more symmetric parent phase below a critical temperature. Such a transformation can be described by (4.9) using a temperature-dependent coefficient A that changes sign from positive to negative at the critical temperature T_c. This is a well-known textbook example of spontaneous symmetry breaking where two equivalent states with ($\eta = \pm\eta_0$) form from the matrix ($\eta = 0$) state (Fig. 4.3).

In order to describe a spatial mixture of the parent phase and the transformed phase, other terms need to be included in the expression of f. The first approach is to expand the energy not only as a function of local order parameters but also as a function of order parameter gradient [8]. To the lowest order, a phase-field type free energy density is then recovered

$$f = \frac{K}{2}|\nabla\eta|^2 + A\eta^2 + B\eta^4. \tag{4.10}$$

Another important contribution to the free energy density is the elastic energy generated from the change of crystal structure in the transformed phase since it has a lattice mismatch with the parent phase. This elastic energy density e can be generally expressed as

$$e = \frac{1}{2}\sigma_{ij}\epsilon_{ij}, \tag{4.11}$$

with stress tensor σ_{ij}, strain tensor ϵ_{ij} and Hooke's law

$$\sigma_{ij} = \lambda\epsilon_{kk}\delta_{ij} + 2\mu\epsilon_{ij}. \tag{4.12}$$

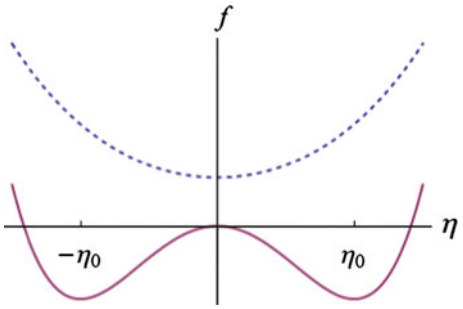

Fig. 4.3 Symmetry breaking as one moves from $T > T_c$ (*dashed line*) to $T < T_c$ (*solid line*)

Here, λ and μ are standard elastic constant (Lamé's first parameter and shear modulus), Kronecker delta δ_{ij}, and Einstein summation is also used. If the transformed phase is not confined by the surrounding parent phase, there is no elastic energy generated and even the system is deformed by the transformation-induced strain ϵ_{ij}^0 (from the change of crystal structure). This indicates that when calculating the elastic energy, contribution from the transformation-induced strain should be removed from the total strain

$$\epsilon_{ij} = \epsilon_{ij}^{total} - \epsilon_{ij}^0 \eta^2. \tag{4.13}$$

For $\eta_0 = 1$, this form of coupling between η and ϵ_{ij}^0 ensures that all the transformation-induced strain is removed in the transformed phase when evaluating the elastic energy. By introducing a displacement field $\vec{u} = u_i \hat{x}_i$, the total strain is then $\epsilon_{ij}^{total} = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right)$. As the transformed phase develops, the displacement field also evolves such that the mechanical equilibrium condition

$$\frac{\partial \sigma_{ij}}{\partial x_j} = 0, \tag{4.14}$$

is satisfied since the transformation is much slower compared with elastic relaxation. Now, to summarize this simple model, the final form of the free energy functional is

$$F = \int \left[\frac{K}{2}|\nabla \eta|^2 + A\eta^2 + B\eta^4 + e\right] dv. \tag{4.15}$$

From here, one can derive the equation of motion for η based on the non-conserved dynamics in (4.3). Evolution of the transformed phase in the parent phase is then governed by the η equation together with the elastic condition (4.14).

The model introduced above is a very simplified version of real world examples [19–22]. When applied to interested materials, several parts in the model can be modified accordingly. For example, more variants can be included by introducing more order parameters $\eta_1, \eta_2, \ldots \eta_n$. In case of three variants, the potential part of the Landau expansion can be written as

$$f_p(\eta_1, \eta_2, \eta_3) = \frac{1}{2}A_1\left(\eta_1^2 + \eta_2^2 + \eta_3^2\right) - \frac{1}{4}A_2\left(\eta_1^4 + \eta_2^4 + \eta_3^4\right) + \frac{1}{6}A_3\left(\eta_1^6 + \eta_2^6 + \eta_3^6\right),$$

with the parent phase given by

$$\eta_1 = \eta_2 = \eta_3 = 0,$$

and three variants at

$$\eta_1 = \pm\eta_0, \eta_2 = \eta_3 = 0$$
$$\eta_2 = \pm\eta_0, \eta_1 = \eta_3 = 0$$
$$\eta_3 = \pm\eta_0, \eta_1 = \eta_2 = 0.$$

The stress-strain relation (4.12) is only valid for isotropic media. When applying to anisotropic materials, a more general relation

$$\sigma_{ij} = C_{ijkl}\epsilon_{kl}, \tag{4.16}$$

with stiffness tensor C_{ijkl} takes over. Anisotropy on the variant-variant surface energy and the variant-matrix surface energy can be introduced into the model by generalizing the expression of the η gradient term in (4.15)

$$\frac{K}{2}|\nabla\eta|^2 \rightarrow \frac{1}{2}\sum_{p=1}^{3}\beta_{ij}(p)\nabla_i\eta_p\nabla_j\eta_p.$$

Using a carefully designed Landau potential, many features in the austenite-martensite transformation in shape-memory alloys can also be captured [22].

For the diffusional solid-state transformation, in addition to the order parameters (and the elastic field) in the previous example, a long-range diffusion field c has to be considered since this category of transformation involves long-range transport of atoms, and the composition of the transformed phase is different from the matrix. The spirit of the Landau theory still applies to the formulation of PF models in this case, but the free energy density form has to be modified such that the effect of concentration is included.

An example in [23] is

$$f = \frac{\alpha}{2}|\nabla c|^2\frac{\beta}{2}|\nabla\eta|^2 + \frac{1}{2}A(c-c')^2 - \frac{1}{2}B\eta^2 - \frac{1}{4}C\eta^4 + \frac{1}{6}D\eta^6 + \frac{1}{2}Gc\eta^2. \tag{4.17}$$

This expression has the order parameter expansion up to 6th order, 2nd order c expansion, and an additional gradient c term. c' is the equilibrium concentration in the matrix. The last term in this energy density is a coupling between the concentration and the order parameter which indicates a very simple dependence between the transformed phase and the concentration. For a specific system, the coefficients in (4.17) can be obtained from other atomistic methods as demonstrated in the multiscale modeling approach [24]. In addition to these modifications of the Landau energy term, a coupling between the transformation-induced strain and the concentration field has to be added to the formulation of the elastic energy.

Considering the simple linear relation given by Vegard's law (lattice parameter is a linear function of concentration), the strain relation in (4.13) becomes

$$\epsilon_{ij} = \epsilon_{ij}^{total} - \epsilon_{ij}^{c} c - \epsilon_{ij}^{0} \eta^2. \tag{4.18}$$

This equation indicates that there are two parts in the total strain that need to be excluded in the elastic energy: one is ϵ_{ij}^{0} which comes from the change of lattice parameter during the formation of the transformed phase; another comes from the dependency of the lattice parameter on the concentration of constituent elements ($\epsilon_{ij}^{c} c$ gives the linear relation in Vegard's law).

Based on (4.17) and (4.18), the equation of motion for η can be readily derived from the free energy functional. However, another equation is required to describe the evolution of the concentration field c during the phase transformation. Since the free energy functional is also a function of c, the same variational method can be applied to derive a conserved equation for the concentration field

$$\frac{\partial c}{\partial t} = \nabla \cdot \frac{D}{KT} \nabla \frac{\delta F}{\delta c}. \tag{4.19}$$

In the case of the single order parameter equation, the kinetic coefficient L in the non-conserved dynamics is not very important since it only changes the rate of transformation and can be scaled out using a proper time constant. Once the diffusion is added, there are now two independent time scales: transformation time (given by L) and diffusion time (given by D). The effect of these two time scales on the system evolution must be evaluated [24].

Other forms of the Landau energy term are possible, such as the one that directly uses the CALPHAD database [25–29]. Recently, another form of the Landau free energy theory, where model parameters can be uniquely determined from interface width and surface energy, starts to appear in solid-state phase transformation PF models. In the study of θ' precipitates in Al–Cu alloy [29], f takes the form

$$f = \sum_{i=1}^{3} \frac{k^2(\varphi_i)}{2} |\nabla \eta_i|^2 + wg(\eta_1, \eta_2, \eta_3) + [1 - h(\eta_1, \eta_2, \eta_3)] f^{\alpha} + h(\eta_1, \eta_2, \eta_3) f^{\theta'}. \tag{4.20}$$

$k(\varphi_i)$ can be related to the surface energy and the interface width, and it depends on the local surface orientation angle φ_i. w is a pre-factor, f^{α} and $f^{\theta'}$ are free energies of α and θ' phase which can be obtained by fitting to a phase diagram [30]. $g(\eta_1, \eta_2, \eta_3)$ is a double-well potential and $h(\eta_1, \eta_2, \eta_3)$ is a monotonous interpolation function. In this form, it not only describes the phase transformation but also correctly incorporates the bulk free energy in the α and θ' phases.

4.3.2 Ferroelectric Materials

Study of ferroelectric materials is an active field due to potential applications in electromechanical systems and electronic devices. Paraelectic-ferroelectric transition happens at Curie temperature, below which spontaneous electric polarization appears as the result of structural transformation. Similar to the formation of different variants in solid-state phase transformation, several crystallographically-equivalent polarized states can appear from the paraelectric phase and form complex domain structures. These polarized states can then be switched from one to another by an external electric field. When deposited on a substrate, ferroelectric thin films are subject to an additional constraint due to lattice mismatch between the substrate material and the film material. The strain energy arising from this constraint applies to the crystal structure of ferroelectric materials and modifies the paraelectic-ferroelectric transition temperature.

Similar to the previously discussed solid-state phase transformation models, one can formulate a PF model to capture the paraelectic-ferroelectric transition. Using polarization P_i as the order parameter, the free energy functional can be written as [31–36]

$$F = \int \left[f_{bulk} + f_{grad} + f_{elas} + f_{elec} \right] dv.$$

According to the phenomenological theory of Devonshire [37], the bulk energy part is

$$f_{bulk} = \frac{1}{2}\alpha_{ij}P_iP_j + \frac{1}{4}\gamma_{ijkl}P_iP_jP_kP_l + \frac{1}{6}w_{ijklmn}P_iP_jP_kP_lP_mP_n + \ldots,$$

with coefficients α, γ and w. The gradient part is

$$f_{grad} = \frac{1}{2}G_{ijkl}\frac{\partial P_i}{\partial x_j}\frac{\partial P_k}{\partial x_l},$$

with gradient coefficient G_{ijkl}. The elastic part is

$$f_{elas} = \frac{1}{2}c_{ijkl}\epsilon_{ij}\epsilon_{kl} - \frac{1}{2}q_{ijkl}\epsilon_{ij}P_kP_l.$$

Clearly, the first term in f_{elas} is the standard strain energy, while the second term is the energy contribution from polarization-induced strain (with electro-restrictive constant q_{ijkl}) and should be removed from the total energy (similar to the case of

transformation-induced strain in the solid state phase transformation). The last electric part is

$$f_{elec} = -P_i\left(E_i + \frac{1}{2}E_{\mathrm{i}}^{\mathrm{d}}\right),$$

with applied electric field E_i and depolarization field $E_{\mathrm{i}}^{\mathrm{d}}$. Ferroelectric domains can evolve to minimize the total energy according to the non-conserved dynamics for the polarization order parameter η_i.

During the microstructure evolution, the force equilibrium condition (4.14) should be satisfied. Assuming no space charge in the film, the electric field should also satisfy electrostatic condition $\nabla \cdot D = 0$. When subject to the thin film boundary condition (constraint from the substrate lattice mismatch), one can, in principle, solve these two equilibrium equations with the relaxation method for each time step of the order parameter equation; however, the high computational cost in this straightforward approach makes experiment-relevant space and time scales difficult to reach in simulations. An efficient Fourier space approach for solving these two boundary value problems was developed [31] based on the elasticity theory of Khachaturyan and Stroh, and has been widely applied in other works in this field. For more details on this topic, we refer readers to the review article [38].

4.3.3 Dislocation

The PF dislocation model was developed based on the free energy scheme similar to the solid-state transformation models [39–42]. By utilizing the fact that dislocations can be described as a set of coherent platelet inclusions [39, 40], contribution of the dislocations to the total elastic energy was expressed through a stress-free strain ε_{ij}^0 (similar to the transformation strain in the solid-state transformation model) which was related to the Burgers vector b by

$$\varepsilon_{ij}^0 = \frac{b_i n_j}{d},$$

where n is the slip plane normal and d is the plate thickness. Using the dislocation density η as the PF parameter, the stress-free strain is then

$$\varepsilon_{ij}^0(r) = \frac{1}{d}\left[\sum_{\alpha}\sum_{m_\alpha} b(\alpha, m_\alpha)_i n(\alpha)_j \eta(\alpha, m_\alpha, r)\right],$$

with spatial coordinate r, slip plane index α, and Burgers vector index m_α. Minimization of the elastic energy gives a dislocation density η that is non-vanishing only within the thin plate; however, this does not put any constraint

on the value of η which should be an integer in case of perfect dislocations. To correctly incorporate this condition, a periodic potential which gives energy minima only at integer values of η was used in the model to replace the multi-well Landau potential in the solid-state transformation model. A simple form of the periodic potential is [39]

$$E_{periodic} = \int \sum_{\alpha} \sum_{m_\alpha} A \sin^2 \pi\eta(\alpha, m_\alpha, r) dv,$$

where the barrier height is given by A. This periodic potential form is similar to the Peierls-Nabarro barrier of dislocation motion, and the amplitude A can be related to shear modulus. From here, evolution equations can be derived following the variational approach in the methods section.

4.3.4 Grain Growth

Grain growth is a well-known phenomenon that happens at high temperature in polycrystalline materials. Associated with an interfacial energy, once thermally activated, grain-boundaries move to minimize the total area and thus the total surface energy in grain growth. Since most materials have a higher yield stress at smaller average grain size, understanding the evolution of average grain size in grain growth is very important in materials design. Due to the complex grain geometry in polycrystalline materials, tracking the grain growth of a statistically significant number of grains is not an easy task in general. Before the PF model appeared, the Monte Carlo Potts model was widely used in the study of grain growth [43].

The first PF grain growth model [44] used multiple order parameters to account for the polycrystalline structure, i.e. each grain was associated with a specific order parameter η_i. The PF free energy functional for this model is

$$F = \int \left[\frac{1}{2} \sum_{i=1}^{p} K_i (\nabla \eta_i)^2 + \sum_{i=1}^{p} \left(-\alpha \frac{\eta_i^2}{2} + \beta \frac{\eta_i^4}{4} \right) + \gamma \sum_{i=1}^{p} \sum_{j<i} \eta_i^2 \eta_j^2 \right], \tag{4.21}$$

with model parameters K_i, α, β, γ and grain index $i = 1$ to p. Evolution equations of the order parameter are then derived using the non-conserved dynamics. This free energy functional can be seen as the total surface energy which is then reduced in grain growth through the non-conserved equation. For each grain, the evolution equation is

$$-\frac{\partial \eta_i}{\partial t} = L_i \left(-K_i \nabla^2 \eta_i - \alpha \eta_i + \beta \eta_i^3 + 2\gamma \eta_i \sum_{j \neq i}^{p} \eta_j^2 \right). \tag{4.22}$$

Compared to the basic PF model in the method section, this one has an additional coupling term between different order parameters. Without this coupling term, the equation has three solutions: two stable solutions at $\eta_i = \pm 1$ with minimum energy and one unstable solution at $\eta = 0$ with maximum energy. With the coupling, linear stability analysis shows that, with $\alpha = \beta$ and $\gamma \leq \alpha/2$, this equation has three stable states at $\eta_i = 0$ and ± 1. For the purpose of grain growth modeling, one can set $\eta_i = 1$ or -1 for the interior of grain i and set $\eta_i = 0$ outside.

By solving (4.22) for a polycrystalline system with p grains, complex grain structure (Fig. 4.4) evolution can be obtained [45]. Very similar grain structure evolutions were obtained in a detailed comparison with the Monte Carlo Potts model [46].

In the PF grain growth model proposed by Steinbach et al. [47], a similar multi-order-parameter approach was used. Rather than adding a coupling term to the standard single PF model, they re-write the free energy based on the coupling between neighboring grains. Detailed numerical analysis of this model was carried out in [48]. Large scale 3-dimensional grain growth simulation based on this model demonstrated a good agreement with predictions from statistical theory [49]. Kobayashi et al. [50, 51] proposed a different approach for this problem. Instead of labeling different grains with different order parameters, they formulated the model with only two order parameters; one was used to distinguish the solid and the liquid state while another parameter took different values based on grain orientation. Such a model reduces the computational cost significantly (2 equations v.s. p equations) but also introduces a time scale from the relaxation of the orientation parameter. Moelans et al. [52] extended PF grain growth models by including surface energy anisotropy and mobility anisotropy in the equation. With a carefully controlled interface width, this model can simulate grain growth in anisotropic systems efficiently. A recent review on this topic can be found in [53].

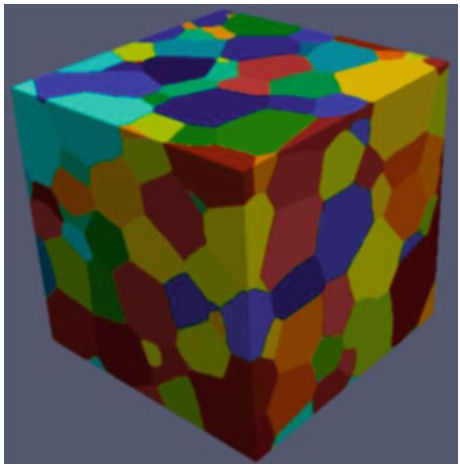

Fig. 4.4 Grain structure from 3-dimensional PF grain growth simulation. Different colors represent different grains

At application level, grain growth is often associated with plastic deformation and re-crystallization. Since the elastic moduli depend on crystal orientation, under applied stress, the stress distribution inside the polycrystalline materials depends strongly on the grain structure and can significantly affect the grain growth. Stress hot spots produced in the evolution can be preferred nucleation sites for further plastic deformation. To understand the internal stress evolution and its effect on grain growth, elasticity has to be added into the PF grain growth model. The effect of strain energy on grain-boundary migration was studied using the PF model [54] in simple bicrystal geometry. For complex polycrystalline materials, the effect of strain energy was studied in [55] by taking into account the difference of elastic moduli in different grains based on Khachaturian's theory for inhomogeneous solids [56] and an efficient spectral iterative method [57].

4.3.5 Dendrite Solidification

The appearance of complex dendrite structures in the solidification process and the sensitive shape dependence of these structures on external conditions have puzzled scientists for a long time (Fig. 4.5). The basic physics of crystal growth can be found in many excellent reviews [58, 59] and is out of the scope of this chapter. Here, we will mainly focus on the application of the PF model to dendrite solidification, discuss its formulation and demonstrate how the PF model reproduces the correct physics in the sharp-interface limit, with only a brief introduction to the underlying physics.

As another major application, the PF solidification model was developed along a different path from its cousins in the solid-state transformation model. Although the equations of motion for PF variables are similar in these two models, the physics behind them are quite different. As we outlined in the introduction and demonstrated in earlier sections, the PF model for solid-state phase transformation and its variants use physical order parameters as the PF variable and evolve the system by minimizing the total free energy, which is written as a function of order parameters based on a Landau-type potential. In the PF solidification model, the reason for

Fig. 4.5 A typical dendrite crystal captured in PF model

introducing a PF variable is rather mathematical. The PF variable, in this case, is nothing but a regularized field to track the evolution of the complex surface shape developed during the solidification process. Therefore, the goal of the PF solidification model is to accurately reproduce the corresponding sharp-interface physics using a diffuse-interface approach. Such a goal is accomplished by carrying out a multiscale analysis of the PF model in the limit of vanishing interface thickness and match the resulting equation to the original sharp-interface problem.

Before discussing the details of PF solidification model, it is important to understand some basic physics in crystal growth and dendrite formation. There are mainly two regimes in growth phenomenon: reaction-limited and diffusion-limited. The first one occurs when the growth is controlled by the attachment barrier of liquid atoms to a solid structure, i.e. the probability of attachment (which is proportional to $e^{-\Delta G/KT}$ with attachment barrier ΔG). The second one happens when the growth is controlled by the availability of atoms, i.e. diffusion transport of atoms from bulk to the interface area. The solidification problem often falls into the diffusion-limited category.

Now, considering a flat solid-liquid interface with a small shape perturbation caused by a local fluctuation of atom availability (Fig. 4.6), an increased diffusion gradient due to this shape perturbation will further enhance the transport to this area through the diffusion equation

$$\frac{\partial c}{\partial t} = \frac{\partial}{\partial y} D \frac{\partial c}{\partial y}.$$

On the other hand, such a shape perturbation increases the surface energy of this solid-liquid interface and thus tends to be eliminated by energy minimization. In case of dendrite formation, the competition of these two mechanisms causes the growth of the main branch initially, and then of the side branches later as the same mechanism works again on the side walls of the main branch.

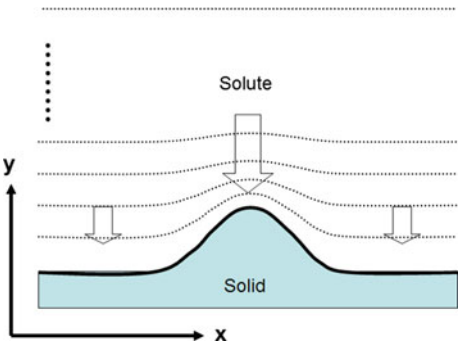

Fig. 4.6 Schematic plot of the concentration profile in response to a surface shape perturbation. Dashed lines are concentration contours

The interface stability problem described above is known as Mullins-Sekerka instability [60] and is the fundamental reason for dendrite formation in crystal growth. Here we have an equation set of solidification

$$\begin{cases} D\nabla^2 u = \dfrac{\partial u}{\partial t} & (4.23a) \\ v = \dfrac{dy_i(t)}{dt} = D\dfrac{\partial u}{\partial n} & (4.23b) \\ u_i = -d_0\kappa & (4.23c) \end{cases}$$

with u as a dimensionless diffusion field, interface position y_i, surface normal n, surface curvature κ and capillary length d_0. Equation (4.23a) simply describes the diffusion of u. Equation (4.23b) states that the interface motion velocity is determined by the normal flux of u. Equation (4.23c) is the Gibbs-Thompson condition at the interface with subscript i for interface.

With a small shape perturbation, the interface profile can be written as

$$y_i(x,t) = y_i(t) + A_k e^{ikx+\omega t},$$

with perturbation amplitude A_k, decay factor ω and wave number k. One can demonstrate that such a perturbation will decay or grow based on the following condition

$$\omega = vk - d_0 D k^3. \tag{4.24}$$

Equation (4.24) indicates that there exists a critical wavelength $2\pi/k_c = 2\pi\sqrt{Dd_0/v}$ beyond which flat interface motion described by equation set (4.23) becomes unstable.

The goal of the PF dendrite solidification model is to recover (4.23) (and the associated interface instability) at the sharp-interface limit. In a generalized form, (4.23) is written as

$$\begin{cases} D\nabla^2 u = \dfrac{\partial u}{\partial t} & (4.25a) \\ v = D\left[\dfrac{\partial u}{\partial n}\Big|^{-} - \dfrac{\partial u}{\partial n}\Big|^{+}\right] & (4.25b) \\ u_i = -d_0\kappa - \beta v & (4.25c) \end{cases}$$

Comparing (4.25) with (4.23), the interface motion described in the 2nd equation is now due to the difference of diffusion transport into and out of the interface area. The 3rd equation is a generalized Gibbs-Thomson condition with a kinetic effect term βv. The PF solidification model with PF variable ϕ is typically written in two equations [61–64]

$$\begin{cases} D\nabla^2 u + \dfrac{1}{2}\dfrac{\partial h}{\partial t} = \dfrac{\partial u}{\partial t} & (4.26a) \\ \tau\dfrac{\partial \varphi}{\partial t} = w^2\nabla^2\varphi - \dfrac{\partial f}{\partial \varphi} - \lambda u\dfrac{\partial g}{\partial \varphi} & (4.26b) \end{cases}$$

The 2nd equation is the standard one shown in the general formulation in the method section. The potential function f gives two minima at the solid and the liquid state, and an energy bias function g, coupled with the diffusion field, drives the solidification process by lowering the free energy of the solid phase. The 1st equation handles the diffusion of u in the bulk phase and the consumption/generation of u due to the interface motion ($\frac{\partial h}{\partial t}$ term). PF models for polycrystalline and multi-component solidification can be found in [50, 65, 66].

The relation between (4.25) and (4.26) is established by carrying out a sharp-interface expansion of the PF equation (4.26). To do that, one needs to look at (4.26) and pattern scale l_c where the interface width is small, i.e. $w/l_c \to 0$. With expansions of the u and φ fields

$$\begin{cases} \varphi = \varphi_0 + p\varphi_1 + \dots \\ u = u_0 + pu_1 + \dots \end{cases}$$

Equation (4.26b) can be exactly reduced to the sharp-interface equation for the Gibbs-Thomson relation (4.25c) while (4.26a) recovers (4.25a) in the bulk and (4.25b) at the solid-liquid interface [61, 62]. Therefore the method of sharp-interface expansion gives a rigorous mathematical background for using the PF method to solve the original sharp-interface solidification problem. However, convergence of computational results using (4.26) to the sharp-interface limit was only achieved under rather stringent conditions [64]. The problem lies at the analysis of the PF diffusion field. Coupled with the PF parameter, it can be deemed as a constant at the interface only when $w/l_c \to 0$ is strictly satisfied. However, in a computational model, with limited grid points (a few hundred) along each dimension, one generally sets the diffused interface region to be a small fraction of the pattern scale ($w/l_c \ll 1$ but remains finite) such that the length scale required to track the complex dendrite evolution can be reached. In this case, the variation of the diffusion field is proportional to wv/D, which is the ratio of interface width w over diffusion length D/v. Such a variation of the diffusion field across the interface region produces additional contributions in PF numerics which are not captured in the sharp-interface expansion [64, 67]. Therefore, (4.26) is only an approximation of the original sharp-interface problem (4.25). Since dendrite growth is quite sensitive to growth conditions, the small contribution introduced in the PF approximation may cause very different growth behavior in numerical simulations. Quantitative PF numerical results are only possible in the limit of very fast diffusion across the interface scale where wv/D becomes very small.

In order to remove this constraint, an improved sharp-interface analysis (so-called thin interface analysis) which takes into account the variation of the diffusion field within the interface region was obtained for pure materials [64] and alloys [68, 69]. The corrections due to the finite interface width are explicitly derived in this limit. The refined PF model based on this analysis can therefore exactly recover the sharp-interface problem (4.25), capture the dendrite formation quantitatively in 3-dimensions, and make experiment-relevant predictions [70].

4.3.6 Fracture

The fundamental physics involved in fracture propagation include the releasing of elastic energy and the increasing of crack surface energy. A commonly agreed upon criterion for crack propagation goes back to the work of Griffith and Irwen based on a crack extension force (for plane strain)

$$G = \alpha\left(K_I^2 + K_{II}^2\right) + \frac{K_{III}^2}{2\mu},$$

where $K_{I,II,III}$ is the stress intensity factor for three standard modes of fracture, μ is the shear modulus, v is Poisson's ratio and $\alpha = (1 - v)/2\mu$. The propagation is possible only for G larger than a critical G_c which is then given by the cost of creating two new surfaces ($G_c = 2\gamma$). From this point of view, the phase-field model coupled with elasticity should be a good approach for this class of problems. Several PF models have been proposed to study the propagation of fracture in brittle materials [71–73]. The basic idea is to treat the fracture propagation as a phase transition process from the solid phase to a "broken" fracture phase. Similar to the solid-state phase transformation models where the elastic energy played an essential role in determining the morphologies of solid-state precipitates, fracture models also use a phase-field energy with both the surface and the elastic contributions as the starting point and produce equations of motion from the minimization of this energy. In the work of Hakim et al. [52], this phase-field energy takes a simple form

$$F = E_{pf} + g(\phi)(E_{strain} - E_c) + E_c,$$

where the surface part takes the commonly seen form $E_{pf} = \frac{k}{2}(\nabla\phi)^2$, the elastic part is $E_{strain} = \frac{\lambda}{2}(\varepsilon_{ii})^2 + \mu\varepsilon_{ij}\varepsilon_{ij}$, and $g(\phi)$ is a smooth interpolation function which describes the softening of elastic energy in the "broken" state. E_c is the threshold for the broken state, above which fracture is favored energetically.

The relation between the phase-field approaches and classic sharp-interface fracture mechanics have been explained in detail and demonstrated numerically in many works. Using asymptotic expansion, Hakim et al. [52] showed that the crack propagation velocity in their PF model is proportional to the departure from the classic Griffith criterion. Spatschek et al. [53] also analyzed the sharp-interface limit of their PF fracture model and showed that the crack propagation can be described as a result of the Asaro-Tiller-Grinfeld (ATG) instability of crack surface under applied stress.

4.3.7 Vesicle Morphology

Most applications of the PF model are within the community of materials sciences; however, the model also garners significant attention from applied mathematicians.

Biben et al. [74] and Du et al. [75] demonstrated its potential in bioscience by developing PF models for the morphology of vesicle blood cells. In their models, by minimizing a membrane bending energy, different experimentally observed morphologies of vesicle cells are successfully constructed using the PF method. Since the energy form they used in the model is different from most other PF models, it is worthwhile to discuss this application in further detail. The elastic bending energy of a vesicle membrane is generally given by

$$\varepsilon_H = \int \left[\frac{b}{2}(H - c_0)^2 + cK + \sigma\right] ds,$$

where b and c are bending rigidities, H and K are the mean and Gaussian curvatures, c_0 is the spontaneous curvature of the system and σ is the orientation-dependent surface energy. The integral is over the entire surface area. This energy form is different from most PF models we introduced earlier, and Biben et al. and Du et al. successfully constructed a simplified version of ε_H by noticing the fact that the mean curvature H can be easily written in terms of the PF parameters.

It is well know that the variational dynamics based on the classic PF energy $F = \int \left(\frac{\epsilon}{2}|\nabla\psi|^2 + \frac{1}{4}\left(\psi^2 - 1\right)\right) dv$ are reduced to a surface motion driven by the mean curvature in the sharp-interface limit. To construct the bending energy with only the mean curvature part using PF parameter ψ, one can make a connection by writing

$$Sharp - interface\text{: } \varepsilon_H = \int H^2 ds$$

$$Phase - field\text{: } \varepsilon_H^{PF} = \int \left(\frac{\delta F}{\delta \psi}\right)^2 ds$$

With the PF form of elastic bending energy, the vesicle membrane shape is then given by the minimization of ε_H^{PF}, i.e. $\delta\varepsilon_H^{PF}/\delta\psi = 0$. Further asymptotic analysis showed that such a minimization of ε_H^{PF} recovered the correct Willmore stress relation in the sharp-interface limit. Later, this vesicle membrane model was generalized to include the Gaussian curvature term in the bending energy, and the effect of an external fluid field was also studied since it is relevant to many biological applications. We recommend [76] for a detailed review on this topic.

4.4 Perspectives

Through the basic PF formulation and its several applications, the strength of the PF model as a powerful approach for complex interface evolution problems has been demonstrated. Its role in multiscale materials modeling is to connect the fundamental properties one gets from atomistic/electronic models to large-scale application-level materials behavior. With naturally incorporated surface energy

and the flexibility to include other physics, this modeling method by itself has also received significant attention from physicists, materials scientists and applied mathematicians. New applications of the PF method are continually being developed in modern society, and a fast increasing category of problems can now be treated using this approach. To conclude this chapter, we summarize a few important points one should keep in mind when adopting the PF method to solve a specific problem of interest. First, one must ensure that the PF method is the correct approach for the problem. The PF method should either offer a length/time scale advantage that other methods cannot reach or should result in improved computational efficiency. Second, the connection between underlying physics and the PF model must be established, typically by using a physics-based energy functional or sharp-interface asymptotics. Third, additional effects introduced by the diffuse interface should be kept in mind. When using the phase-field variable only, the interaction distance between diffuse interfaces could be much longer than the interaction distance between sharp interfaces. When coupled with other fields, additional steps (i.e., a thin-interface analysis) are needed to treat the variation within the diffuse interface region. Additionally, application of the PF model with dynamics other than the two classical variational dynamics (Allen-Cahn or Cahn-Hilliard equation) is possible.

References

1. L.Q. Chen, Annu. Rev. Mater. Res. **32**, 113 (2002)
2. W.J. Boettinger, W.A. Warren, C. Beckermann, A. Karma, Annu. Rev. Mater. Res. **32**, 163 (2002)
3. H. Emmerich, Adv. Phys. **57**, 1 (2008)
4. I. Steinbach, Modell. Simul. Mater. Sci. Eng. **17**, 073001 (2009)
5. Y. Wang, J. Li, Acta Mater. **58**, 1212 (2010)
6. I. Steinback, Annu. Rev. Mater. Res. **43**, 89 (2013)
7. N. Provatas, K. Elder, *Phase-field Methods in Materials Science and Engineering* (Wiley, Weinheim, 2010)
8. J.W. Cahn, J.E. Hilliard, J. Chem. Phys. **28**, 258 (1958)
9. S.M. Allen, J.W. Cahn, Acta Metall. **27**, 1085 (1979)
10. L.Q. Chen, J. Shen, Comp. Phys. Comm. **108**, 147 (1998)
11. J. Shen, X. Yang, Discrete Continuous Dyn. Sys. A **28**, 1669 (2010)
12. K.R. Elder, M. Grant, Phys. Rev. E **70**, 051605 (2004)
13. K.R. Elder, N. Provatas, J. Berry, P. Stefanovic, M. Grant, Phys. Rev. B **75**, 064107 (2007)
14. B.P. Athreya, N. Goldenfeld, J.A. Dantzig, Phys. Rev. E **74**, 011601 (2006)
15. K. Wu, A. Karma, Phys. Rev. B **76**, 184107 (2007)
16. Z.F. Huang, K.R. Elder, Phys. Rev. Lett. **101**, 158701 (2008)
17. R. Spatschek, A. Karma, Phys. Rev. E **81**, 214201 (2010)
18. H. Emmerich, H. Lowen, R. Wikktowski, T. Gruhn, G.I. Toth, G. Tegze, L. Granasy, Adv. Phys. **61**, 665 (2012)
19. Y. Wang, L.Q. Chen, A.G. Khachaturyan, Acta Metall. Mater. **41**, 279 (1993)
20. A. Artemev, Y. Wang, A.G. Khachaturyan, Acta Mater. **48**, 2503 (2000)
21. A. Artemev, Y. Jin, A.G. Khachaturyan, Acta Mater. **49**, 1165 (2001)
22. V.I. Levitas, V.A. Levin, K.M. Zingerman, E.I. Freiman, Phys. Rev. Lett. **103**, 025702 (2009)

23. Y. Wang, A.G. Khachaturyan, Acta Metall. Mater. **43**, 1837 (1995)
24. V. Vaithyanathan, C. Wolverton, L.Q. Chen, Phys. Rev. Lett. **88**, 125503 (2002)
25. U. Grafe, B. Böttger, J. Tiaden, S.G. Fries, Scripta Mater. **42**, 1179 (2000)
26. J.Z. Zhu, Z.K. Liu, V. Vaithyanathan, L.Q. Chen, Scripta Mater. **46**, 401 (2002)
27. I. Loginova, J. Odqvist, G. Amberg, J. Ågren, Acta Mater. **51**, 1327 (2003)
28. J. Eiken, B. Böttger, I. Steinbach, Phys. Rev. E **73**, 066122 (2006)
29. S.Y. Hu, J. Murray, H. Weiland, Z.K. Liu, L.Q. Chen, Comput. Coupling Phase Diagrams Thermochem. **31**, 303 (2007)
30. J.L. Murray, Int. Metall. Rev. **30**, 211 (1985)
31. Y.L. Li, S.Y. Hu, Z.K. Liu, L.Q. Chen, Appl. Phys. Lett. **78**, 3878 (2001)
32. A.K. Soh, Y.C. Song, Y. Ni, J. Am. Ceram. Soc. **89**, 652 (2006)
33. K. Dayal, K. Bhattacharya, Acta Mater. **55**, 1907 (2007)
34. J. Wang, T.Y. Zhang, Acta Mater. **55**, 2465 (2007)
35. J. Slutsker, A. Artemev, A. Roytburd, Phys. Rev. Lett. **100**, 087602 (2008)
36. N. Ng, R. Ahluwalia, D.J. Srolovitz, Acta Mater. **60**, 3632 (2012)
37. A.F. Devonshire, Philos. Mag. **40**, 1040 (1949)
38. L.Q. Chen, J. Am. Ceram. Soc. **91**, 1835 (2008)
39. Y.U. Wang, Y. Jin, A.G. Khachaturyan, J. Appl. Phys. **92**, 1351 (2002)
40. D. Rodney, Y. Le Bouar, A. Finel, Acta Mater. **51**, 17 (2003)
41. N. Ghoniem, S. Tong, L. Sun, Phys. Rev. B **61**, 913 (2000)
42. M. Koslowski, A.M. Cuitiño, M. Ortiz, J. Mech. Phys. Solids **50**, 2597 (2002)
43. Elizabeth A. Holm, Corbett C. Battaile, JOM **53**, 20 (2001)
44. L.Q. Chen, W. Yang, Phys. Rev. B **50**, 15752 (1994)
45. C. Krill, L.Q. Chen, Acta Mater. **50**, 3059 (2002)
46. V. Tikare, E.A. Holm, D. Fan, L.Q. Chen, Acta Mater. **47**, 363 (1998)
47. I. Steinbach, F. Pezzolla, B. Nestler, M. Seesselberg, R. Prieler, G.J. Schmitza, J.L.L. Rezendeb, Physica D **94**, 135 (1996)
48. H. Garcke, B. Nestler, B. Stoth, SIAM J. Appl. Math. **60**, 295 (1999)
49. S.G. Kim, D.I. Kim, W.T. Kim, Y.B. Park, Phys. Rev. E **74**, 061605 (2006)
50. R. Kobayashi, J.A. Warren, W.C. Carter, Physica D **119**, 415 (1998)
51. J.A. Warren, R. Kobayashi, A.E. Lobkovsky, W.C. Carter, Acta Mater. **51**, 6035 (2003)
52. N. Moelans, B. Blanpain, P. Wollants, Phys. Rev. Lett. **101**, 025502 (2008)
53. N. Moelans, B. Blanpain, P. Wollants, Calphad **32**, 268 (2008)
54. M. Tonks, P. Millett, W. Cai, D. Wolf, Scripta Mater. **63**, 1049 (2010)
55. S. Bhattacharyya, T.W. Heo, K. Chang, L.Q. Chen, Modell. Simul. Mater. Sci. Eng. **19**, 035002 (2011)
56. A.G. Khachaturyan, S. Semenovskaya, T. Tsakalakos, Phys. Rev. B **52**, 15909 (1995)
57. S.Y. Hu, L.Q. Chen, Acta Mater. **49**, 1879 (2001)
58. J.S. Langer, Rev. Mod. Phys. **52**, 1 (1980)
59. M. Asta, C. Beckermann, A. Karma, W. Kurz, R. Napolitano, M. Plapp, G. Purdy, M. Rappaz, R. Trivedi, Acta Mater. **57**, 941 (2009)
60. W.W. Mullins, R.F. Sekerka, J. Appl. Phys. **34**, 323 (1963)
61. G. Caginalp, Phys. Rev. A **39**, 5887 (1989)
62. A.A. Wheeler, W.J. Boettinger, G.B. McFadden, Phys. Rev. A **45**, 7424 (1992)
63. J.A. Warren, W.J. Boettinger, Acta Metall. Mater. **43**, 689 (1995)
64. A. Karma, W.-J. Rappel, Phys. Rev. E **57**, 4323 (1998)
65. J. Tiaden, B. Nestler, H.J. Diepers, I. Steinbach, Physica D **115**, 73 (1998)
66. S.G. Kim, W.T. Kim, T. Suzuki, Phys. Rev. E **60**, 7186 (1999)
67. R.F. Almgren, SIAM J. Appl. Math. **59**, 2086 (1999)
68. B. Echebarria, R. Folch, A. Karma, M. Plapp, Phys. Rev. E **70**, 061604 (2004)
69. A. Karma, Phys. Rev. Lett. **87**, 115701 (2001)
70. T. Haxhimali, A. Karma, F. Gonzales, M. Rappaz, Nat. Mater. **5**, 660 (2006)
71. I. Aranson, V. Kalatsky, V. Vinokur, Phys. Rev. Lett. **85**, 118 (2000)
72. V. Hakim, A. Karma, J. Mech. Phys. Sol. **57**, 342 (2009)

73. R. Spatschek, C. Müller-Gugenberger, E. Brener, Phys. Rev. E **75**, 066111 (2007)
74. T. Biben, C. Misbah, Phys. Rev. E **67**, 031908 (2003)
75. Q. Du, C. Liu, X. Wang, J. Comput. Phys. **198**, 450 (2004)
76. Q. Du, Philo. Mag. **91**, 165 (2011)

Chapter 5
Peridynamics

Ibrahim Guven

Abstract In this chapter, we present an integral representation of continuum mechanics which serves as the underlying theory for the peridynamics model. Practical implementation of this model into simulations illustrate the power of this approach. In particular, mechanical deformation from impact and external strain display that the peridynamic model captures failure processes that are difficult, if not impossible, to obtain from other continuum models.

5.1 Introduction

The applied materials science is closely linked with engineering design through the common interest in properties of materials. The mechanical engineers are concerned with functionality and reliability of engineered products while materials scientists explore the link between the macro-scale properties and atomistic/molecular level structure of the constituent materials. Considerations surrounding damage and failure of materials are important parts of the design process. While an engineered part must be practical and functional, it must also serve reliably for a predictable amount of time. Analyses that simulate the response under real physical conditions are commonly utilized during the design of engineering systems, which include, among others, stress analysis with the purpose of identifying concentrations of stress. The interest in finding locations of high stress concentrations is almost exclusively due to concerns of fracture and failure, which is directly related to reliability of an engineering system.

Traditionally, the prediction of deformations and stresses have been performed using the finite element method (FEM), which is a well-established and robust method. While providing accurate predictions for stress components, traditional FEM's capability of capturing realistic material response in modeling and

I. Guven (✉)
Virginia Commonwealth University, Richmond, VA, USA
e-mail: iguven@vcu.edu

P.A. Deymier et al. (eds.), *Multiscale Paradigms in Integrated Computational Materials Science and Engineering*, Springer Series in Materials Science 226,
DOI 10.1007/978-3-319-24529-4_5

predicting failure is questionable. This is due to the underlying assumption that the material remains continuous during deformation; there is no inherent mathematical formulation accounting for possibility of discontinuities arising, which is the case when a crack initiates. One possibility is to include the crack geometry in the analysis. However, since FEM is formulated based on partial differential equations, computations at/near the geometric discontinuities defined by crack faces and crack tips break down due to spatial derivatives being undefined. While a number of remedies are proposed to address this shortcoming, they involve cumbersome special techniques that treat fracture as a special case rather than including it as an inherent material behavior.

In addition to the difficulties of handling discontinuities mathematically, one must be aware that fracture is almost always a dynamic phenomenon which involves propagation of cracks. In engineered systems with complex three-dimensional geometries, cracks rarely propagate along a flat plane; they often turn, branch, oscillate and stop. A predictive computational approach must be able to address this variable nature of fracture in three-dimensions and in a dynamic sense. This task has proven to be most challenging in traditional FEM approaches. In recent past, a number of enhancements have been proposed; two such methods that gained general acceptance in the fracture prediction community are (i) XFEM [1] and (ii) cohesive zone methods [2]. Both of these methods made significant improvements to FEM's handling of problems involving fracture, they require externally supplied kinetic relations for crack growth without providing a damage model as part of the constitutive relations.

Silling [3, 4] introduced a new theory, peridynamic (PD) theory, that represents the continuum mechanics in integral form, which removes the mathematical artifacts present in classical formulations involving spatial derivatives. In the PD theory, internal forces are expressed through nonlocal interactions between the material points within a continuous body, and damage is part of the constitutive model. Additionally, interfaces between dissimilar materials have their own properties and damage can propagate when and where it is energetically favorable for it to do so. The integral equation representation allows crack initiation and growth simultaneously at multiple sites with arbitrary paths.

In this chapter, first, peridynamic theory formulation and its numerical implementation are presented briefly. A number of applications are considered in order to demonstrate the capabilities of the theory.

5.2 Integral Representation of Continuum Mechanics

In PD theory, a point in a solid body interacts with all points within its range as illustrated in Fig. 5.1. Even though the discretization considers material points (as opposed to elements), the body is assumed continuous. The PD theory reformulates the deformation of a continuous body in terms of integral equations without the spatial derivative terms of the classical continuum theory. This feature allows PD

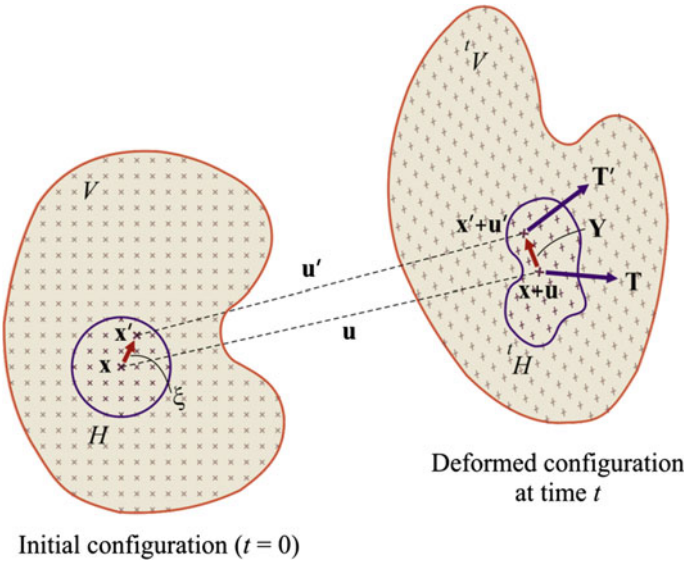

Fig. 5.1 Kinematic description of a pair of PD material points

theory formulation to apply everywhere in the solid body even when geometric and material discontinuities are present.

In peridynamic theory, motion of a material point $\mathbf{x}$ is evaluated through analysis of its interactions with other material points $\mathbf{x}'$ in the body. There exists a distance beyond which the influence of $\mathbf{x}'$ on motion of $\mathbf{x}$ become negligible. This distance is termed *horizon* and the region defined by the horizon is called the neighborhood of $\mathbf{x}$. For example, the circular region in the initial configuration shown in Fig. 5.1 is the neighborhood of $\mathbf{x}$. The equations describing the motion of $\mathbf{x}$ at the deformed configuration at time t is written as

$$\rho \frac{\partial^2 \mathbf{u}}{\partial t^2} = \int_H \{\mathbf{T}(\mathbf{x}, t; \mathbf{x} - \mathbf{x}') - \mathbf{T}(\mathbf{x}', t; \mathbf{x}' - \mathbf{x})\} dH + \mathbf{b}(\mathbf{x}, t) \tag{5.1}$$

in which ρ is the mass density, $\mathbf{u}$ is the displacement vector of material point $\mathbf{x}$, and $\mathbf{b}(\mathbf{x}, t)$ is the body load vector. The integral is the vector summation of internal forces between the material points $\mathbf{x}$ and $\mathbf{x}'$, and is taken over the volume surrounding the material point, H, which defines the neighborhood of $\mathbf{x}$. The initial distance between $\mathbf{x}$ and $\mathbf{x}'$ is denoted by the vector $\boldsymbol{\xi}$ shown in Fig. 5.1, initial configuration. At time t, the body is deformed with $\mathbf{x}$ and $\mathbf{x}'$ moving to new positions with corresponding displacement vectors $\mathbf{u}$ and $\mathbf{u}'$, respectively. The initially spherical (circular in 2-D) neighborhood of $\mathbf{x}$ becomes, in general, an irregularly shaped domain. The force vectors between $\mathbf{x}$ and $\mathbf{x}'$ are described formally by force state vectors $\mathbf{T}(\mathbf{x}, t; \boldsymbol{\xi})$ and $\mathbf{T}(\mathbf{x}', t; \boldsymbol{\xi})$, former belonging to $\mathbf{x}$ and

latter to $\mathbf{x}'$, then the corresponding deformed configuration is described by the deformation state vector $\mathbf{Y}$. It should be noted that the force state vectors are functions of time t, initial positions $\mathbf{x}$ and $\mathbf{x}'$, and the relative position $\boldsymbol{\xi}$, rendering the PD formulation Lagrangian. Detailed derivations related to force and deformation state vectors can be found in Silling et al. [1] and Madenci and Oterkus [2].

In this chapter, the force state vectors, $\mathbf{T}(\mathbf{x}, t; \boldsymbol{\xi})$ and $\mathbf{T}(\mathbf{x}', t; \boldsymbol{\xi})$, describing the interaction between the material points $\mathbf{x}$ and $\mathbf{x}'$ will be assumed to have same magnitude and opposite directions. This assumption leads to the so-called "bond-based" peridynamic formulation with the following PD equation of motion

$$\rho \frac{\partial^2 \mathbf{u}}{\partial t^2} = \int_H \mathbf{f}(\mathbf{x}, \mathbf{x}', t; \boldsymbol{\xi}) dH + \mathbf{b}(\mathbf{x}, t) \tag{5.2}$$

where the interaction between $\mathbf{x}$ and $\mathbf{x}'$ is described by the response function $\mathbf{f}$. It is explicitly written for an isotropic material as

$$\mathbf{f}(\boldsymbol{\eta}, \boldsymbol{\xi}) = \frac{\boldsymbol{\xi} + \boldsymbol{\eta}}{|\boldsymbol{\xi} + \boldsymbol{\eta}|} \mu c s \tag{5.3}$$

in which $\boldsymbol{\xi}$ is the relative position ($\boldsymbol{\xi} = \mathbf{x} - \mathbf{x}'$) and $\boldsymbol{\eta}$ is the relative displacement ($\boldsymbol{\eta} = \mathbf{u}' - \mathbf{u}$). Based on the relative position and displacements, the stretch between $\mathbf{x}$ and $\mathbf{x}'$ can be defined as

$$s = \frac{|\boldsymbol{\xi} + \boldsymbol{\eta}| - |\boldsymbol{\xi}|}{|\boldsymbol{\xi}|} \tag{5.4}$$

The stretch in Peridynamic theory is analogous to the strain in classical continuum theory. The relationship between the force and stretch, (5.3), involves a material parameter, c, which is commonly referred to as the *bond constant*. Its value is evaluated by calculating the strain energy of a PD domain and equating it to the corresponding strain energy that is found under classical continuum mechanics formulation; Silling and Askari [3] used isotropic extension as the loading for this operation to find c as

$$c = \frac{18\kappa}{(\pi \delta^4)} \tag{5.5}$$

Finally, the binary function μ is utilized for introducing failure between two material points as follows

$$\mu(t, \boldsymbol{\xi}) = \begin{cases} 1 & \text{if } s(t', \boldsymbol{\xi}) < s_0 \text{ for all } 0 \leq t' \leq t \\ 0 & \text{otherwise} \end{cases} \tag{5.6}$$

where s_0 is the critical stretch. When the stretch between a material point pair exceeds the critical stretch, the interaction between the material points is permanently terminated. During the simulation, each interaction is checked against this criterion $(s < s_0)$ at each time step and number of broken interactions emanating from each material point is recorded. Thus, damage at a material point is defined as the ratio of broken interactions to the original number of interactions; damage is a value between zero (no damage) and one (completely detached from other material points). Therefore, damage at a material point along a well-defined crack face in a brittle fracture problem is expected to be around 0.5.

5.3 Practical Implementation

Solution of peridynamic equations of motion (5.1) or (5.2) is achieved through numerical means. This involves discretization of the solution domain into small and simple sub-volumes. These sub-volumes should be simple 3-D shapes as they are used in numerical volume integration of PD equations; the simplest sub-volume that enables the easiest numerical handling of PD analysis is a cubic sub-volume. In this case, a peridynamic material point is placed at the center of the cubic sub-volume. After discretization, the volume integral of (5.2) can be written for material point i in terms of summations as

$$\rho \ddot{\mathbf{u}}_i^n = \sum_j \mathbf{f}\left(\mathbf{u}_i^n, \mathbf{u}_j^n, \mathbf{x}_i^n, \mathbf{x}_j^n\right) V_j + \mathbf{b}\left(\mathbf{x}_i^n\right) \tag{5.7}$$

in which V_j represents the volume of the sub-volume represented by material point j, and the superscript n denotes the time step. This equation is written at each material point i to calculate the acceleration vector at point i by equating its product with the mass density to the forces acting on point i within its horizon.

The numerical implementation of the PD theory involves a number of practical issues surrounding the volume integration, time marching and convergence. Each of these issues have been addressed in varying levels of detail in literature; a comprehensive discussion of these topics and further considerations can be found in [2]. A brief summary is provided herein.

The summation of (5.7) approximately represents the volume integral of (5.2) where the volume is the neighborhood of point i defined by a sphere with a certain radius (horizon). In the summation representation, each sub-volume is lumped in a material point located at the mass center of the sub-volume. Thus, if the material point j who belongs to the neighborhood of material point i is within this radius, its whole volume is included in the summation. As it can be foreseen, this situation may pose inaccuracies for material points j that are near the boundary of the sphere where part of the volume may fall outside of the sphere. A similar situation arises when the material point j narrowly falls outside the horizon while a non-trivial

portion of its volume is still inside the horizon. These two scenarios should be addressed in order to increase the accuracy of the computations. The remedy may involve a simple algorithm that estimates the amount of volume that falls within the horizon and adjust the contribution accordingly; this is implemented by [4].

Another common issue related to numerical implementation is the effect of free surfaces that exist within the horizon of material the material point *i*. Evaluation of the material parameters of PD model (5.5) involves comparison of energy quantities between classical continuum mechanics and PD theory; this is performed based on the assumption that the sphere defining the neighborhood of a material point is completely populated (whole sphere). However, as with every finite geometry, existence of physical boundaries will contradict this assumption and a correction must be made. Silling [4] describes this operation as the *surface correction*. The computed material parameters of PD must be corrected according to the local geometry of the material point *i*. Detailed discussion on surface corrections is given in [2].

The peridynamic equation of motion, while not having spatial derivatives, contains time derivative term (acceleration of material point *i*). In the numerical implementation of PD, a time marching scheme must be employed. The PD simulation starts with complete list of known positions, velocities and accelerations (initial conditions). The acceleration term on the left hand side of (5.7) is expressed in terms of a combination of displacements and velocities belonging to current (known) and next (unknown) time steps such that a complete list of displacements, velocities and accelerations in the next time step is calculated. This is a rather routine operation common to many transient problems of continuum mechanics; any traditional method can be employed (e.g., finite difference, Runge-Kutta, Adams-Bashforth). The well-known error considerations apply; a stability criterion for each of these methods can be constructed to find a suitable time step size.

The peridynamic theory works with the equations of motion, and thus a dynamic formulation by nature. However, in many problems of mechanics, a static (or steady-state) solution is sought after. Kilic [5] addressed this issue by implementing an adaptive dynamic relaxation method based on the work of Underwood [6]. The method is based on introduction of damping to the equations of motion, towards finding the equilibrium state much quicker than solving the problem dynamically. The adaptive dynamic relaxation method used by Kilic [5] leads to calculation of damping coefficients for each material point at each time step.

5.4 Applications

The peridynamic theory has been used for simulation of numerous problems of applied mechanics. A few applications are presented in this section: First, fracture performance of an advanced ceramic layer under impact by a sand particle is discussed. Next, investigation of fracture patterns in glass is presented followed

by anodized aluminum thin film cracking. Finally, preliminary work on three-dimensional fracture observed in polycrystalline materials is shown.

5.4.1 Damage in a Ceramic Layer Due to Small Particle Impact

Materials that transmit Long Wave InfraRed (LWIR) are inherently softer and weaker than their Mid-Wave Infrared (MWIR) counterparts because the price paid for extending transparency to longer wavelengths is a reduction in lattice bond strength. Zinc selenide has excellent transmittance in the LWIR but is too soft for use in mechanically harsh environments, as are the many chalcogenide-based LWIR transmitting glasses. The semiconducting materials, Ge and Si, are mechanically strong and hard but do not transmit into the visible. Diamond has excellent optical and mechanical properties but the cost for diamond optics will likely remain prohibitive for many years. Consequently, most LWIR systems exposed to mechanical stresses use Zinc Sulfide (ZnS) for windows and domes, as it represents the best compromise between strength, hardness, and transparency. Over the last three decades much effort has gone into developing coating materials to improve the modest erosion resistance of ZnS to both rain and sand impacts damage, including sputtered ZnS with an over-layer of Y2O3 (Raytheon's DAR-REP system), amorphous boron phosphide based glasses, diamond-like carbon, and germanium (see [7] for thorough review through 1999). Much of this effort is empirical in nature due to the lack of predictive tools for characterizing impact damage.

To date, most theoretical models of particle impact damage in brittle materials seek to define the stress fields caused by the elastic/plastic nature of an impact event so that crack propagation can be predicted (see [8–12]). Unfortunately, the inherent continuum nature of these models prevents them from properly accounting for crack initiation/introduction. In addition, most of the models focus on characterizing hardness indentation damage as a surrogate for particle impact damage, addressing impact damage largely by generalization, in a semi-quantitative manner. Also, most of the current models are based on an idealized geometry of a single impact event that has difficulty accommodating complexities such as irregularly shaped, frangible, tumbling particles, non-normal incidence, multiple impacts, and the use of protective coatings with variable surface adhesion and internal stresses.

In this subsection, we present an application of the peridynamic theory for accurate quantification of damage/failure in a ceramic layer. A new methodology for extraction of critical stretch parameter and yield strength based on indentation experiments is also given. These material parameters are subsequently used in simulation of sand impact of ZnS.

Vickers indentation testing produces an impression that is square in cross-section with cracks visible on the surface emanating from the corners (see Fig. 5.2).

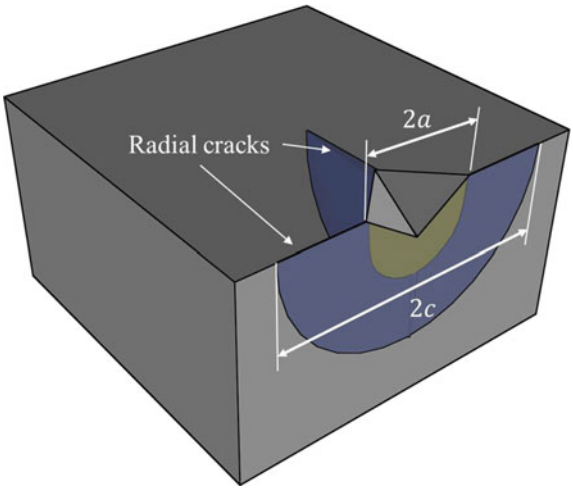

Fig. 5.2 Schematic of Vickers indentation and definition of 2*c*

Table 5.1 Hardness test data for multispectral grade ZnS

Load (N)	Crack length, 2*c* (μm)
0.98	57
0.98	45
1.96	74
1.96	82
2.94	82
2.94	101
2.94	97
4.90	139
4.90	116
4.90	128

The lengths of these cracks, *c*, are measured from the center of the indent out to the crack tip. However, it is common to report crack lengths in terms of 2*c*, the distance from one crack tip through the impression to the tip of the crack emanating from the opposite corner. The measurements in this study were made on multispectral grade ZnS using a Leco model M-400-H1 Micro Hardness Tester. Dead weight loads of 100, 200, 300 and 500 g (equivalent to about 1, 2, 3 and 5 N) were applied using a dashpot-controlled descent taking about 5 s. Dwell time was 10 s. The resulting indents were characterized in the SEM and the values of 2*c* are reported in Table 5.1.

Peridynamic simulations of Vickers indentation tests were performed next. The PD model is comprised of two material regions corresponding to the indenter and the ceramic substrate. The indenter is modeled as the pyramid shape of a standard Vickers hardness apparatus while for the substrate a cylindrical shape having a 120 μm height and radius is considered. Total number of material points in

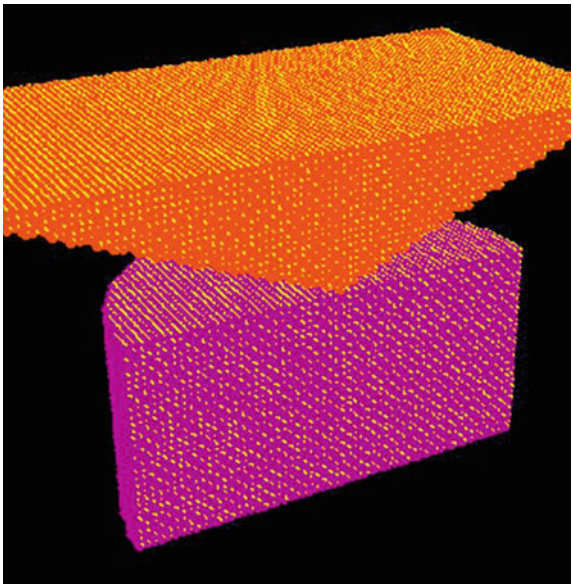

Fig. 5.3 The PD model of the indenter and the ceramic substrate used in Vickers indentation simulations

the PD model including both the indenter and the substrate is 148,550. Figure 5.3 shows the PD models of the indenter and ceramic substrate prior to indentation. The bottom surface of the substrate is constrained from displacement in all directions. The magnitude of the indenter velocity is held constant during the entire simulation. During the "loading" phase, it is directed downward, into the substrate, and during the "unloading" phase it is directed upward or away from the substrate. The constant velocity condition imposed on the indenter implies that no damage occurs in the indenter as it moves down (or up) as a rigid body. Specified velocity conditions combined with specified loading/unloading times dictate the indentation depth for each simulation; this is in contrast to test conditions where maximum force is controlled. A preliminary study was conducted to identify an approximate relationship between the indentation depth versus the maximum force. This relationship is in turn utilized to decide on four distinct cases to simulate with specified indentation depths. The damage developed in the ceramic substrate is analyzed and the crack length $2c$ is measured for each simulated case. An example of the damage pattern visible on the substrate surface is shown in Fig. 5.4. The shape of the high damage region matches the surface cracks observed in the experiments. However, PD simulations do not produce clearly defined crack faces, rather concentration of high damage at sites cracks are expected to form. This is likely due to the PD grid density and that a more refined grid is expected to lead to a better defined fracture; in this study, the damage pattern is used as approximation to fracture. The total force in the vertical direction, F_3 is also calculated in PD simulations. A typical indentation force versus indentation depth response calculated by the PD simulations is shown in Fig. 5.5. Several simulations were performed with a range of values for f_y and s_0 in search of the values which produce the best match of F_3 and

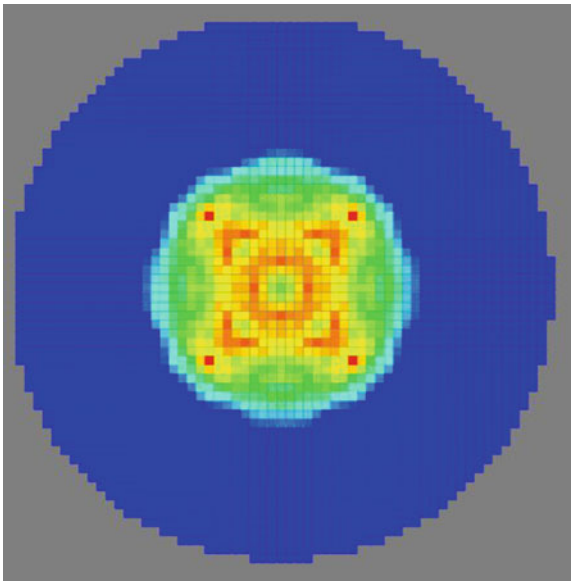

Fig. 5.4 Damage pattern consistent with hardness test fracture visible on the surface of the substrate

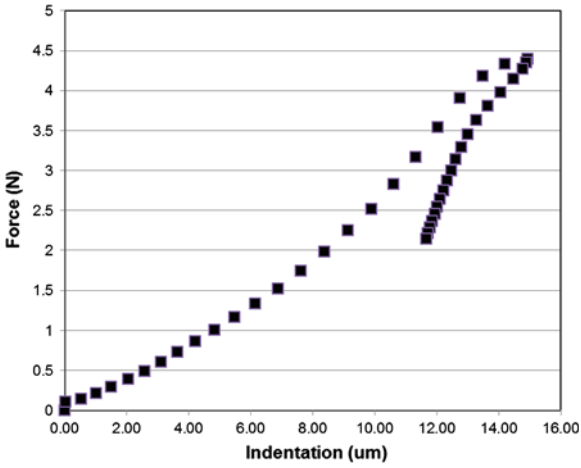

Fig. 5.5 A typical force versus indentation response simulated by PD

$2c$ to the data in Table 5.1. In this study, for ZnS, the optimum material parameters are found to be $\sigma_y = 525$ MPa and $s_0 = 0.0007$. Figure 5.6 shows the good agreement between the PD simulated force versus crack length response of the four cases when the optimal parameters are used and the experimental measurements.

The dynamic fracture of a zinc sulfide layer due to a rigid impactor is considered in order to demonstrate the effectiveness of the use of peridynamic theory for high-fidelity simulation of impact events experienced by electromagnetic windows, The experimental guidelines listed in Harris [7] for testing IR windows and domes under sand impact are taken as the starting point for peridynamic simulation model

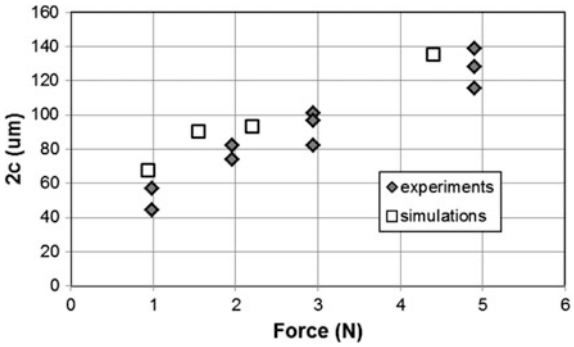

Fig. 5.6 PD simulated force versus crack length response compared against experimental measurements

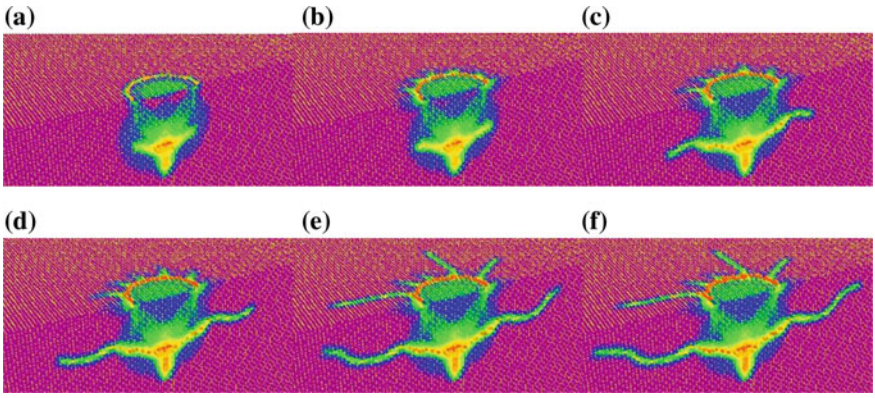

Fig. 5.7 Damage evolution in ZnS substrate at times, **a** 27 ns, **b** 47 ns, **c** 68 ns, **d** 99 ns, **e** 102 ns and **f** 136 ns

parameters. The sand particle is represented by a rigid impactor having a cylindrical geometry with 160 μm diameter and 60 μm height; the mass of the impactor is calculated based on a density of 3300 kg/m^3. The impactor is given an impact velocity of 75 m/s. The ZnS layer is a cylinder with 3 mm diameter and 1.5 mm height. The peridynamic model parameters governing ZnS behavior (deformation and failure) are calculated based on the following properties: density of 4070 kg/m^3, bulk modulus of 87.38 GPa and critical stretch value of 0.0007 as extracted from hardness tests. The model uses a grid spacing of 8 μm, and the time-step size is 0.68 ns.

Figure 5.7 shows the evolution of damage due to sand impact as predicted by the peridynamic model; only half of the solution domain is shown so that the damage through the thickness can be clearly observed. In these plots, the damage varies from zero (purple, no damage) to unity (red, maximum damage). The impactor is

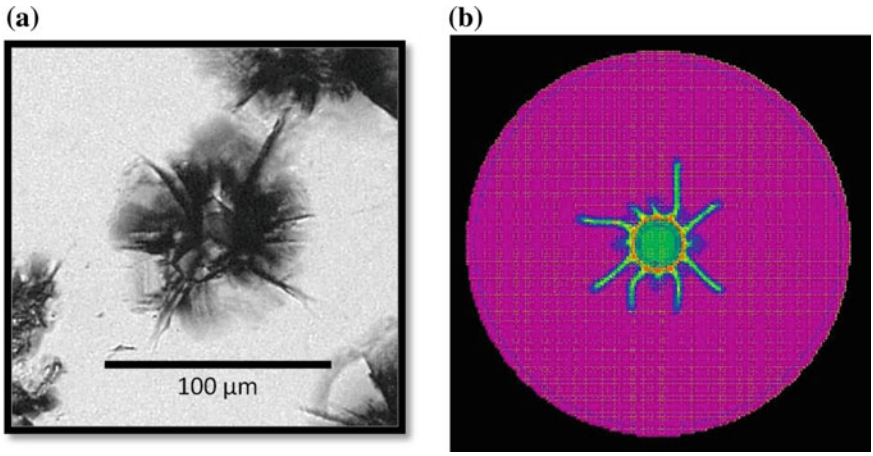

Fig. 5.8 **a** *Top view* of experimentally observed fracture pattern on multi-spectral ZnS caused by impact of 144–177 µm sand particle at velocity of 75 m/s, **b** simulated fracture pattern of ZnS under impact by 160 µm diameter sand particle at velocity of 75 m/s

not show in order to provide a better view of the surfaces. The damage progresses straight down early in the process (Fig. 5.7a, b) with increasing impact load. As unloading from the impact begins, lateral cracks below the surface and shallow radial cracks start to emanate from the center radially (Fig. 5.7c) curving upwards towards the free surface (Fig. 5.7d). Simultaneous with lateral and radial crack initiation, completely damaged material directly below the impactor rebounds upward and separates from the ZnS layer as it forms a conical impact impression, clearly visible in Fig. 5.7f. The crack formation visible on the top surface of the substrate exhibits a combination of short and long radial cracks. The overall fracture pattern from the peridynamic simulation matches well the morphology observed during experimental testing of sand impact damage in ZnS. Such damage is shown in Fig. 5.8a, where the lateral and radial cracks caused by impact of a 149–177 µm sand particle traveling at 75 m/s are clearly evident. For comparison, a top view of the PD simulated crack pattern is shown in Fig. 5.8b, which exhibits striking similarity.

As noted earlier, many researchers have recognized the strong correlation in fracture behavior between particle impact damage and indentation damage. Figure 5.9, taken from Evans [8] documents crack formation and growth in ZnS during the loading and unloading stages of an elastic/plastic spherical indentation test. While the sequence of events, as outlined in Fig. 5.9, do not exactly match those observed from the PD simulation, where radial crack formation lags that of lateral cracks, the similarity in the phenomenology of the two fracture events is striking, despite differences in load conditions. In both cases, radial cracks, lateral cracks, extensive damage below the impact site (median crack in Fig. 5.9) and impact crater formation occurs.

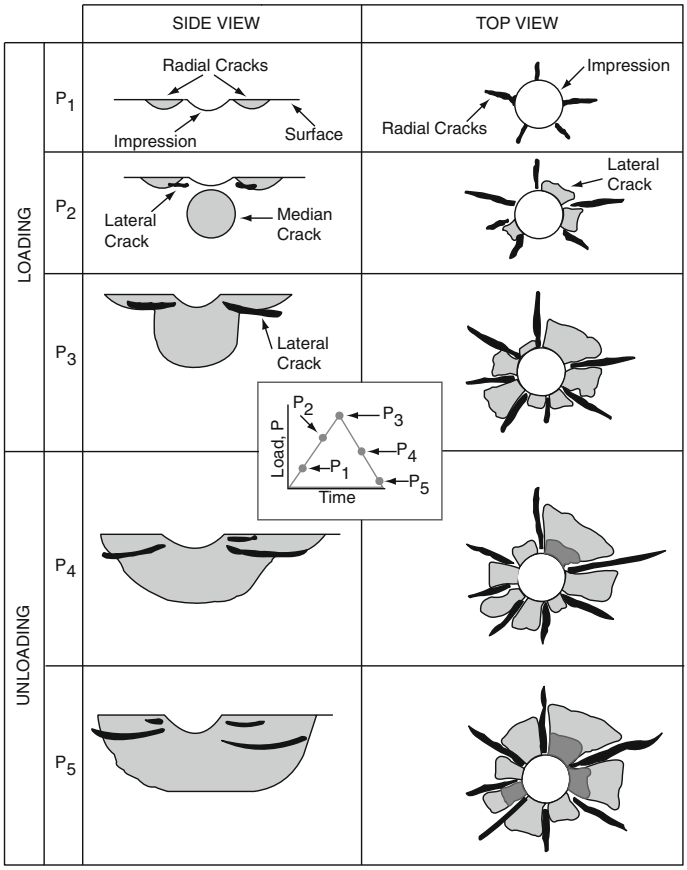

Fig. 5.9 Schematic showing the sequence of crack formation and growth during spherical indentation studies of ZnS (after evans [8])

In order to further highlight PD theory's ability to simulate 3-D, non-planar multiple fracture surfaces, the results are presented in Fig. 5.10, for material points with damage greater than 0.2; i.e., material points with less damage is removed from view. Four different views (left to right: oblique, side, top and front) at five different times (top to bottom: 47, 61, 88, 95 and 102 ns) are shown. The selection of 0.2 as the cut-off damage parameter for these figures ensures viewing of all regions with any significant damage, including localized regions that do not have fully developed cracks.

These results clearly demonstrate the utility of peridynamic theory to model mechanical damage, especially for those situations where continuum mechanics has difficulty capturing the complexities of the evolving stress fields and crack formation. Future work to improve and expand the capabilities of this model include investigation of various coating materials with varying thicknesses, simulations

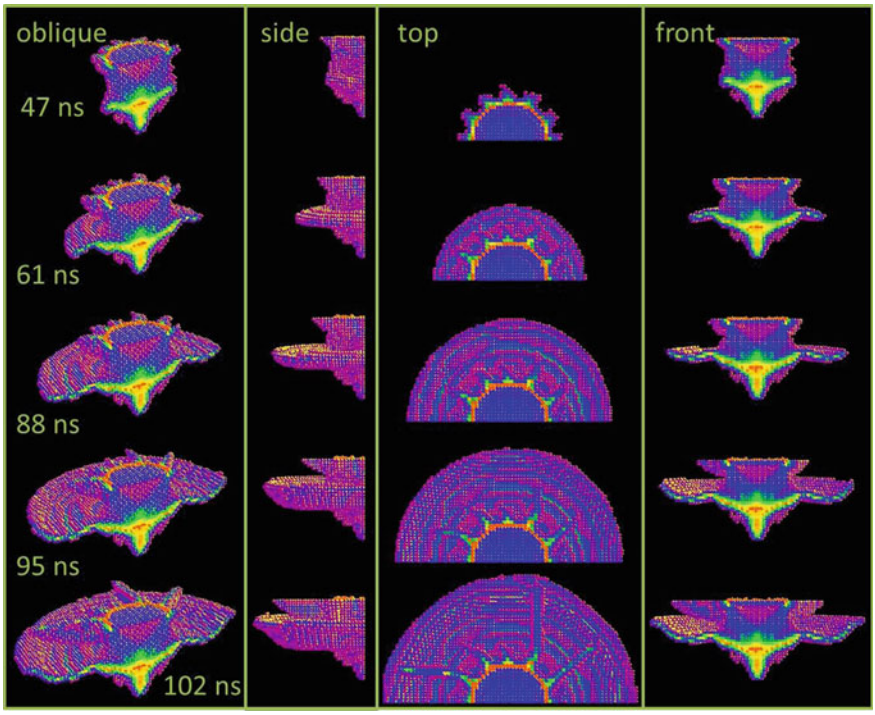

Fig. 5.10 PD simulated damage evolution from different angles with material points with damage less than 0.2 removed from view

involving deformable and friable impactors to better understand the physics of "soft" particle impact on "hard" substrates, and investigation of raindrop impact.

5.4.2 Fracture Patterns in Anodized Aluminum

Plasma etching is an important step of wafer manufacturing in integrated circuit industry. The process aims to precisely engineer the surface features of the wafer. It involves generation of plasma, which requires a sealed chamber. The plasma is exposed to the target (wafer) and any other material surface inside the chamber including the plasma chamber walls. Anodized aluminum is a common plasma chamber and component material in the IC industry; an oxide coating (alumina) is produced on the aluminum surface leading to strong anodic polarization, significantly reducing reactivity. A common problem is the landing of foreign particles on the wafer while it is being etched rendering part of the wafer with zero yielding die. These particles are believed to originate from the plasma chamber components surrounding the wafers. One of the models is mechanical failure in the form of

fracturing of the coating and flaking off. There are two main reasons for the coating failure [13]: (i) the difference between the coefficients of thermal expansion of the coating and the underlying metal, and (ii) change of surface chemistry due to plasma environment. These two mechanisms may or may not work together. Very few studies have been conducted to address the problems associated with integrity of anodized aluminum parts used in plasma chambers; most focus on cleaning and conditioning of the surface [14, 15]. A physics based understanding of the failure mechanisms at play is essential for improving the surface integrity of anodized aluminum parts.

This study aims to investigate the underlying physics of the failure modes observed in anodized aluminum parts used in plasma chambers. Experimental investigation involves mechanical property characterization and clear identification of failure modes examination through SEM. Simulations of aluminum substrate with alumina coating under expansion loading conditions are performed. Computational investigation also explores the potential effect of an additional thin layer between the coating and the substrate.

Ground electrodes encase wafers processed in a plasma chamber. These electrodes are exposed to plasma and corrosive gases used during the etching process. The chamber temperature varies a few tens of degrees from room temperature during the etch process. Commercially available ground electrode after use was cross sectioned for this study. The images were taken under bright field conditions at 15 kV with both the interior and exterior surfaces showing similar sized crazing (Fig. 5.11). The mean crack-free area for the inside surface was 3952 μm^2 while that for the exterior was 3521 μm^2. The thicknesses of exterior and interior coatings were approximately 22 and 19 μm, respectively (Fig. 5.12). EDS mapping of the interior surface showed the same elements common in the top-down spectra but without the *F* signal. No morphological or compositional layering was evident.

The fracture due to coefficient of thermal expansion mismatch between the alumina coating and the aluminum substrate is simulated using the peridynamic theory. Results for the material system with a single alumina layer are presented

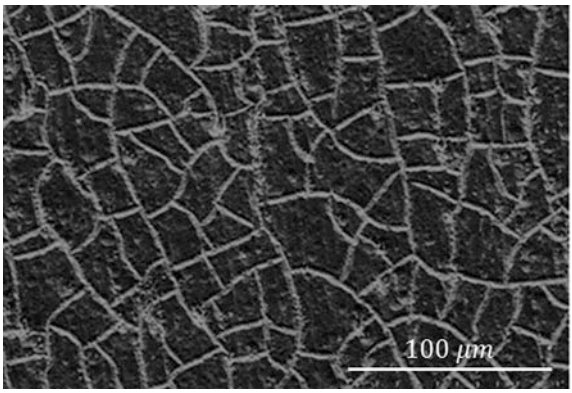

Fig. 5.11 Fracture pattern observed along the inside surface of the plasma chamber

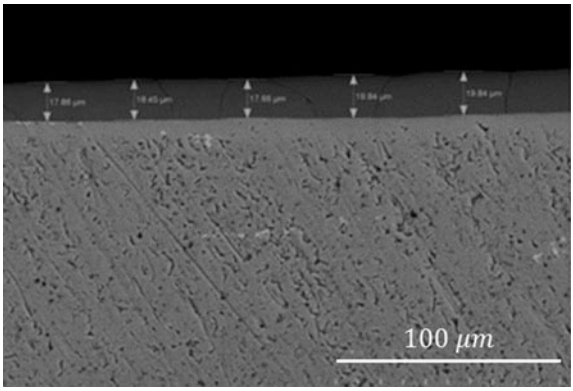

Fig. 5.12 Cross-sectional view of the alumina coating and aluminum substrate. Fracturing of alumina coating is observed

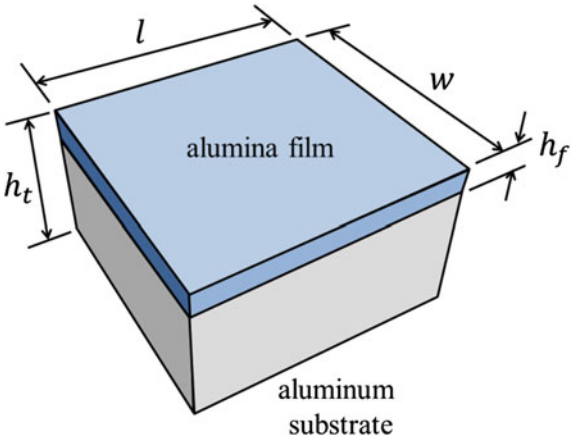

Fig. 5.13 Geometry of the anodized aluminum material system

first, followed by the investigation of effect of adding a compliant layer between the aluminum and the alumina.

The geometry of the material system (two-layer system) is shown in Fig. 5.13. The problem domain is a rectangular prism with length l = 400 μm, width w = 400 μm, and a total height of h_t = 220 μm. The thickness of the alumina film is taken as h_f = 20 μm. The grid spacing in the model is taken as 3.33 μm leading to approximately 950,000 grid points.

The material properties for the aluminum substrate are: elastic modulus of 68 GPa and density of 2,700 kg/m^3. Similarly, for the alumina film, the elastic modulus is taken as 370 GPa while its density is kg/m^3. The critical stretch for the thin alumina coating is calculated to be around 5 % while aluminum is about 20 % owing to its ductile nature.

In order to simulate the deformation field due to the mismatch between coefficients of thermal expansion values of the constituent materials, a large portion of the substrate in the depth direction (80 %) was subjected to isotropic expansion in

the planar directions (length and width as defined earlier). The expansion is then transferred to the film through deformation. The loading is applied in a ramped fashion gradually so as to prevent premature fracture near the boundary regions.

Figure 5.14 shows the top view of the damage progression along the surface of the coating. Four different time steps are shown, with time increasing from top to bottom. Fracture starts at one edge and propagates as the expansion of the substrate continues. Due to the nature of the isotropic extension, a number of branches emerge leading to the final fracture configuration, shown in the bottom segment of Fig. 5.14. However, there is further fracture beneath the surface that is of interest. In order to examine the fracture morphology inside the coating, damage contours are plotted such that only those material points with damage values ranging between 0.2 and 1.0 are shown in Fig. 5.15 from top view. Same set of results are shown from an oblique angle in Fig. 5.16 in order to clarify the extent and geometric distribution of the damage in 3-D. Examination of Figs. 5.15 and 5.16 and their comparison to Fig. 5.14 reveal that a considerable damage is accumulated along the interface between the aluminum substrate and the alumina coating. This is consistent with the failure mode observed in experiments, where part of the crazed coating peels/flakes off, suggesting interface delamination. Peridynamic simulations appear to capture the correct failure modes observed in experimental setup.

Additionally, the top surface fracture pattern predicted by peridynamic theory is compared to the micrographs of the crazed coating inside the plasma chamber in Fig. 5.17. A scale bar of 300 μm is included; it applies to both the SEM micrographs (left) and the peridynamic results (right). The size and shape of the surface fracture predicted by peridynamic theory closely resemble those observed experimentally. The "cellular" characteristic of the fracture is prominent on both sides. Peridynamic theory is able to predict the shape and size of the fractured-coating pieces with satisfactory accuracy.

Modification of the anodization process may allow a layered coating system with intention of designing a layer sequence to alleviate or eliminate cracking in the exposed surfaces. The anodized coatings are inherently porous; depending on the durations and concentrations of acid bath process, different porosity values are attained. Therefore, there may be configurations of layered anodized coatings that might lead to better fracture performance of the chamber interior coating. With this hypothesis, the same problem is considered with an added thin layer between the alumina coating and the aluminum substrate (leasing to a three-material system) as sketched in Fig. 5.18. This intermediate layer has the same thickness as the alumina coating but with different material properties. The relationship between the mechanical properties and porosity is well-documented for alumina [16–18]. As the porosity increases, elastic modulus decreases, which makes the material more compliant allowing the material to deform more before fracturing. By making the intermediate layer compliant, the effect of the mismatch between the substrate and the original coating will be lessened, leading to less cracks and flaking off. This added layer will serve as a buffer layer.

Therefore, in this hypothetical test case, the intermediate layer elastic modulus was decreased to half of the original value while the critical stretch was increased to

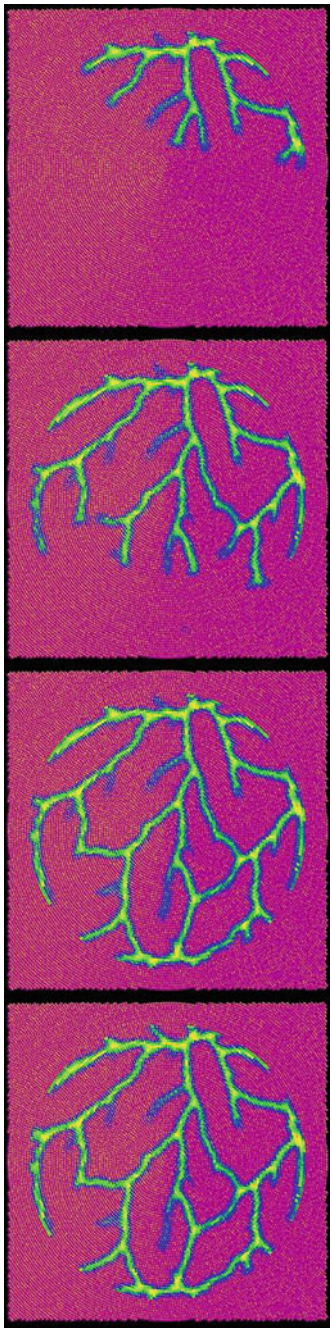

Fig. 5.14 Damage progression along the top surface of the alumina layer predicted by peridynamic theory for the two-layer system

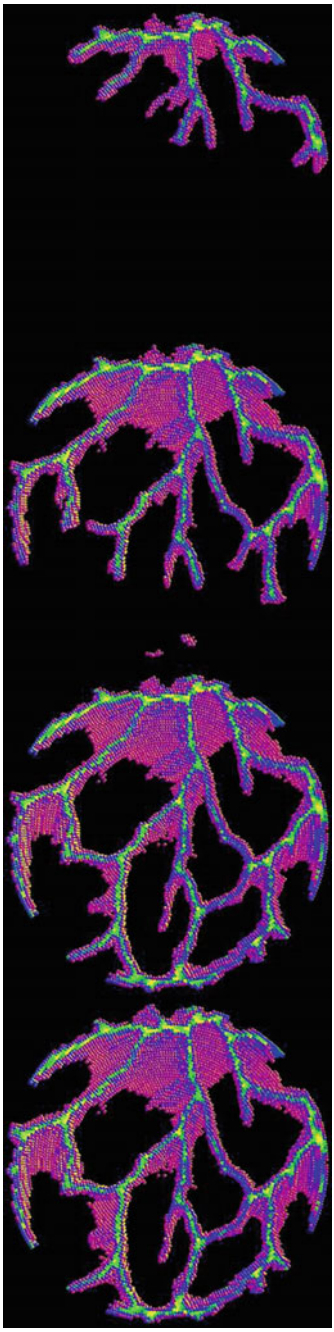

Fig. 5.15 Damage progression in the alumina layer through the thickness from top view predicted by peridynamic theory for the two-layer system

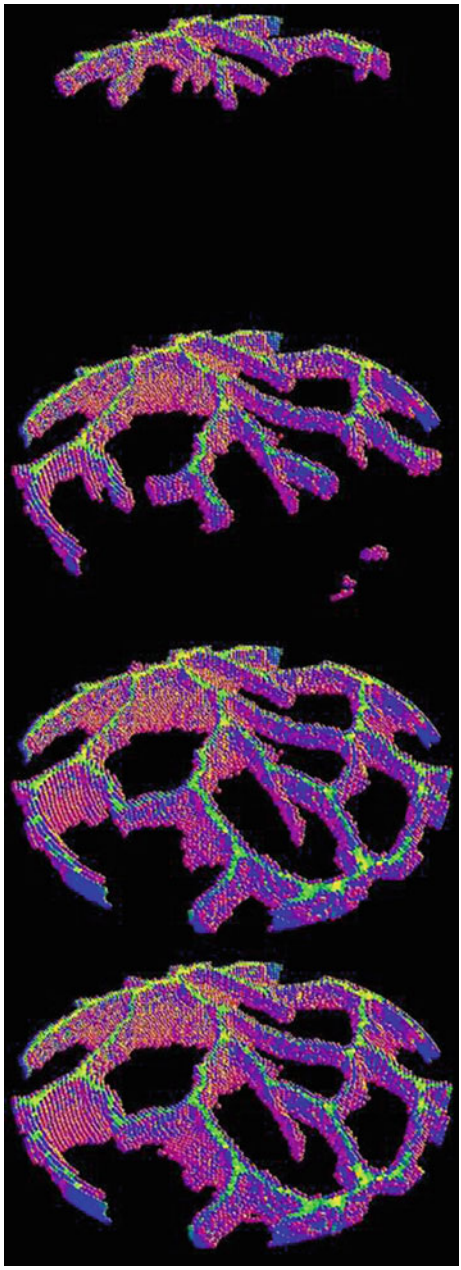

Fig. 5.16 Damage progression in the alumina layer through the thickness from oblique view predicted by peridynamic theory for the two-layer system

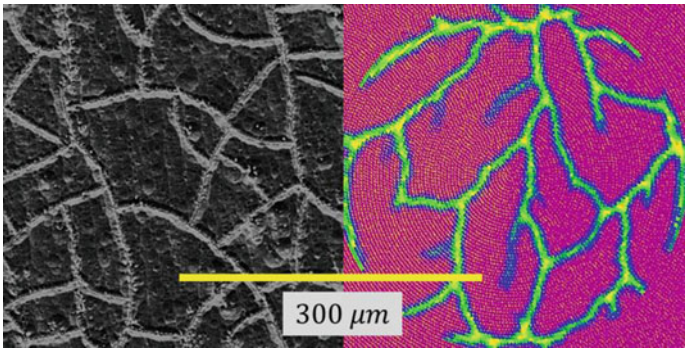

Fig. 5.17 Comparison of experimentally observed fracture patterns (*left*) against those predicted by peridynamic theory for the two-layer system

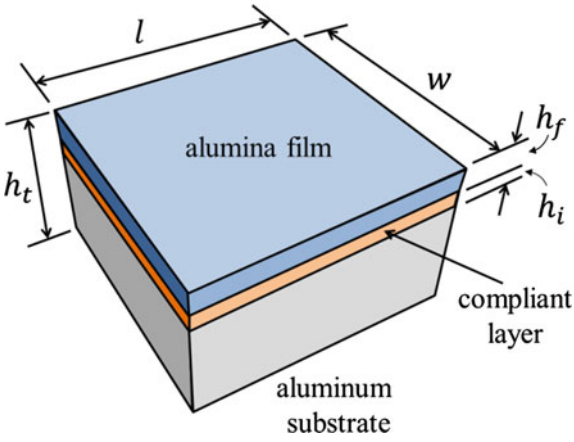

Fig. 5.18 Geometry of the anodized aluminum material system with the compliant intermediate layer

about 7.5 % (compared to 5 %). The remaining parameters of the model is identical to the previous problem.

The damage progression along the top surface of the exposed coating is shown in Fig. 5.19. The amount and distribution of the damage in the three-material system are less than those of the two-layer system. Further, similar to the previous case, the damage contours for the current configuration are plotted for values between 0.2 and 1.0 in order to examine the damages in the thickness direction as shown in Fig. 5.20. In order to make the comparison easier, side-by-side comparison of damage patterns for the two-layer and three-layer systems are shown in Fig. 5.21. The two-layer system results are shown in the left column while the right column is shows the three-layer system damages. Top row in Fig. 5.21 shows only the top surface while bottom row shows damages greater than 0.2 through the thickness. It is clear that the interface delamination problem is significantly reduced; the predicted cracks have vertical faces. This prediction suggests that the

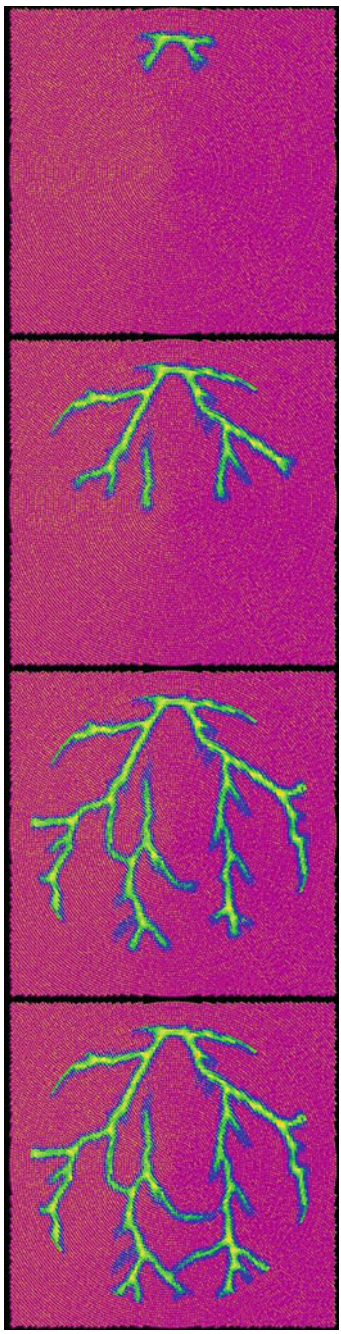

Fig. 5.19 Damage progression along the top surface of the alumina layer predicted by peridynamic theory for the three-layer system

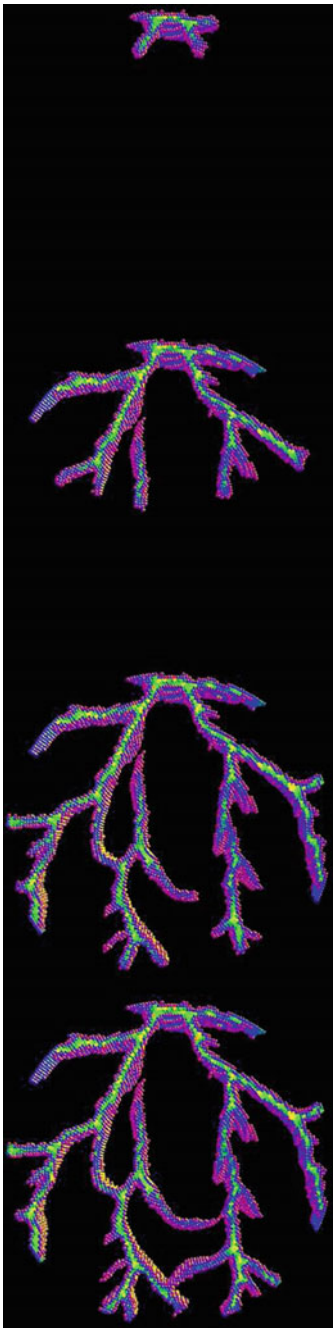

Fig. 5.20 Damage progression in the alumina layer through the thickness from top view predicted by peridynamic theory for the three-layer system

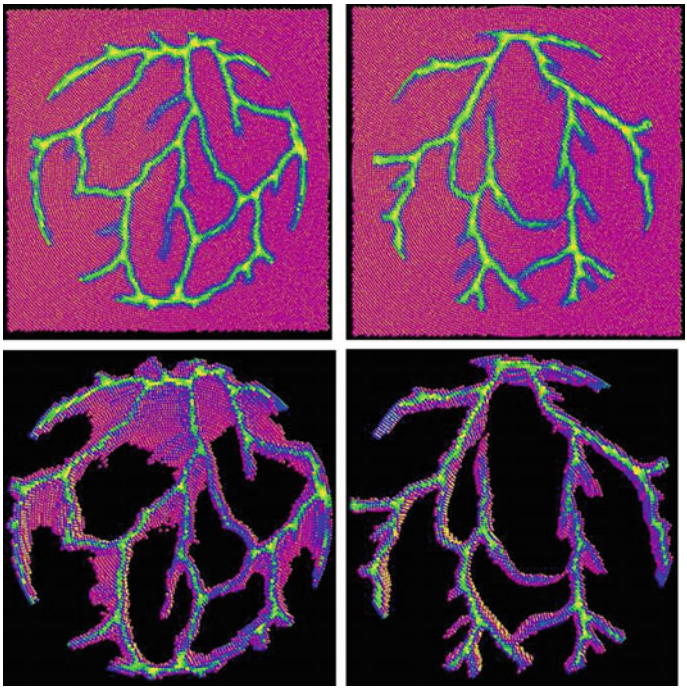

Fig. 5.21 Side-by-side comparison of damage patterns for the two-layer (*left*) and three-layer (*right*) systems. *Top row* shows only the top surface while *bottom row* shows damages greater than 0.2 through the thickness

peeling/flaking off phenomena could be significantly reduced or eliminated by carefully engineering an anodization process leading to a desired layered system.

In this study, a common problem in IC manufacturing, which involves peeling/flaking off of fractured coating of anodized aluminum plasma chamber walls, is described. A computational approach, peridynamic theory, is used to simulate the surface fracturing due to uniform expansion is demonstrated. It was shown that the peridynamic theory captures the correct failure modes observed in experiments. Also, it is able to predict the shape and size of the fractured-coating pieces with satisfactory accuracy.

In order to reduce or eliminate the peeling/flaking off problem, incorporation of an intermediate layer between the coating and substrate is considered. Peridynamic simulation of this hypothetical structure under the identical expansion conditions suggests that a compliant intermediate layer between the coating and substrate has potential to reduce the problem through eliminating interface fracture (delamination) and reducing the surface fracture.

Further study would explore the effects a range of elastic properties and critical stretch values as well as geometry parameters (e.g. thickness).

5.4.3 Dynamic Fracture of Glass

Glass has been used as engineering material over a long period of time. Due to their amorphous micro/nano structure, their behavior under mechanical loading has not been fully understood. Currently, glass materials are ubiquitous in electronic, automotive and aerospace industries; methods with high fidelity fracture predictions would pave the way to components with increased reliability and durability.

In this study, we demonstrate use of peridynamic theory to simulate mechanical response of brittle glass materials under impact loading. The material system involve polymethyl methacrylate (PMMA) plates of 18 mm radius with different thicknesses: 0.5, 1.0, and 3.0 mm. The plates are impacted by a steel ball of radius 1.8 mm and a mass of 16 g. The experimental results are taken from [19]. The impact velocity varies between 10 and 120 m/s.

Peridynamic models for the PMMA plates were generated to have approximately 2.2, 2.6 and 3.2 million material points for thicknesses 0.5, 1.0 and 3.0 mm, respectively. Elastic moduli of 2.42 and 180 GPa were used for PMMA and stainless steel, respectively. In all cases, the peridynamic grid was generated using arbitrarily oriented concentric rings layered orthogonally to the direction of impact in order to eliminate the potential bias in fracture morphology that might be introduced by using a regular rectangular grid. Figure 5.22 demonstrates such a bias introduced when using a rectangular grid (Fig. 5.22a), and a more realistic fracture morphology when circular grid is used (Fig. 5.22b) for a 1.0 mm plate impacted by the projectile with velocity of 80 m/s.

The experimental observations focused on correlating the number of cracks and impact velocity [19]. The cracks counted mostly consisted of radial cracks emanating from the impact site, but the circumferential cracks were also counted when present. For example, Fig. 5.23 shows the progression of the fracture pattern predicted by the peridynamic simulation for a 1 mm thick plate impacted at 20 m/s.

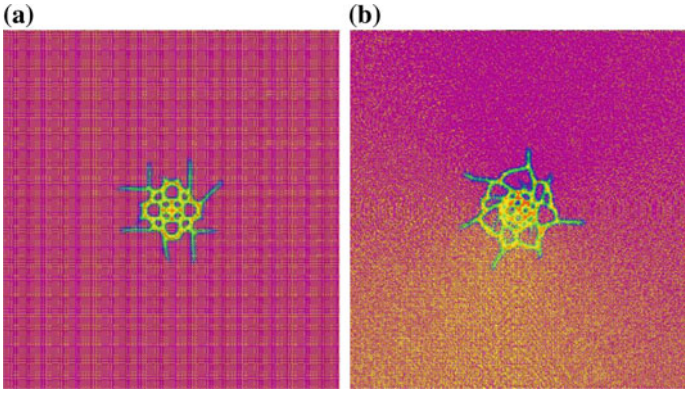

Fig. 5.22 Comparison of fracture morphologies in a 1.0 mm thick plate impacted at 80 m/s, modeled with **a** rectangular grid, and **b** circular grid

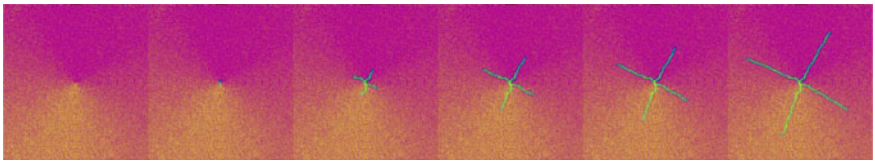

Fig. 5.23 Time progression of fracture for a 1 mm thick plate impacted at 20 m/s

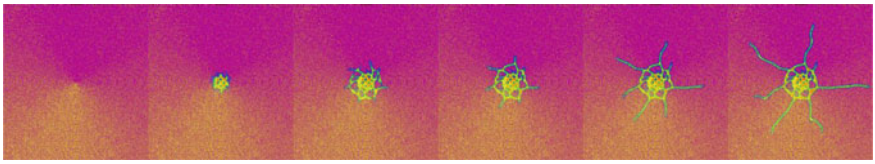

Fig. 5.24 Time progression of fracture for a 1 mm thick plate impacted at 80 m/s

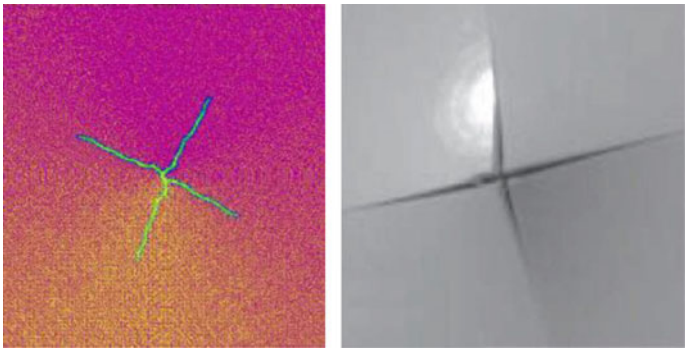

Fig. 5.25 Comparison of peridynamic simulation fracture morphology against experimental observations for a plate with a thickness of 1 mm impacted at 20 m/s

Four radial cracks emanate from the impact point and continue to grow as time progresses. As the velocity of the impactor increases, the number of radial cracks that form increases. After a critical velocity is reached, circumferential cracks form allowing the impactor to break through the plate. Figure 5.24 shows formation of a circumferential crack for a 1 mm plate is impacted at 80 m/s.

Comparisons to the experimental work was done both qualitatively and quantitatively. Figures 5.25 and 5.26 compare the peridynamic simulations performed at 20 and 60 m/s impact velocities against the experiments for 1 mm thick PMMA plate with impact velocities measured at 22.2 and 66.2 m/s, respectively. In the case of 22.2 m/s impact velocity, in experiments four radial cracks are observed with no circumferential cracking. The peridynamic simulation with 20 m/s captures the morphology and the number of radial cracks exactly (Fig. 5.25). Similarly, the comparison of higher velocity impact response is shown in Fig. 5.26; it is worth

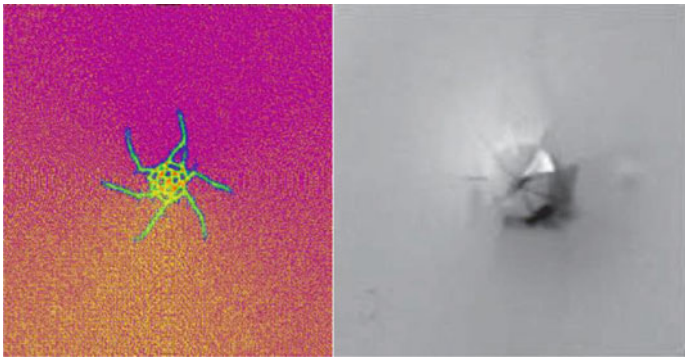

Fig. 5.26 Comparison of peridynamic simulation fracture morphology against experimental observations for a plate with a thickness of 1 mm impacted at 60 m/s

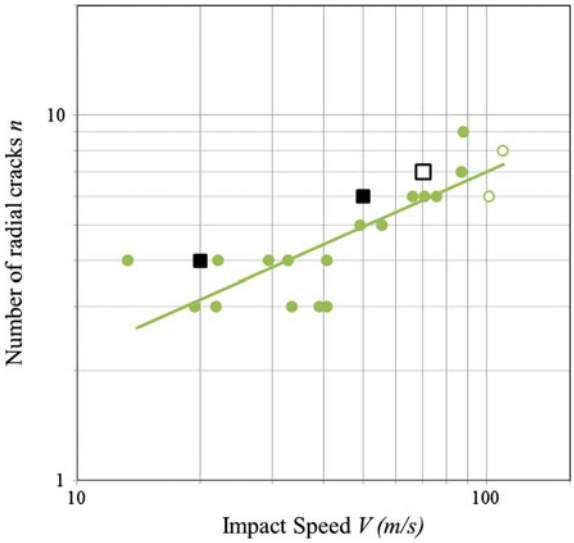

Fig. 5.27 Quantitative comparison of number of radial cracks observed experimentally against captured through peridynamic simulations for a plate with a thickness of 0.5 mm impacted at various velocities

noting that peridynamic simulation accurately captures both the radial and circumferential cracks.

The quantitative study involved 10 distinct peridynamic simulations involving the three different plate thicknesses and impact velocities ranging from 20 to 80 m/s. Results of the peridynamic simulations are compared against the experimental results in Figs. 5.27 through 5.29. In these figures, the circular symbols are the experimental observations while the square symbols are the peridynamic simulation results. In both experimental and peridynamic results presentations, solid symbols indicate that only radial cracks formed and open symbols indicate that circumferential cracks were present. The lines in the plots are least-squares linear fit of the experimental observations to guide the eye. In all three thickness cases

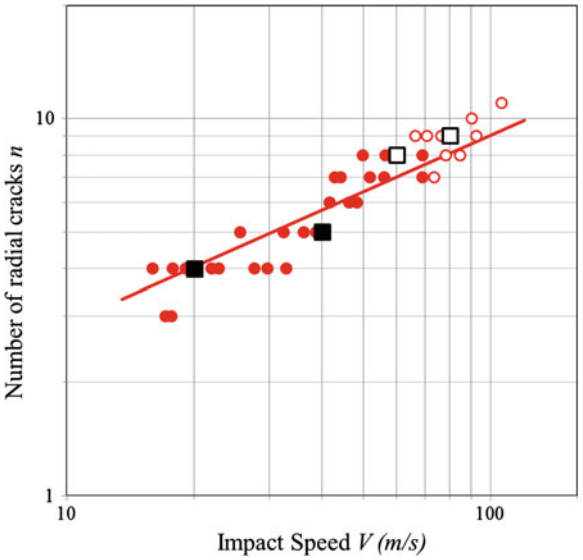

Fig. 5.28 Quantitative comparison of number of radial cracks observed experimentally against captured through peridynamic simulations for a plate with a thickness of 1.0 mm impacted at various velocities

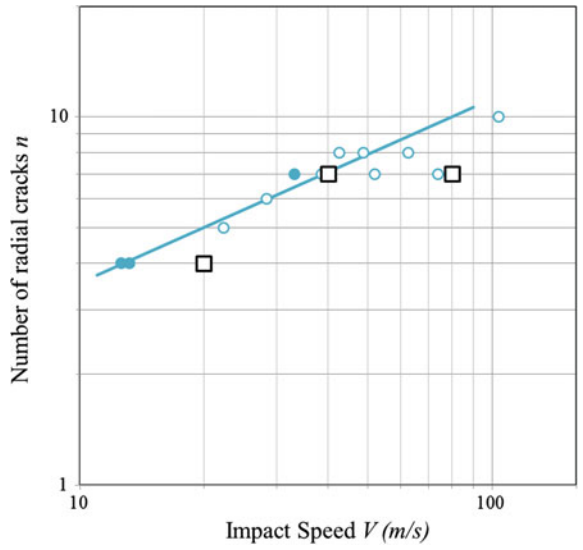

Fig. 5.29 Quantitative comparison of number of radial cracks observed experimentally against captured through peridynamic simulations for a plate with a thickness of 3.0 mm impacted at various velocities

peridynamic predictions of number of radial cracks are close to experimentally observed ones.

The qualitative (Figs. 5.25 and 5.26) and quantitative (Figs. 5.27, 5.28 and 5.29) comparisons clearly illustrate that peridynamic theory is able to capture specific fracture morphology (radial vs. circumferential) as well as the number of radial cracks under dynamic fracture conditions due to impact in PMMA.

References

1. S.A. Silling, M. Epton, O. Weckner, J. Xu, A. Askari, J. Elast. **88**, 151 (2007)
2. E. Madenci, E. Oterkus, *Peridynamic Theory and its Applications* (Springer, New York, 2014)
3. S.A. Silling, A. Askari, Comput. Struct. **83**, 1526 (2005)
4. S.A. Silling, J. Mech. Phys. Solids **48**, 175 (2000)
5. B. Kilic, Peridynamic theory for progressive failure prediction in homogeneous and heterogeneous materials, PhD Dissertation, University of Arizona, 2008
6. Underwood, *Philip in Computational Methods for Transient Analysis* (*A 84-29160 12-64*). (Amsterdam, North-Holland, 1983), pp. 245–265
7. D.C. Harris, *SPIE—the International Society for Optical Engineering*, (1999). ISBN: 0-8194-3482-5
8. A. Evans, T. Wilshaw, Acta Metall. **24**, 939 (1976)
9. A.G. Evans, T.R. Wilshaw, J. Mater. Sci. **12**, 97 (1977)
10. B.R. Lawn, A.G. Evans, D.B. Marshall, J. Am. Ceram. Soc. **63**, 574 (1980)
11. D.B. Marshall, B.R. Lawn, A.G. Evans, J. Am. Ceram. Soc. **65**, 561 (1982)
12. Xi Chen, J.W. Hutchinson, A.G. Evans, J. Am. Ceram. Soc. **88**, 1233 (2005)
13. S. Ponnekanti, J. Vac. Sci. Technol. A Vacuum Surf. Film **14**, 1127 (1996)
14. G. Cunge, B. Pelissier, O. Joubert, R. Ramos, C. Maurice, Plasma Sources Sci. Technol. **14**, 599 (2005)
15. R. Ramos, G. Cunge, B. Pelissier, O. Joubert, Plasma Sources Sci. Technol. **16**, 711 (2007)
16. M. Asmani, C. Kermel, A. Leriche, M. Ourak, J. Eur. Ceram. Soc. **21**, 1081 (2001)
17. W. Pabst, E. Gregorov, G. Tich, J. Eur. Ceram. Soc. **26**, 1085 (2006)
18. J. Kovacik, J. Mater. Sci. Lett. **18**, 1007 (1999)
19. N. Vandenberghe, R. Vermorel, E. Villermaux, Phys. Rev. Lett. **110**, 174302 (2013)

Chapter 6
Consistent Embedding Frameworks for Predictive Multi-theory Multiscale Simulations

Krishna Muralidharan, Keith Runge and Pierre A. Deymier

Abstract In this chapter, we present the predictive multi-theory multiscale framework called consistent embedding. The characteristics of a consistently embedded set of theories is defined and illustrated by four examples. These examples show the coupling between quantum mechanical and atomistic scales, short and long time microstructural evolution, and chemical reactivity and flow dynamics. A final example carries a cautionary message in that all embedding are not necessarily consistent. Overlapping tests are essential to the establishment of a predictive multi-theory multiscale simulation.

6.1 Introduction

The ability to predict and characterize the complex structure-property relations of condensed matter embodies the central objective of the field of computational materials science and engineering. Material properties are collectively governed by phenomena that occur over a broad span of temporal and spatial scales, and an accurate predictive description of the underlying processes cannot be achieved by solely relying on traditional materials modeling methods. Such methods include quantum mechanics based wave-function techniques or density functional theory (DFT), Newtonian mechanics based atomistic molecular dynamics (MD) and finite element methods (FEM), as well as statistical mechanical phase-field models (PF). Each of the above-listed methods address distinct spatial and temporal scales

K. Muralidharan (✉) · K. Runge · P.A. Deymier
Department of Materials Science and Engineering, University of Arizona, Tucson, AZ 85721, USA
e-mail: krishna@u.arizona.edu

K. Runge
e-mail: krunge@u.arizona.edu

P.A. Deymier
e-mail: deymier@u.arizona.edu

P.A. Deymier et al. (eds.), *Multiscale Paradigms in Integrated Computational Materials Science and Engineering*, Springer Series in Materials Science 226,
DOI 10.1007/978-3-319-24529-4_6

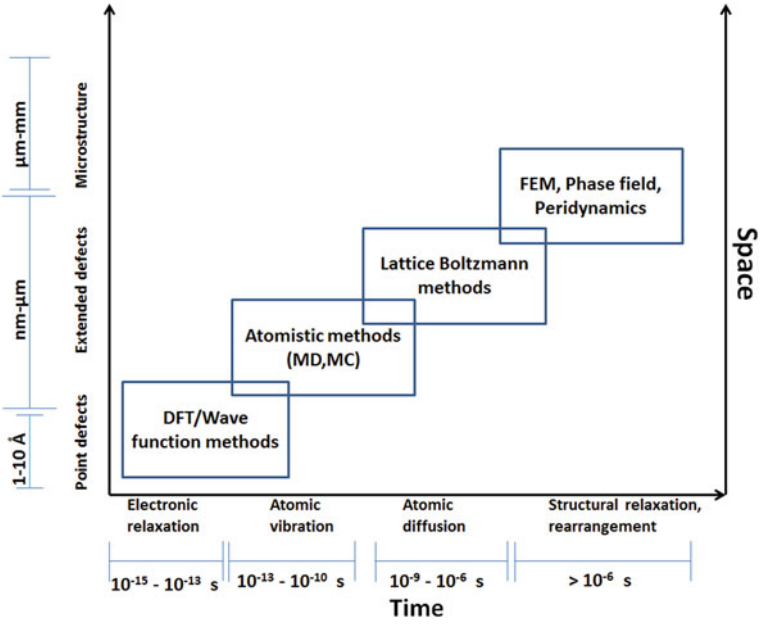

Fig. 6.1 The temporal and spatial domains corresponding to the traditional materials modeling tools

(see Fig. 6.1), and importantly, represent models based on different theories. This necessitates the implementation of multi-theory, multiscale computational paradigms for which the following are required: (i) compatibility between the employed scale-specific and theory-specific models as well as (ii) seamless exchange of information across scales. Compatibility between the scale-specific methods can be ensured, if and only if the overarching multiscale paradigm conforms to the principles of consistent embedding. Specifically, consistent embedding requires that the parameters for a coarser-scale model be derived exclusively from an underlying finer-scale method that corresponds to a more rigorous level of theory. An equally important requirement is to ensure that the coarser-scale method emulates the behavior of the finer-scale model, albeit in an average fashion, at the spatial and temporal scales accessible to the finer-scale model.

Multiscale models can be classified as serial or concurrent. In serial methods, a set of calculations at a fundamental level (fine length scale) is used for evaluating constitutive relations for coarser-scale models that describe a phenomenon of interest at longer length scales. For example, atomistic simulations can be used to obtain the constitutive behavior of finite elements, which are then used to simulate larger scale problems [1, 2]. Other examples include the coarse-grained molecular dynamics methods (CGMD) [3] as well as the parameterization of interatomic potentials based on ab initio/first principles data as demonstrated in Chap. 3. Concurrent methods represent dynamical extensions of serial multiscaling and rely

on real-time coupling of different computational methodologies applied to different spatial regions of the material [4, 5]. A relevant example is the problem of crack propagation that was tackled early on by multiscale methods. Atomistic simulation techniques were used to model the crack tip where large deformations (even bond breakage) occur and continuum approaches (finite element methods) were used to model the region far away from the crack tip. Broughton and co-workers [5] proposed an algorithm involving hand shaking between Finite Elements (FE), molecular dynamics (MD) and simplified quantum chemical Tight-Binding (TB) regions. This algorithm was able to dynamically track a crack propagating through silicon. The handshaking between the MD and FE regions was achieved by drawing an imaginary surface between them. Within the range of the MD interatomic potential from this surface, FE mesh points were located at equilibrium atomic sites. Moving away from this overlapping region into the FE region, the mesh spacing was made larger. Any FE element that crosses the interface contributed half its weight to a conservative Hamiltonian. Similarly any MD interaction that crossed the interface also contributed half its weight. Kohlhoff et al. [6] introduced a similar transition region between the atomic and continuous regions. They also scaled down the finite element size to the atomic scale in this transition region. Unlike Broughton et al.'s work, the interface was of finite size and not sharp. Abraham et al. [7] combined the above two techniques by constructing an explicit Hamiltonian for the atoms and the FE nodes in the transition region by weighing their contributions with respect to their distance away from the middle of the interface. Ogata and co-workers [8] used a similar algorithm to study chemical reactions and their interplay with mechanical phenomena in materials, such as in the oxidation of Si(111) surface.

The principles of consistent embedding represent a necessary condition for robust predictive multiscale modeling, but does not guarantee seamless transfer of information across scales. Specifically, in serial multiscaling, upscaling of information results in homogenization of the underlying finer-scale data and consequently, fluctuations inherent to the finer-scales cannot be captured. In concurrent multiscale paradigms, the utilization of different modeling techniques to represent different spatial segments of the material gives rise to 'interfaces' that are merely an artifact of the implementation method. In the case of multiscale simulations that examined crack-propagation as described above, interfaces are a consequence of 'elastic impedance-mismatches' between the chosen methods, resulting in spurious scattering of stress waves that emanate from the crack-tip. Thus, steps to minimize possible systematic errors that arise due to the artificial interfaces within the multiscale implementation are necessary. Further, from a computational cost point of view, there is a need for the judicious use of the finer-scale methods without compromising the accuracy of the concurrent multiscale simulation framework. In this context, a discussion of available methods that facilitate the seamless coupling of information derived from scale-specific methods, while limiting the use of computationally expensive methods without loss in accuracy, will serve as the basis for this chapter. In particular, using relevant examples, we illustrate multiscale methods that adhere to the consistent embedding philosophy and when necessary

invoke wavelet based scale-decomposition methods to glean spatial and temporal information at all relevant scales, in order to develop comprehensive, accurate representations of multiscale materials phenomena.

As a first step towards characterizing multiscale simulations involving coupling between atomistic and continuum methods, we examine elastic wave propagation in such 'coupled' media.

6.2 Example 1: The Consistent Embedding Framework for Quantum-Classical Coupling

In this section, we demonstrate the successful implementation of a concurrent multiscale methodology that couples classical atomistic simulations with quantum mechanical methods, in tune with the consistent embedding principles. In particular, the choice of quantum chemical theory will be made subject to consistent embedding constraints, where higher level quantum chemical theory will be used to train less computationally demanding semi-empirical quantum chemical forms as well as classical interatomic potentials. Here it is important to note that the form of the quantum chemical Hamiltonian used is known as 'semi-empirical', but our training will be based solely on computed results from correlated calculations. The successful development of the semi-empirical Hamiltonian or the Transfer Hamiltonian (TH) and the classical potentials will enable a concurrent multiscale methodology capable of modeling the mechanical deformation of silica-based systems. The interface between TH modeled region and the classical region will be handled by 'pseudo atoms' that are appropriately trained to enable a seamless coupling between the regions.

6.2.1 Transfer Hamiltonian

For our higher level of quantum chemical theory we choose a method that includes the effects of electron correlation, known as coupled cluster theory including single and double excitations (CCSD). This highly accurate level of quantum chemical theory is also, of necessity, very computationally demanding. Hence it is necessary to choose a training molecule which exhibits the chemical bonding characteristic of silica, but is limited to a relatively small number of atoms. We choose pyrosilicic acid ($H_6Si_2O_7$) to create a CCSD training set for the 'semi-empirical' Hamiltonian. As seen in Fig. 6.2 [9], the Si–O bond length is varied through compressions and stretches to generate a training set for the Transfer Hamiltonian. In this case, we have chosen to use a neglect of diatomic differential overlap (NDDO) Hamiltonian as our less computationally demanding quantum chemical model. NDDO is one of a set of approximations collectively referred to as zero differential overlap methods [10]. When computers were much less powerful than they are today, these methods

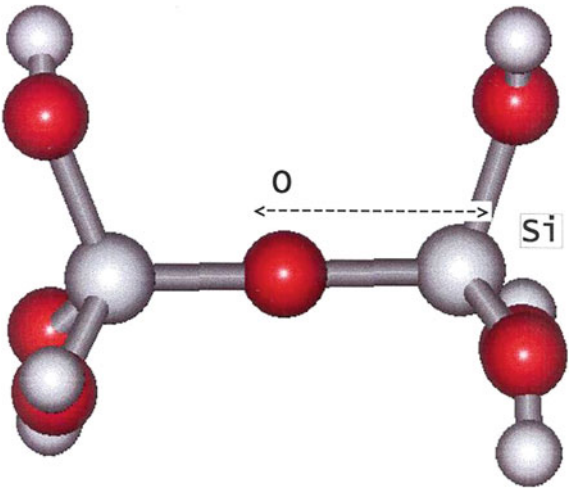

Fig. 6.2 Pyrosilicic acid is used to create a training set of forces from CCSD for the Transfer Hamiltonian

were developed to be computationally tractable and used mathematical forms derived from theory. These forms were parameterized to reproduce certain empirical data, e.g. heats of fusion, for simple molecules and the resulting parameterized Hamiltonians were applied to more complex problems with some success. The empirical parameterization of theoretically derived forms came to be known as 'semi-empirical' theory. By choosing to parameterize the NDDO Hamiltonian based on high accuracy, ab initio quantum chemistry, CCSD, we remove the empirical information from the procedure and replace it with a more detailed theoretical model. This substitution of theoretical for empirical information characterizes one possible implementation of the consistent embedding framework for the development of predictive modeling. The training of the Transfer Hamiltonian is accomplished using genetic algorithms, which are tuned to reproduce CCSD forces for the training set of molecular geometries. The choice of training on forces is motivated by our interest in stress-strain relations. In the next section, the implications of this Transfer Hamiltonian will be examined in a somewhat larger system.

6.2.2 Classical Interatomic Potential

The proposed method for constructing a classical pair potential for use in multiscale modeling has several components: (1) it should be constructed for accuracy of the specific properties to be studied, (2) it should be "trained" on quantum data generated in the same way as for the quantum domain at its interface (i.e. the TH), (3) it should predict accurate equilibrium structure, but include training on appropriate near equilibrium states as well, and (4) simplicity of form for parameterization and implementation in MD codes should be maintained. The steps in constructing a potential are the following. First the specific quantum method to be used in the

multiscale modeling (NDDO) is applied to a large cluster or representative sample of the solid to be modeled. The forces on atoms for both equilibrium and near equilibrium states are then calculated quantum mechanically. Next, a simple functional form for the pair potentials with the correct physical shape (e.g., one of the existing phenomenological forms) is chosen and the parameters controlling that shape selected for optimization. The forces on the ions are calculated for these chosen potentials at the configurations used in the quantum force calculations, and compared with those quantum forces. The parameters are adjusted for a good fit to the quantum force data. When a good fit has been obtained, the potential energy at equilibrium is tested for stability using a gradient algorithm to establish that a local minimum of the potential has been obtained. Finally, the property of interest (e.g., elastic response) is tested by comparing its calculation using MD simulation with the fitted potentials and that with the quantum forces. If necessary, the fitting procedure can be repeated with differing weights for the quantum force data in equilibrium and near equilibrium states, or other additional input from the quantum calculations.

Following the above enumerated strategy, an interatomic potential for silica with a functional form as given in (6.1) is developed (the parameter-set is available in [11]). The training set for this potential consists of using force data corresponding to small strain configurations as well as the equilibrium configuration of a silica nanorod as shown in Fig. 6.3 [11]. Fitting a set of the ten potential parameters, constrained by charge neutrality, from hundreds of data points requires an optimization approach that is capable of exploring a large parameter space efficiently without being trapped in local minima. For this task, a genetic algorithm (GA) in conjunction with standard geometry optimization techniques is used as discussed in [11].

$$U = \sum_{i<j} V_{ij}(r); V_{ij}(r) = \frac{q_i q_j}{r_{ij}} + a_{ij}\exp\left(-b_{ij} r_{ij}\right) - \frac{c_{ij}}{r_{ij}^6} \tag{6.1}$$

6.2.3 Pseudo Atoms

The Link Atom method (LA) is a commonly used method to spatially couple ‘quantum mechanics (QM)’ and ‘classical mechanics (CM)’ regions [12]. In this method, hydrogen atoms are added to the CM side to satisfy the valency of the QM system. There are many variations within the implementation of the LA method, for example, the double Link Atom method [13], the Add-Remove Link Atom method [14] or the Scaled-Position-Link-Atom Method (SPLAM) [15]. Other termination schemes for the QM/CM boundary include the ‘pseudobond’ scheme [16], ‘IMOMM’ [17, 18] and ‘ONIOM’ [19, 20] procedures, and the ‘effective group potential (EGP)’ [21].

In this work, to ensure adherence to the consistent embedding principles, we have developed the “pseudo-atom” method as applicable to silica. The pseudo-atom serves as the QM-CM interfacial atom that provides the appropriate boundary conditions (chemical environment) for the QM region. In our formulations, we have

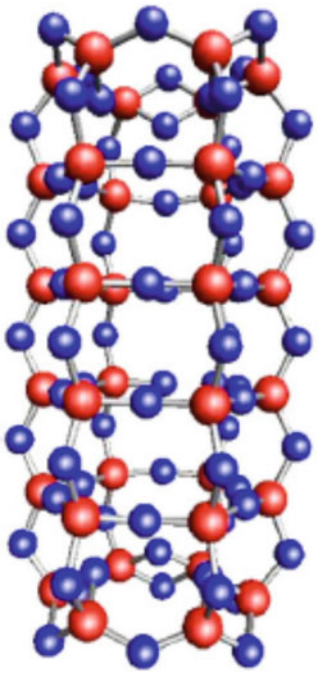

Fig. 6.3 Structure of the SiO_2 nanorod; *red spheres* represent Si atoms while *blue spheres* represent oxygen

chosen to have silicon atoms as the terminal atoms of the QM region, while the pseudo-atom corresponds to terminal oxygen atoms within the CM region. The training cluster is a pyrosilicic acid molecule (Fig. 6.4). The part of the molecule within the dotted lines is replaced by an effective fluorine (F) atom whose NDDO parameters are then adjusted to give the correct QM forces (which implies correct geometry as well) and charge density in other parts of the molecule (outside the dotted lines). The specific NDDO parameters modified are: one-center-one-electron integrals, Coulomb integrals, exchange integral, and two-center-one-electron resonance integrals [10]. The primary advantage of using pseudo-atoms is that they are trained to give the correct QM forces and charge densities for the QM-CM interfacial Si–O bonds. This allows use of the pseudo-atoms while studying dynamics in the system and enables embedding a QM region within a CM system. In addition, the surrounding CM region is effectively represented as a dipolar field, which is taken into account by the NDDO Hamiltonian that treats the QM region.

6.2.4 Quantum-Classical Multiscale Framework

As a first step towards examining the QM-CM interface, we characterize the ability of the multiscale representation to accurately model a silica nanorod as shown in

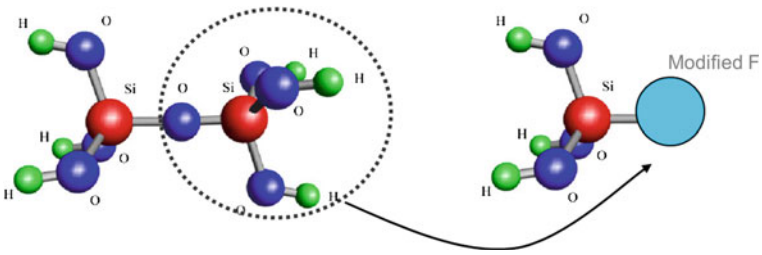

Fig. 6.4 Pseudoatoms (labeled Modified F) are trained to reproduce local effects in the electron density

Fig. 6.5. The two CM regions are represented as two dipoles for the top and bottom portions of the rod (Fig. 6.5). The values of the dipole have been calculated using the TH-NDDO charge density for these two portions of the rod. These domains are taken to be charge neutral, but are polarized by the presence of the QM domain. The validity of the approximation by dipoles has been checked by comparing the force on a silicon atom within the central ring (see Fig. 6.6) that arises due to all charges and that due to the dipole. Clearly the CM-QM is able to match the full QM representation, which is not the case when link atoms are used instead of the pseudo atoms. Further, it is found that the normalized difference of charge density is reproduced to within 0.1 % in the plane of the ring in the coupled QM-CM representation. This scheme was found to be applicable to both equilibrium and strained configurations, as illustrated in Fig. 6.6 for the following cases: (a) equilibrium, (b) the ring of the QM domain radially expanded by 5 %, and (c) a distorted ring in which one Si atom is radially pushed out and one Si pushed in. Also shown are the corresponding results for a longer 10 ring rod (cases (d), (e), and (f)).

In order to determine the efficacy of the CM-QM coupling, we next determine the elastic response of the 'composite' nanorod and compare it with QM-only and classical potential-only simulations. Here, the forces on the atoms in the CM region are calculated from the pair potential, while the forces on the atoms in the QM domain are provided by the TH-NDDO method. It has to be noted that CM atoms interact with atoms in the QM region via the interatomic potential, while the QM atoms 'see' the CM atoms via the dipolar field(s) that are calculated based on the positions of the atoms in the CM region. Figure 6.7 shows the stress-strain behavior of the nanorod obtained from three different methods: (i) TH method for the entire rod, (ii) pair potentials for the entire rod, and (iii) the composite rod constructed as described above. The three overlaying curves indicate that the composite rod is identical to the rod obtained from TH and the pair potential nanorod in terms of small strain elastic properties and structure. The stress-strain results show the success of our multiscale method indicating that the composite rod is indistinguishable from the underlying quantum mechanics for small-strains. But, at higher strains (>10 %), there is deviation in the respective responses. This is because the pseudo-atoms, trained at regions only close to the equilibrium configuration, fail to

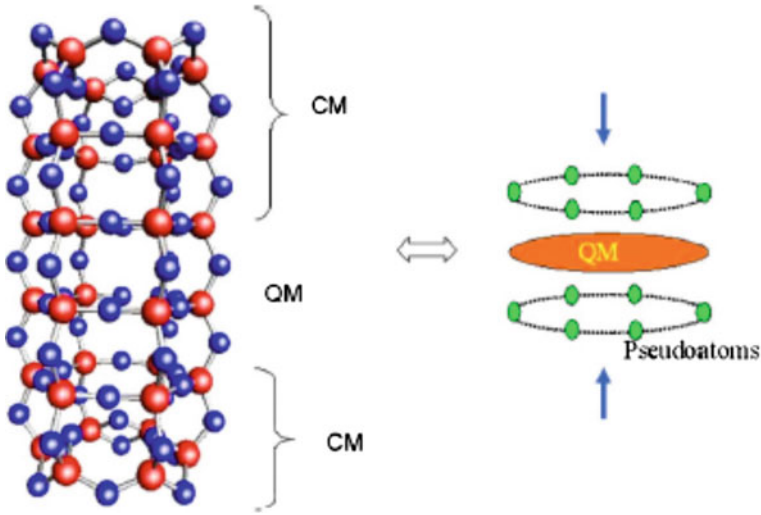

Fig. 6.5 Embedding scheme: approximating the CM region by pseudo-atoms and dipoles

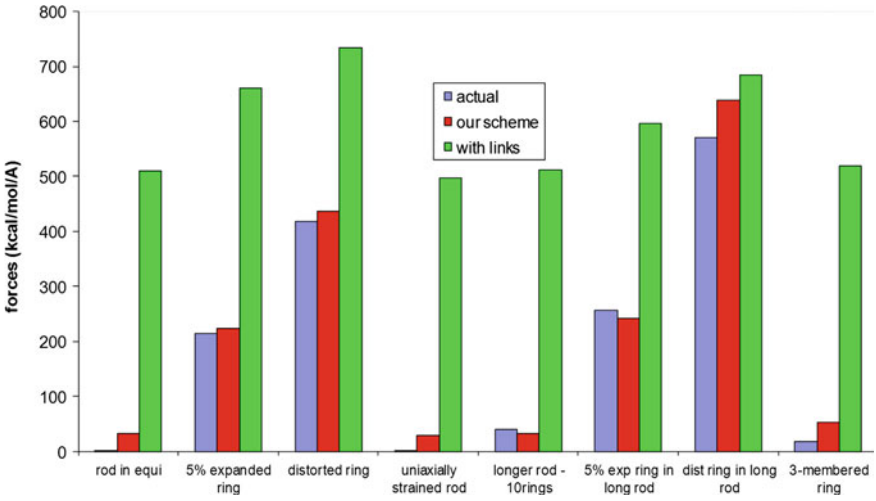

Fig. 6.6 Forces on Si nuclei for various cases studied

give the correct charge densities at such high strains. This diagnosis was checked by comparing the charge density in the QM domain of the 12 % strained rod with that of equilibrium configuration and a difference of 6 % was found. Retraining of the pseudo-atoms improve the stress-strain performance of the composite rod beyond a strain of 10 % is possible, however real systems have many inherent defects. These defects act as stress concentrators that cause the material to break at much lower strains than observed for the pristine nanorod, so retraining of the pseudo-atoms was deferred while the behavior of a defective rod was examined.

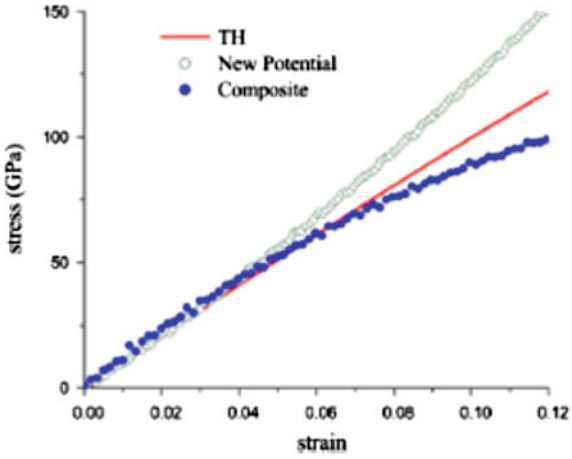

Fig. 6.7 Stress-strain curve for the TH (quantum), new potential (classical), and composite rods

Towards this end, a defect notch was placed in the 108 atom nanorod by removal of an oxygen atom as shown in Fig. 6.8. The MD stress-strain curves for this notched rod were found using TH quantum mechanics, the trained classical potential, and the composite. We note that the presence of only a small defect can significantly reduce the yield stress of the material and make it more prone to fracture. The TH curve for the defect-free rod is plotted in the same figure to contrast the value of the yield stress. As can be seen there is a reduction of $\sim$60 GPa in the yield stress. For the composite rod in this case, the QM domain was chosen to consist of 2 silica planes and the intermediate 5 oxygen atoms (see Fig. 6.8), so that the defect could be located in the QM region. The stress-strain curve for this composite notched rod agrees well with that for the TH quantum calculation up to 10 % strain. Above 4 % strain the curve for the composite notched

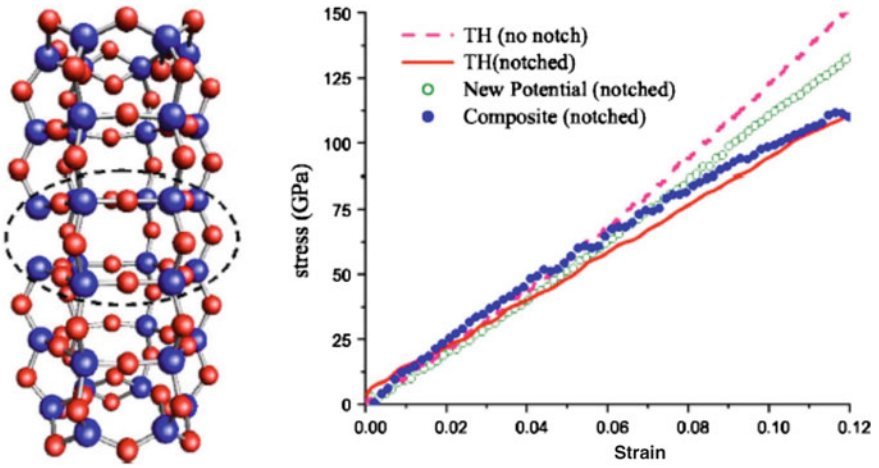

Fig. 6.8 The notched nanorod and corresponding stress-strain curve

nanorod follows that of the TH instead that of the trained classical potential nanorod, showing that the composite rod is representing the "real" material. This is exactly what is required of multiscale modeling.

So far, we examined the role of the interface when methods that correspond to different scales and theories are spatially coupled. Now we turn our attention to multiscale simulations where the scale-specific and/or the theory-specific methods are not spatially coupled in an explicit and dynamic fashion; rather the method of wavelets is invoked to enable coupling of different methods. The method called the compound wavelet matrix (CWM) method is given below and examples that demonstrate its use are then discussed.

6.3 The Compound Wavelet Matrix Method

The ability of wavelets to separate scale information effectively through the construction of local orthogonal basis functions makes them invaluable in the examination of local, global and scale-wise properties of any data set (or signal) that requires multiresolution analysis. As examples, such data could represent fractals, multifractals, turbulence patterns, or earthquakes [22–24]. In such applications, typically, wavelets operate on a signal obtained from a method restricted to a single discretization in time and/or space. CWM differs from such multiresolution techniques in that it bridges solutions obtained from multiple methods operating on different discretizations in time and space to yield a composite multiscale representation, while using wavelets as the basis. Utilizing wavelets as a multiscale tool rather than a pure multiresolution tool makes the CWM approach a *unique* as well as a *powerful* technique for coupling disparate simulation methods that address the same process, albeit at different levels of accuracy, or, equivalently different levels of theory. CWM also differs from up-scaling techniques [25, 26] used in adaptive bridging of scales, given that in the CWM framework, physical phenomena are examined using two or more different methods at their respective scales not necessarily based on the same variables or in the same mathematical form.

CWM is capable of combining short-time (and/or small spatial domain) fluctuations that could only be obtained from computationally expensive fine-scale methods, with the large-scale long-time (and/or large spatial domain) mean behavior of coarse-scale method(s), while assuming the overall trajectory fluctuations are stationary. It has been applied to various multiscale phenomena such as grain-growth and microstructure evolution [27], heterogeneous porosity/inclusions [28], dispersion problems [29], and diffusion from a reactive boundary [30]. In these studies, the fine-scale and coarse-scale information were combined via a compound wavelet matrix to yield the coupled multiscale behavior of the system over the entire extent of space/time. In other words, CWM was used as a *corrective* tool, involving the transfer of fine-scale fluctuations of the accurate method on to

the coarse-scale trajectory. In this paper, the CWM is further generalized to serve as an accurate *predictive* tool, by extending it to take into account the interaction of the fine- and coarse-scale methods as the simulated system evolves in phase space. Later in this chapter, we will explore some of these applications.

A discrete function can be hierarchically transformed into wavelet space by using discrete wavelet transforms (DWT) [31–35] via a series of 'scale-decompositions' with each stage of decomposition representing a level of description of the function specific to that scale. The transforms involving a set of wavelet filters in the form of linear convolution operators are applied hierarchically first to the full data vector (assuming the discrete function to be a 1-D data vector of length 2^N), then to a smoother vector of length 2^{N-1}, then to a vector of length 2^{N-2} and so on and so forth (as shown in Fig. 6.9a) until only a trivial number of smooth components (scaling or mother-function coefficients) remain; the number of scaling coefficients depend on the number of wavelet filters chosen. At any scale, the 'h_i' components of the transformed vector (Fig. 6.9b), which are essentially decorrelated, represent the wavelet coefficients corresponding to that scale, and characterize the scale-specific fluctuations inherent to the original function. The inverse transform is the exact reverse of the above process, with the wavelet operator (analysis operator) replaced by its inverse (synthesis operator). Different boundary conditions (e.g. periodic, fixed, zero, or on the interval), are implemented by modifications of the appropriate operator terms [35].

Every scale decomposition stage results in two sets of components: the smooth (or coarse) and the fine components, each representing a localized 'moving' average of the function (in wavelet space) with the fine components being regarded as the 'true' wavelet coefficients corresponding to the given scale. For convenience, at each stage, the data are rearranged as shown in Fig. 6.9b. Since the hierarchical decomposition at each stage acts only on the smoother components of the previous stage, only very 'coarse' information is available at the final stages of the wavelet transform procedure.

The given function (or signal) can be examined at any given range of scales simply by retaining the wavelet transform (WT) coefficients corresponding to the chosen scales and setting all other coefficients equal to zero and then sequentially reconstructing via the inverse transform to obtain a modified signal; this method is sometimes referred to as the padding technique. Extending these concepts, coupling of two different signals in wavelet space can be done by first identifying their overlapping scales (spatial/temporal) and then substituting corresponding wavelet coefficients from each signal into a compound wavelet matrix (CWM) that can be inverted to provide a hybrid signal that has the desired characteristics of both signals over the union of scales of the two signals. The identification of the overlap is critical to the CWM method and the extent of overlap decides the relative accuracy of the hybrid signal as explained below.

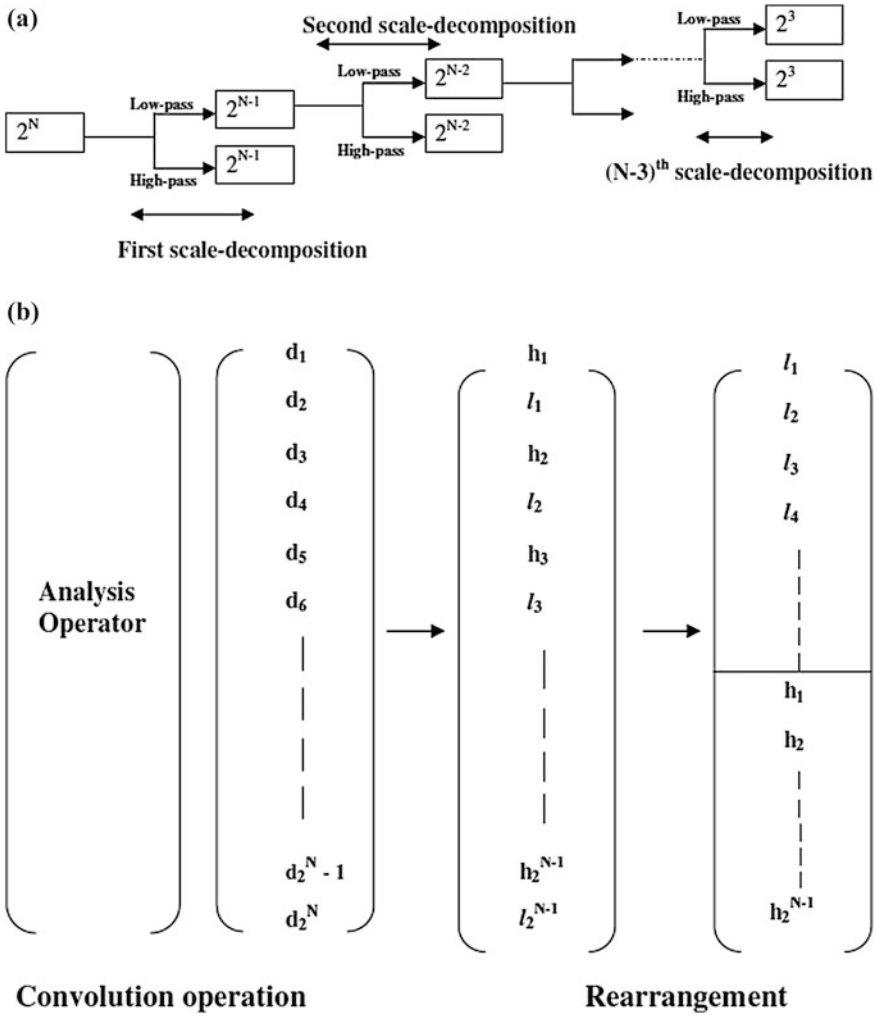

Fig. 6.9 **a** Stage by stage wavelet decomposition of a 2^N vector. At each stage, the data vector is convolved with two kinds of wavelet filters low-pass and high-pass resulting in two sets of wavelet-transformed data. **b** Illustration of the first stage of wavelet decomposition. The initial convolution operation is followed by rearrangement. Here *h* and *l* represent the high-pass and low-pass components, respectively

6.3.1 Compounding Methodology for Coupling Scales: Forming the CWM

Without the loss of generality, the methodology will be described based on two temporally varying 1-D signals, one representing the high-resolution, short-time

trajectory—the fine signal, while the other corresponds to a coarse-resolution approximation obtained over a much longer time interval—the coarse signal. Let R denote the ratio of the size of the coarse (Δt_c) over the fine (Δt_f) time steps, and N_c and N_f represent the number of coarse and fine data points respectively. The following steps result in a compound signal that contains essential information from both signals:

1. Coarse signal interpolation: Depending on R, the coarse signal is interpolated in order to obtain a new signal with a larger set of data points (I_c), with the time interval between successive data points equaling Δt_f. Note that I_c equals the product of R and N_c.
2. Wavelet decomposition: The wavelet transforms of the coarse and fine signals are performed until a trivial number of mother wavelet or scaling coefficients (S_N) are left. If $S_N = 8$ (2^3), then, p_c-3, and p_f-3 scale decompositions are carried out on the coarse and fine signal, yielding I_c-8 and N_f-8 WT coefficients respectively, where p_f and p_c are defined in (6.2).

$$\begin{aligned} I_c &= 2^{p_c} \\ N_f &= 2^{p_f} \end{aligned} \tag{6.2}$$

3. Selection/identification of overlapping scales: Overlapping scales refer to common or equivalent scales between the two signals; given that 2^{p_c} coarse data points and 2^{p_f} fine data points are sampled at the same rate (i.e. size of respective time-steps are equal), the first p_f-3 scale-decompositions for each signal are equivalent and yield $2^{(p_c-n)}$ and $2^{(p_f-n)}$ coefficients respectively at the nth scale decomposition (see Fig. 6.10).
4. Prolongation: At all overlapping scales, the fine WT coefficients are replicated to ensure consistency in the number of coarse and fine WT coefficients. The number of replications M_p at a specific scale is given by $M_p = I_c/N_f$ and represents periodic repetitions of the fine fluctuations at every scale with the inherent assumption that the fluctuations are quasi-stationary over the period of coupling. The primary role of this step is to set the stage for replacement of a majority of the coarse coefficients by the repeated fine coefficients to enable the transferability of all fine information at every overlapping scale.
5. Multiscaling via mixing and scaling of coefficients: This stage enables an 'intelligent' extension of the fine signal via appropriate selection of relevant wavelet coefficients at all scales from both signals to form a compound wavelet matrix (or vector in this case): in order to facilitate explaining the merging of coefficients, we will use for the coarse signal a smooth logarithmic function consisting of 1024 points (Fig. 6.11), while the shorter fine signal is comprised of the first 256 data points of the coarse signal with added white noise (via random numbers). Figure 6.12 presents (i) the WT of the coarse signal, (ii) WT of the first 256 points of same coarse signal, and (iii) WT of the fine signal; the above transforms use the symmetry-preserving bi-orthogonal CDF(4,6) wavelet

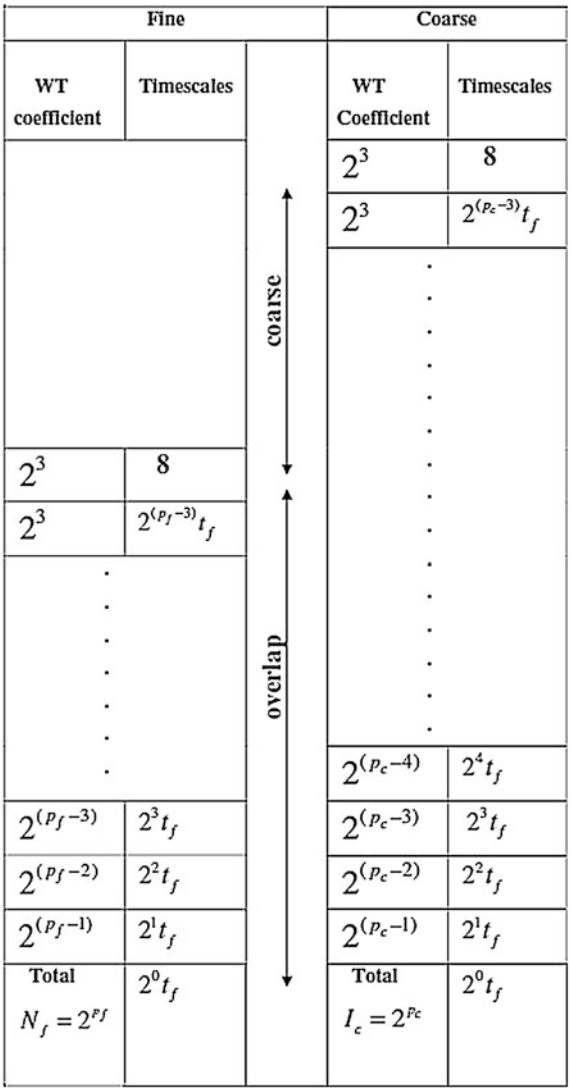

Fig. 6.10 Tabulation of the number of overlapping scales and "coarse-only" scales for fine and coarse functions consisting of N_f and I_c data points

filters with on-the-interval boundary conditions [35]. Specifically, in Fig. 6.12, the values of the various coefficients at all scales of decomposition are given (except the first scale).

The following are clearly evident from Fig. 6.12: (i) the initial few and the final few coefficients at every scale are significantly different in magnitude from the rest of the coefficients, (ii) excluding the final few coefficients, the $2^{(p_f-n)}$ coefficients of the shorter coarse signal and the first $2^{(p_f-n)}$ coefficients of the larger coarse signal (where n refers to the nth scale decomposition, and $2^{p_f} = 256$) at equivalent scales

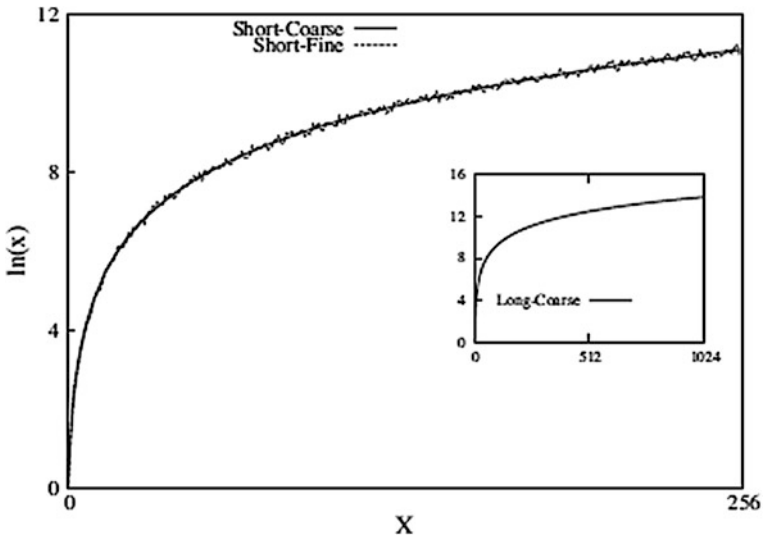

Fig. 6.11 Coarse and fine logarithmic signals used for explaining the CWM process. The inset figure represents the entire coarse signal

are identical, and (iii) the fine coefficients are slightly larger in magnitude than the corresponding short coarse coefficients at every scale.

The difference in magnitude between the initial (or final) few coefficients and the rest is due to the imposition of boundary conditions, which contribute significantly to the net magnitude of the boundary (i.e. the initial and final few) coefficients [35]. The fact that there is a one-to-one match between the short coarse and the corresponding long coarse coefficients (barring the final few coefficients) at every equivalent scale is a consequence of the moving average representation characteristic of wavelets. The deviations between corresponding short coarse and fine coefficients are a measure of the fluctuations (in this case, the white noise) inherent to the fine signal as also any possible deviations in the overall trajectories of the coarse and fine trajectories (in this case the overall trajectory is the same).

While boundary conditions do change the initial and final coefficients at each scale significantly, it has to be noted that when both short and long coarse signals are wavelet transformed, the initial coefficients are the same, while it is only the final coefficients that differ; this is due to the fact that the same domain is sampled via the convolution operation to yield the initial coefficients while the longer time data is not available to the shorter signal resulting in differences in the magnitudes of the final boundary coefficients. The repetition of fine coefficients at overlapping scales, replicates the essentially uncorrelated fluctuations associated with the fine signal over the entire domain of the long, coarse signal. This repetition has to be done with care to (i) prevent any unphysical behavior that could arise due to the periodic appearance of the boundary coefficients at each scale, while (ii) ensuring appropriate boundary conditions are still imposed. One could identify these

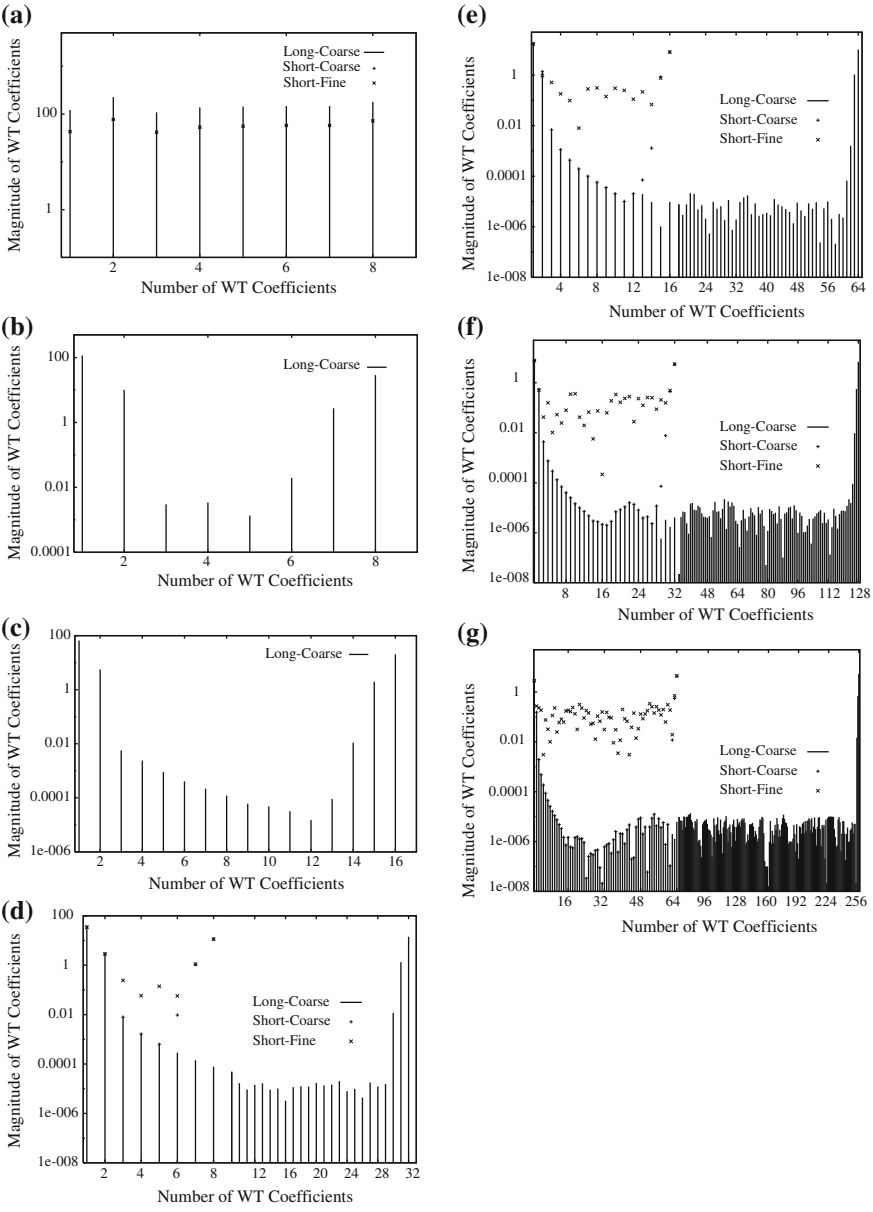

Fig. 6.12 Magnitude of WT coefficients at different stages (excluding the first stage) of decomposition of a 1024 data-set logarithmic signal as given in Fig. 7.11. Except for the initial and final three coefficients, note that the short-coarse coefficients coincide with the long-coarse coefficients and are always found *right on top of the impulses* that represent the long-coarse coefficients

boundary coefficients either from the knowledge of the extent (length) of the boundary elements in the analysis/synthesis matrix or by adopting an empirical procedure: at each scale, the deviation in magnitude between corresponding short coarse and short fine coefficients averaged over all coefficients (except for the first and last) is calculated. The number of boundary coefficients is identified as the number of successive coefficients greater in magnitude than the average deviation calculated for that given scale. Typically, for the interval boundary conditions used herein, the first three and the last three coefficients represent the boundary coefficients.

After replication, any fine coefficient that is greater in magnitude than the average deviation is replaced by the corresponding long coarse coefficient, except at certain locations, as explained below:

For clarity, consider the second equivalent scale decomposition of the long coarse (i.e. coarse) and short coarse signal (Fig. 6.12g). Only the first 61 (=64 − 3) coefficients are identical, while the 62nd, 63rd, and 64th coefficients are obviously different, though, in fact, the ratios of the respective boundary coefficients (i.e. ratio of 254–256 of long coarse to the 62–64 of short coarse) are very similar (analogous trends are true for all other equivalent scales too). In order to transfer the entire spectrum of fine information available at this scale, and to ensure that only the true fine signal fluctuations and trends are periodically replicated, an initial repetition of all (in this case $M_p = 4$) fine coefficients; next, the 62–67, 126–131, 190–195 of the repeated fine coefficients (corresponding to the boundary coefficients) are replaced by the equivalent coarse coefficients that are scaled by a multiplicative factor; this factor equals the ratio of the sum of absolute magnitudes of the short-fine coefficients to the sum of absolute magnitudes of the short coarse coefficients, with the boundary coefficients being excluded in the respective sums.

Keeping in mind that the ratio of the end boundary conditions of the respective signals at each scale are comparable, the 254th, 255th and 256th fine coefficients are scaled by a factor *Es*(1,2,3), where *Es*(1,2,3) are the ratios of 254th long-coarse to 62nd short coarse, 255th to 63rd and 256th long-coarse to 64th short coarse coefficients respectively. In a similar fashion, the above algorithm is repeated for all other overlapping scales too.

After appropriate repetitions and replacements at every overlapping scale, 32 remaining coarse coefficients are available (out of which, the first 8 represent the scaling coefficients) in addition to the 8 remaining scaling coefficients of the fine signal. Since no equivalent scales are available, the remaining 24 (9–32) of the 32 coarse coefficients are retained; in contrast, the first eight coefficients of the compound matrix are chosen to equal scaled values of the corresponding eight scaling coefficients of the fine signal, with the scaling factor similar to the definition of *Es* in the previous paragraph.

This method of repetition, scaling and replacement of the coefficients, enforces transferability of almost all fine information (excepting the repeated boundary coefficients), while ensuring an ‘informed’ modification of the end boundary coefficients and the scaling coefficients, which combine in tandem to enable the suitable extension of trends exhibited by the short time fine trajectory.

6. Compounding: The selected WT coefficients are assembled to form a compounded matrix (vector in this case) that has both coarse and fine WT coefficients, ready for inversion.
7. Reconstruction: The hybrid signal is now obtained via a series of convolution operations involving the compound matrix (vector) and the synthesis matrix.

The relative number of overlapping scales affects the quality of the compounding; a larger number of overlap scales leads to a better description of the fine-signal features, while a decrease in the number of overlap scales with respect to the total number of scales leads to reduction in the accuracy of the description of fine signal features. The rescaling of the eight mother-wavelet coefficients by the respective *Es* ensures that the resultant compounded trajectory represents a logical projection of the shorter fine trajectory extended in time.

6.4 Example 2: Microstructural Evolution of Materials Using the Compound Wavelet Matrix Method

In this example, we illustrate a simple yet powerful application of the CWM framework for bridging an atomistic model with a lattice Q-states Potts model to describe grain growth in nanocrystalline materials over different ranges of spatial and time scales. The primary advantage of this method is that the simulation time of the coarsest methodology is not controlled by the methodology with the slowest dynamics.

The polycrystalline state of matter is multilevel with information at several scales, Point defects, dislocations, grain boundaries, grain boundary junctions and iso-orientation aggregates of grains span a very wide range of scales from the atomic to the mesoscopic to the macroscopic. The geometric complexity is complicated further by the interactions among structural features at different scales. To cite only a few: grain boundaries which may act as source or sink for point defects; grain boundary/dislocation (from bulk to primary to secondary grain boundary dislocations) interactions and their effect on grain boundary migration; grain boundary/grain boundary interactions during grain annihilation or topological transitions such as grain switching; and low energy/high energy grain boundary interaction in anisotropic polycrystals. In addition to the structural hierarchy of polycrystalline materials it is essential to also recognize their dynamical nature [36]. Processes at various length scales usually possess different relaxation times suggesting that in addition to the spatial hierarchy a temporal one exists as well. For instance, although point defects may exist at equilibrium, dislocations and grain boundaries may never be at equilibrium. Microstructures and the underlying grain distribution are therefore only a reflection of the fact that a polycrystalline material has not had enough time to reach equilibrium (lowest energy corresponds to the single crystal state).

While it is well known that normal grain growth achieves a quasi-stationary distribution of grain sizes after a transient period, a complete picture of grain growth of polycrystalline materials is still to be achieved, the main reason being that the overall kinetics are controlled by geometrical features and relaxation times at largely different scales. In particular, grain growth kinetics with anisotropic grain boundaries is governed by physics at three different scales, namely atomic scale at the level of individual grain boundaries, a larger scale associated with the grain size and the even larger scale of clusters of grains with similar grain boundary energy.

Atomistic computer simulation techniques such as MD can capture the physics of the phenomenon from the atomic scale and up. However, to study grain growth at the scale associated with clusters of grains as well as its interplay with processes at the two other smaller scales, one would need to simulate an excessively large number of atoms for an excessively long time unachievable with present and foreseeable computer capabilities. On the other hand, several mesoscopic computational techniques such as Monte Carlo (MC) simulations based Q-states Potts model [37], continuous models [38], Vertex model [39] have been developed to simulate the evolution of microstructure in polycrystalline materials, but all these techniques are usually limited in their inability to include atomic scale features. Thus, given the hierarchy of spatial scales and time scales associated with grain growth, it is important to couple simulation techniques such as MD and mesoscopic models to study grain growth over the complete spectrum of scales. Towards this end, MD simulations of Lennard-Jones (L-J) systems and MC simulations based on the Q-states Potts model are employed within the CWM framework to capture the multiscale physics of grain growth in materials. Specifically, a 2-D grain growth is modeled and both MD and MC are applied simultaneously at the coarse and fine spatial scales.

The two-dimensional simulated system is modeled at the atomic scales using the 6–12 Lennard-Jones potential [40] with parameters $\varepsilon = 119.79$ K and $\sigma = 3.405$ Å. This atomic system consists of 22,500 particles in a cell with edge length ~ 0.0503 μm. Interactions between atoms are extended up to third nearest neighbors. A polycrystalline microstructure with fine grains is initially obtained by quenching a liquid. This initial microstructure is then evolved with a constant temperature-constant volume (NVT) MD algorithm. The temperature is maintained at approximately 70 % of the melting point. Periodic boundary conditions are used for about 400,000 MD integration time steps or nearly 1.7 ns, after which a coarser microstructure and a 63 % drop in energy are obtained. This microstructure is then characterized by calculating the excess atomic potential energy of each individual atom (relative to the potential energy of an atom in a perfect lattice at the same temperature). The excess atomic energy is then normalized by the total excess energy of the microstructure at $t = 0$. The spatial distribution of the normalized excess atomic energy is then mapped onto a 256×256 square matrix to obtain what will be referred to in the rest of the discussion as an energy map. We note that the L-J system includes grain boundary anisotropy, that is, the grain boundary energy may vary depending upon the degree of misorientation between adjacent grains. A microstructure may therefore contain low-energy low-angle grain boundaries and

high-energy high-angle grain boundaries. Moreover, the anisotropy in grain boundary energy also results at fixed misorientation from the inclination of the grain boundary plane.

In addition, MD simulations of a system four times larger are carried out. The same conditions are imposed and as a result a microstructure that was 63 % less than the initial structure was obtained, similar to the smaller MD system. The spatial variation in energy for this system is mapped onto a 512 × 512 square matrix. This large MD system serves as the yardstick to assess the CWM based coupling of smaller MD system with the Potts model.

Next, a Monte Carlo (MC) simulation of grain growth with the Potts model is carried out. In general, the Potts model maps the microstructure onto a discrete lattice coarser than the atomic scale, and the "spin" state S = 1,…,Q of each lattice site represents the orientation of the grain in which it is embedded [37]. A grain boundary exists between two adjacent lattice sites with different orientations. For the sake of simplicity, the Potts model we use here does not include orientational anisotropy; the energy associated with two neighboring sites with different spin orientations does not vary with the magnitude of the difference in spin.

The system size of the Potts model used here is the same size as that of the bigger MD system or equivalently four times that of the smaller MD system described above. Periodic boundary conditions are imposed and only nearest neighbor interactions are considered. The spatial discretization corresponds to 128 × 128 sites. In the MC algorithm, the thermal energy, k_BT, (where T is the temperature and k_B is Boltzmann constant) is 0.2 J with J being the energy associated with two misoriented neighboring sites. This temperature is sufficient to alleviate the pathological problem of growth on first nearest neighbor square lattices at low temperature. Indeed it is known that for the square lattice with nearest neighbor interaction, domain grain growth at zero temperature stops when the domain vertices absorb all initial wall curvature [41]. However, Anderson et al. [42] have shown that at finite temperature the square lattice with nearest neighbor interaction and Q > 3 exhibits significant coarsening, grain boundary curvature and microstructure resembling that obtained with a triangular lattice. In this work, we chose Q = 10, to ensure the consistent embedding requirements are met. Specifically, it was seen that this choice yielded a similar (as compared to MD) time evolution in the potential energy (averaged over four trajectories) of the Potts model as shown in Fig. 6.13. Note that the potential energy in the Potts model represents an excess energy relative to a perfectly ordered system (perfect crystal). In this context, it is worth pointing out that the conversion from the MC time to real time relies on a scheme, which is based on grain boundary migration [43].

Figure 6.14a, b represents the normalized energy maps as obtained at the end of the large MD and MC simulations respectively. From these figures it is evident that fine-scale features that characterize the MD energy map are not present in the Potts model maps. Thus a true multiscale representation of the microstructure can be obtained as a union between the smaller MD system and Potts model microstructures. Towards this end, the CWM method in 2-D is implemented as follows. In 2-D, given an energy map of 512 × 512 points such as that generated in the large

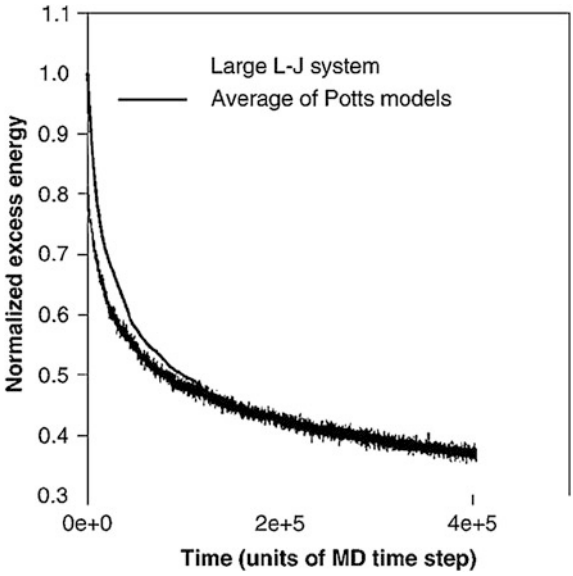

Fig. 6.13 Excess energy versus time obtained with the MD simulation of a L-J system and average of MC simulations of four Potts models with different initial configurations. The excess energy relative to the perfect crystal at the same temperature is normalized to the excess energy of the initial microstructure. The MC time is scaled to match the MD time (see text for details)

MD simulation, the wavelet transform consists of three 256 × 256 matrices (corresponding to x, y, xy directions), three 128 × 128 matrices, and so on; each decomposition level is at half the resolution from the previous one. The final level of decomposition represents the map at the coarsest resolution. Wavelet analysis of a MD energy map provides its wavelet transform coefficients from the atomic to its coarsest scale (corresponding to the physical dimensions of the system). Similarly the wavelet coefficients of an MC energy map extend over scales ranging from the lattice spacing of the grid to the system size.

Note that the Potts model system is four times larger than the smaller MD system and the corresponding number of scales is 7 ($2^7 = 128$). In contrast the larger MD system of comparable size has 9 ($2^9 = 512$) scales, while the smaller MD system has 8 scales though its size is a quarter of the Potts model/large MD. Thus to ensure appropriate comparisons, a 512 × 512 CWM matrix is formed by compounding the wavelet coefficients from the smaller MD and Potts representations as per the strategies elucidated earlier. Only the coefficients from the Potts model corresponding to the coarsest scales (1–2) are retained while the wavelet coefficients corresponding to scales 2–8 of the smaller MD system are used for representing the finer scales of the CWM (i.e. 3–9 scales of the CWM). We have used biorthogonal symmetric wavelets with four vanishing moments and the corresponding so-called scaling functions [33] as our wavelet basis sets in this work.

In order to carry out a scale-by-scale comparison of the MD/Potts representation with the larger MD system, the energy corresponding to each scale is calculated. We have used energy maps from frozen microstructures in order to minimize the contribution of thermal vibration to the energy of the small and large L-J systems. Figure 6.15 shows the energy of the small and large MD systems, the MC Potts

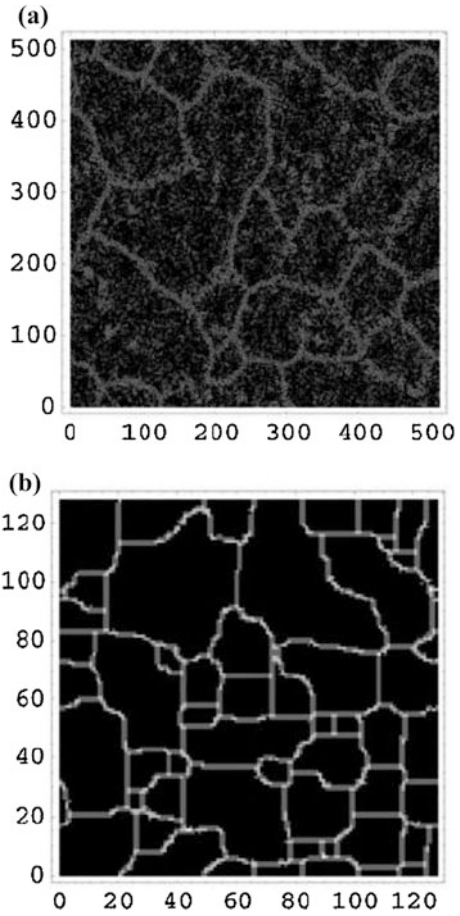

Fig. 6.14 Gray scale representation of the energy maps (see text for definition) for: **a** L-J system. **b** Potts system. The energy increases from *black* to *white*

model and CWM coupled MD/Potts model, with respect to scale. The energy associated with a given scale is evaluated from the wavelet representation of the energy maps at that same scale. For the wavelet representation of the energy map at scale n, the wavelet coefficients at all scales except those at n are set equal to zero; with this set of coefficients, the inverse wavelet transform is performed. This inverse wavelet transform represents the spatial distribution of the contribution of scale n to the energy map. The total energy associated with scale n is then calculated as the sum of the energies in the representation at scale n. Figure 6.15 points out to the fact that the Potts model deviates significantly from the larger MD model especially at finer scales as expected. But, by implementing the CWM to integrate the smaller MD and MC together, the ability to emulate variation in energy of the more accurate larger MD system is significantly enhanced leading to good agreement between the CWM based MD/Potts model and larger MD system.

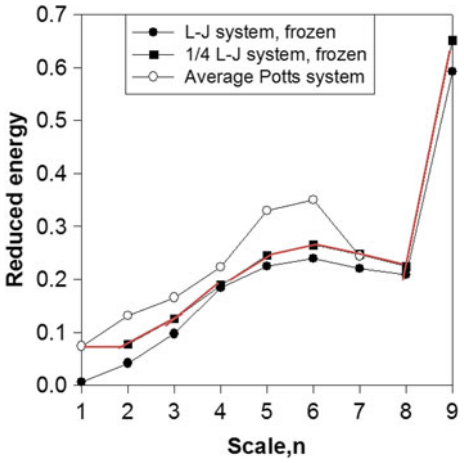

Fig. 6.15 A scale-by-scale representation of the energy of the small and large LJ system and the Potts system. Note that the *red-line* represents the energy of the CWM-combined MD-Potts system (Colour figure online)

Using the above example, we have illustrated how the mathematical tool of wavelet transformation can be used to analyze the microstructure of two-dimensional polycrystalline systems as well as their evolution during normal grain growth. The CWM based method presented here enables characterization of structural features at different spatial scales and their respective contribution to the energy of a microstructure. More generally, using an illustrative example, we have demonstrated the potential of using wavelet analysis to bridge computer simulation techniques that model a given phenomenon at different scales.

6.5 Example 3: The Dynamic-CWM (dCWM) Approach Applied to Reactive Flows

The dCWM method extends the CWM approach to enable the concurrent coupling of different scale-specific methods by dynamically combining the fine and coarse simulation methods over successive sub-intervals (dynamic updates) to obtain the compound trajectory. At each dynamic step, the fine and coarse trajectories are coupled via CWM, and the resulting end point of the compound signal in trajectory space becomes the starting point (initial condition) for computing the next set of trajectories respectively; this procedure is then repeated for the desired number of intervals as illustrated in Fig. 6.16. dCWM has the advantage of inherently coupling the dynamics of each method, in addition to lending further stochasticity to the compound trajectory, which could lead to a better exploration of the trajectory space. Furthermore, in the 'static' CWM formulations, it was implicitly assumed that the fluctuations obtained from the fine signal, which was evaluated over much shorter time, was representative of fluctuations occurring over a much longer duration. This is not necessarily true as the magnitude as well as the nature of

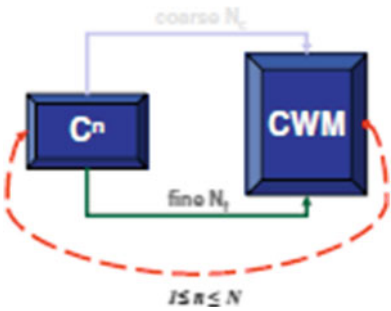

Fig. 6.16 Illustration of the dCWM algorithm

fluctuations can and in general does change as the system evolves. This is overcome in the dCWM approach, since the fine method is employed at various stages of the evolution of the system. At the same time, the advantages of the CWM method are retained, without any measurable increase in computational expense. Thus dCWM can be viewed as a predictor-corrector method, while the CWM is simply a corrector method. In order to illustrate the dCWM approach, we examine a model system, consisting of two species A and B that participate in a first-order reversible reaction occurring at a surface as given in (6.3). As discussed later in this section, this problem demonstrates consistent embedding in the time domain.

$$A \underset{k_{ba}}{\overset{k_{ab}}{\leftrightarrow}} B \tag{6.3}$$

The surface represents the boundary of a semi-infinite positive half space, with the half space suitably discretized to handle diffusion of both species to (from) the reactive surface. For convenience, the two reaction rates k_{ab} and k_{ba} are set to equal unity, and the initial concentration of A at the reactive surface is chosen to be much larger in order to bias the forward reaction to be more active at the initial stages of the diffusion-reaction process. Further, initial concentration of both species is set to zero everywhere other than at the reactive surface. Our numerical model of the system is 1-D in space, such that the diffusion domain is discretized into line increments and nodes, and the reactive surface is represented by a single boundary node. The infinite extent of the diffusion domain is approximated by choosing a sufficiently large length and large number of nodes such that for the duration of the simulations, there is no change in concentration of species at nodes far from the reactive boundary and that concentration remains equal to zero.

Two levels of description (coarse and fine) are used. The accurate 'fine' method employs the stochastic kinetic Monte Carlo (KMC) method [44] in conjunction with finite differences to simulate the chemical reaction at the reactive boundary and diffusion respectively, while using relatively smaller time increments as dictated by KMC. In contrast, the coarse method models both reaction and diffusion deterministically using much larger time increments, thereby addressing much larger temporal dimensions, in a less precise fashion. For simplicity, both coarse and fine spatial dimensions as well as spatial discretizations are taken to be the same.

Using an explicit Euler scheme, the diffusion equation (6.4) for species diffusion is solved numerically, where C_s and D_s represent the species concentration ([A] and [B]) and their respective diffusion constants. The size of the time step for the fine method is obtained from KMC calculations (described below), while the coarse time step is chosen to be a multiple of the fine time step size. The flux of the diffusing species is set to zero at the boundary node (i.e. at the reactive site).

$$\frac{\partial C_s}{\partial t} - D_S \nabla^2 C_S = 0 \tag{6.4}$$

In the fine model, the occurrence of the two reactions (6.3) at the reaction site is determined stochastically via KMC based on Gillespie's stochastic simulation algorithm [44]. The reaction times for both reactions are given by

$$\delta t_{ab} = -\log(1 - R_1)/[A]k_{ab} \tag{6.5}$$

$$\delta t_{ba} = -\log(1 - R_2)/[B]k_{ba} \tag{6.6}$$

where R_1, and R_2 are distinct numbers obtained from a uniform random number distribution and [A] and [B] *r*epresent the atomic/molecular concentrations of the respective species. The occurrence of either one of the reactions is decided by the minimum of the two reaction times as given by the following equations.

If

$$\delta t_{ab} < \delta t_{ba}, \begin{matrix}[A] = [A] - 1 \\ [B] = [B] + 1\end{matrix} \tag{6.7}$$

while if

$$\delta t_{ab} > \delta t_{ba}, \begin{matrix}[A] = [A] + 1 \\ [B] = [B] - 1\end{matrix} \tag{6.8}$$

Next, the smaller of the two times is chosen to be the size of the time step for solving the diffusion equation. If the resultant time step size (Δt) is found to be incompatible with the Courant condition for numerical stability of the diffusion equation i.e. if $\Delta t > \Delta x^2/2D_S$, where Δx is the size of the finite difference element, Δt is then sub-divided into smaller time steps as follows:

$$\begin{aligned} \Delta t &= n_t \Delta t' + dt \\ n_t &= \mathrm{int}\left(\frac{\Delta t}{\Delta t'}\right) \end{aligned} \tag{6.9}$$

where $\Delta t' = \Delta x^2/2D_S$.

At every dynamic step, the fine simulation is propagated for N_f time steps and a resultant variation in concentration of the two species at all nodes is obtained as a function of time. Keeping in mind that KMC results in unequal spacings in time, a spline interpolation routine is invoked to get the concentration variation as a function of equally spaced time steps (Δt_f). Once Δt_f is obtained, Δt_c is evaluated ($=R\Delta t_f$) and the coarse signal is propagated for N_c time steps.

In the coarse simulation, the concentration of each species at the reaction site is obtained deterministically in a straightforward fashion from the respective reaction rates (6.10) using a finite difference scheme. Once the coarse and fine responses are obtained, the steps enumerated in the section on CWM is used to form the compounded matrix (vector), with which the compound trajectory for the given dynamic step is obtained, with the final concentrations of each species at all nodes being the starting concentrations for the next set of coarse and fine calculations. Since both kMC as well as the deterministic method model the same reaction, the consistent embedding principles are satisfied; interestingly, this work demonstrates consistent embedding multiscaling in the time domain.

$$\begin{aligned} A(t+\Delta t) &= A(t) - k_{ab}A(t)\Delta t + k_{ba}B(t)\Delta t \\ B(t+\Delta t) &= B(t) + k_{ab}A(t)\Delta t - k_{ba}B(t)\Delta t \end{aligned} \tag{6.10}$$

In order to present the advantages of using the dCWM framework for multiscale simulations such as the 1-D reaction-diffusion problem described above, we chose a model system that is simple enough, so that a stochastic 'fine' simulation of the reaction-diffusion process (i.e. until the system attains chemical equilibrium), is possible. Thus one can benchmark and quantify the ability of the dCWM framework to emulate the multi-scale behavior of the system. In this context, in addition to a coupled fine-coarse representation, the problem is also modeled using the (i) stochastic fine method and the (ii) deterministic coarse method. In each case, the number of nodes (i.e. extent of spatial discretization) equaled 1000 and the size of each node was 0.0625 in arbitrary units of length (l); note that in addition, time was also expressed in arbitrary units (t). Thus the units of concentration are expressed as *atoms/l*, while the units of the diffusion constant are expressed as l^2/t. The initial concentration of [A] and [B] at the reactive site was set to 2500 and zero respectively, while the respective concentrations at all other sites were identically set to zero. The diffusion constant of both species was chosen to be the same and equaled 0.01. Figure 6.17a, b present the long-time variation in concentration of [A] and [B] at the reactive site as obtained from the fine method (KMC + finite differences); the simulation was carried out until there was a net saturation in the respective concentrations, though there was still the persistent presence of fluctuations in both trajectories. To achieve equilibrium 65,536 time steps were used, with the spline interpolated time step size equaling 5.66×10^{-3} t (total time = 371.0t). The overall trend (i.e. the mean trajectory) was independent of the choice of the random number seed, while there were very minor differences in the occurrence of fluctuations for different random number seeds. Also, the amount of fluctuations greatly diminished

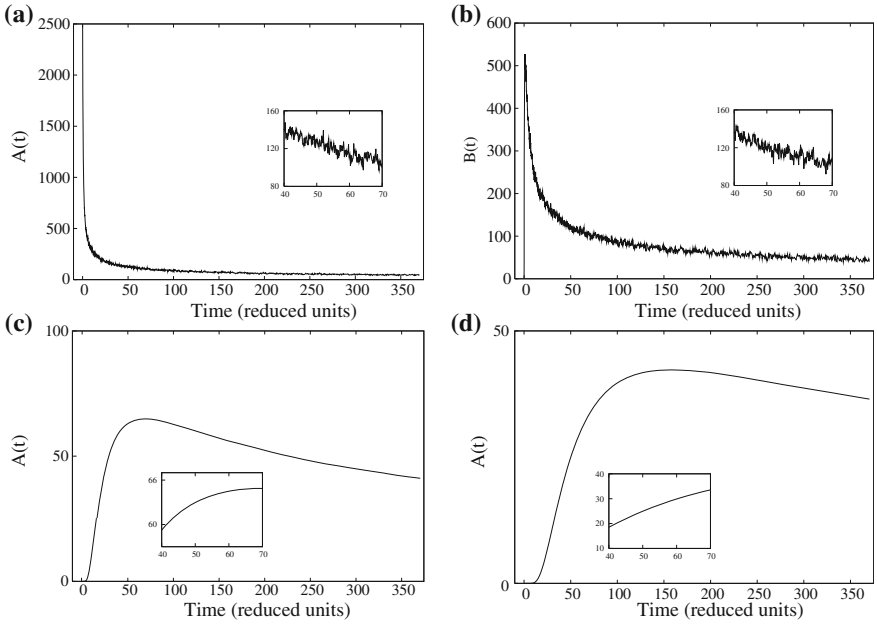

Fig. 6.17 **a** and **b** represent the fine-method variation in concentration of species *A* and *B*, respectively, with time. Note the fluctuations in the fine signal. **c** and **d** represent the time variation in concentration of species *A* at nodes 50 and 100, respectively. The *insets* illustrate the variation in concentration within a much narrower period of time. *Note* y-axis scale is not constant across the plots

with distance from the reaction site (Fig. 6.17c, d) and in addition, far away from the reactive site, the concentrations essentially remained zero. As a study in contrast, a full deterministic simulation was carried out for the same number of time steps and the same time-step size as that of the fine method. Clearly, as evident from Fig. 6.18a, b, the fluctuations characteristic of the fine method were absent from the respective coarse trajectories.

Having the results at both ends of the scale spectrum, i.e., the results as obtained using the accurate stochastic fine method and the purely deterministic coarse method, attention is turned to examining the ability of the dCWM approach to model the reaction-diffusion process accurately. The dCWM method depends critically on the choice of N_f and the number of dynamic updates (N_{dyn}). Ideally, as the system evolves in time, the fine simulation has to be invoked as frequently as possible to enable sampling at various points in phase space, while N_f has to be large enough to capture the stationary fine-scale fluctuations as and when the fine simulation is invoked. Since $N_f \times N_{dyn}$ represents the total number of fine computations, it would be judicious to maximize N_{dyn} and reduce N_f as much as possible to limit any redundancy in the fine computations, while still being able to capture all essential physics inherent to the fine simulation at different stages of the simulation

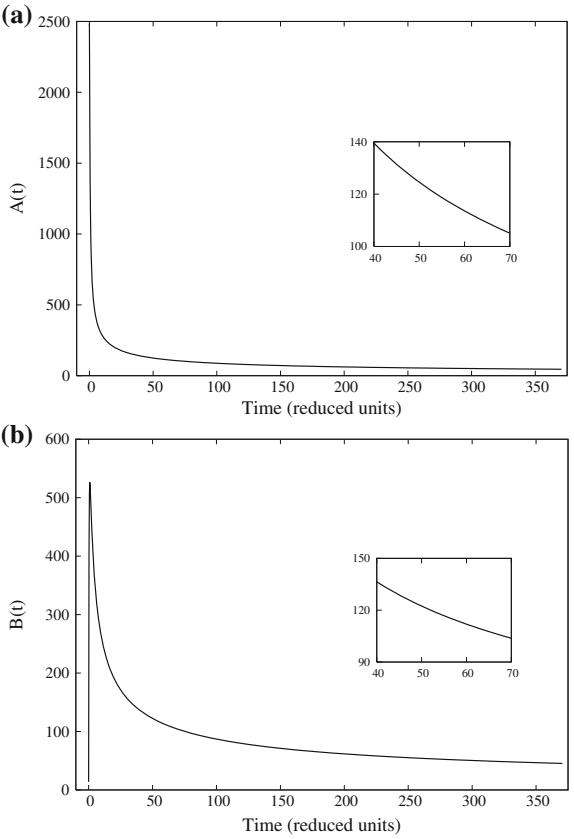

Fig. 6.18 **a** and **b** represent the coarse-method variation in concentration of species *A* and *B*, respectively, with time. The *insets* illustrate the variation in concentration within a much narrower period of time

for a given number of fine computations. Equally important choices are that of M_p in tandem with N_c, as implicitly, it is assumed that the fluctuations captured over N_f time steps are stationary over the time equaling $M_p \times N_c$. In order to comprehensively characterize the importance of the different simulation parameters as well as also to aid comparison between the dCWM and the above-discussed stochastic and deterministic results, the following were investigated: (i) the interplay between N_f and N_{dyn} for a given number of total fine ($N_f \times N_{dyn}$) and coarse ($N_c \times N_{dyn}$) computation steps, and (ii) the importance of the relative number of scales beyond overlap (or equivalently M_p) for a given N_f. Note that for all investigations, the compound trajectory obtained via the dCWM method consisted of 65,536 points—the same as the pure coarse and fine trajectories.

As a first part to this study, the interplay between N_f and N_{dyn} for a fixed value of M_p and a given number of fine computation steps ($N_f \times N_{dyn}$) is investigated. Obviously, N_c equals $M_p \times N_f$. Figures 6.19 and 6.20 illustrate the variation of the two concentrations at the reactive site for the specific value of $M_p = 8$ and different

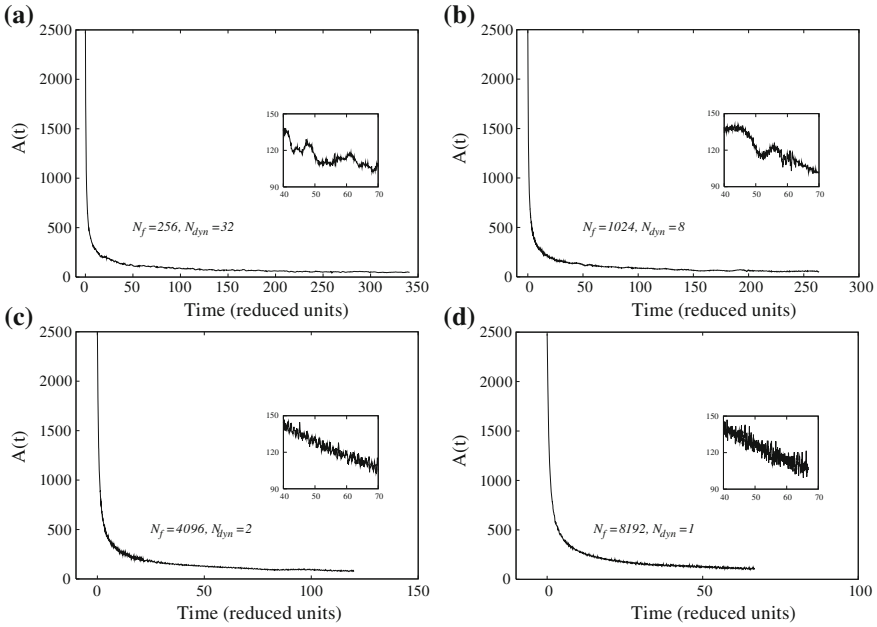

Fig. 6.19 Time-variation in concentration of species *A* for $M_p = 8$ and different values of N_f and N_{dyn} such that their product is constant. The *insets* illustrate the variation in concentration within a much narrower period of time

combinations of N_f and N_{dyn}, while Table 6.1 gives the list of values of N_f, N_{dyn} and N_c used in Figs. 6.19 and 6.20. It has to be pointed out that the time step size for the coarse method was set to equal that of the spline interpolated fine method time step size at each dynamic step, thereby avoiding the need for interpolation of the coarse information at each dynamic step.

A central observation regarding the effect of N_{dyn} and N_c emerges; for a given number of coarse-computation steps (i.e. 65,536 points) and fine computation steps ($=N_{dyn} \times N_f$), the amount of time addressed in each case increases with increasing N_{dyn} (or decreasing N_f) as noted in Table 6.1. The fact that during the initial stages of each case, there is a monotonic and a relatively rapid decrease (or increase) in the concentration of *A* (or *B*) at the reactive site, indicates that, initially, the forward reaction is selectively preferred by KMC. More importantly, at the initial stages, relatively much smaller time step sizes are selected by KMC due to the higher concentration of [*A*] (see (6.5) and (6.6)), while in the later stages, the probability of occurrence of larger time steps is much higher. This fact is reflected in Fig. 6.21, which depicts the variation in the time step size (evaluated at the end of every N_f fine computations), as a function of the number of dynamic steps. Thus, for cases involving large N_f (and therefore large N_c, for a given M_p), where the dynamic

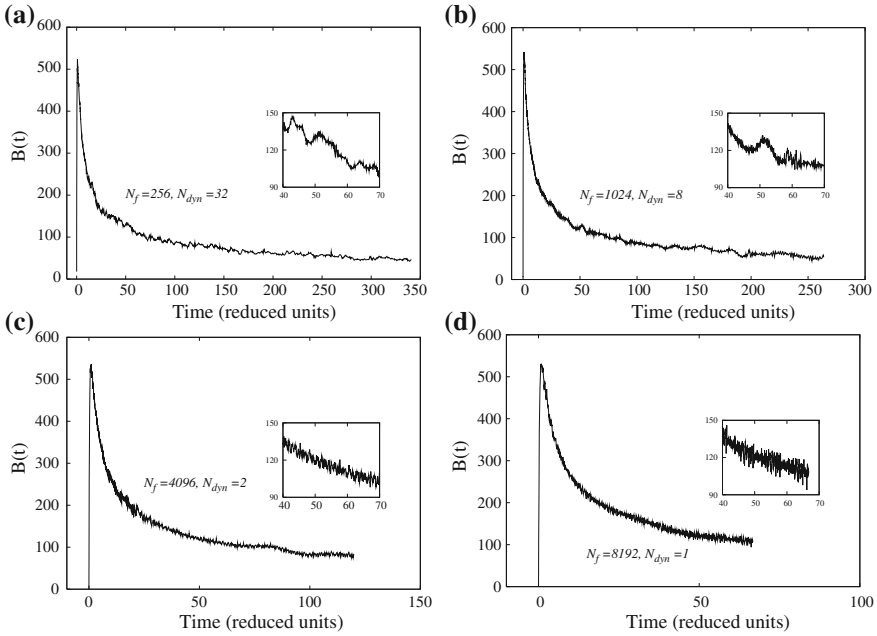

Fig. 6.20 Time variation in concentration of species *B* for M_p = 8 and different values of N_f and N_{dyn} such that their product is constant. The *insets* illustrate the variation in concentration within a much narrower period of time

Table 6.1 Values of various simulation parameters used for the case when M_p = 8 and $N_f \times N_{dyn}$ are constant

M_p = 8	N_{dyn}	N_f	N_c	Total time (t)
$N_f \times N_{dyn}$ = 8192	32	256	2048	341.3
$N_c \times N_{dyn}$ = 65,536	8	1024	8192	264.0
	2	4096	32,768	120.2
	1	8192	65,536	66.8

updates occur less frequently, smaller time step sizes are used over much larger chunks of computational steps, leading to much smaller time coverage. In fact, for the N_{dyn} = 1 case, which represents the original CWM approach, the least time coverage can be seen. Also, the dependence of total-time coverage was very weakly dependent on the choice of the random number seed (for KMC), and any difference in the times were negligible for all practical purposes.

The importance of N_{dyn} is shown by the insets of Figs. 6.19 and 6.20. One can observe clear differences in the overall shape of the concentration variation as a function of N_{dyn}. Particularly, for larger N_{dyn}, there are distinct oscillatory features (as pointed out in the inset figures), similar in nature to those present in the fine trajectories (Fig. 6.17), that are noticeably absent in the other cases; in other words,

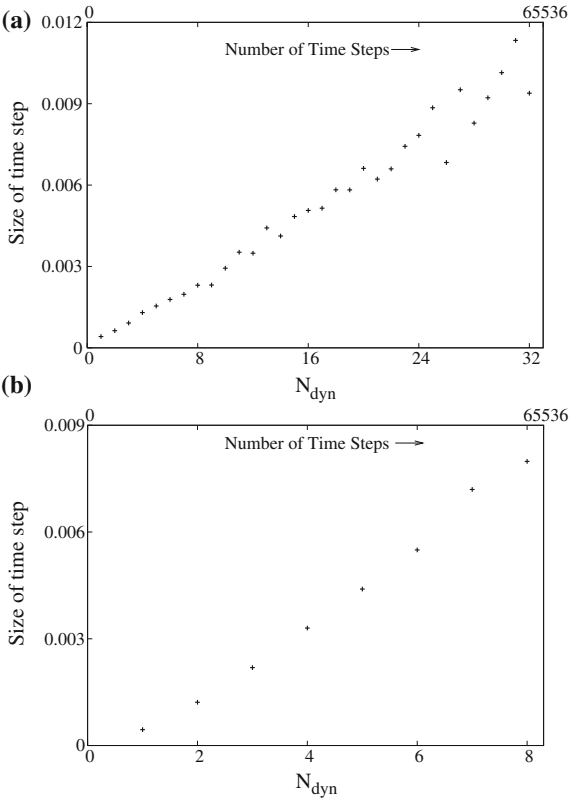

Fig. 6.21 Variation in size of time step as a function of the number of dynamical steps for **a** $N_{dyn} = 32$ and **b** $N_{dyn} = 8$

if one were to compare the CWM (Figs. 6.19d, and 6.20d) and the dCWM results, it is obvious that in addition to larger time coverage, trends characteristic of the concentration variations which were not captured by the CWM approach are described by the dCWM method.

Next, the role of non-overlapping scales (S_B) and M_p (for a given N_f) in controlling the accuracy of the resulting trajectories is examined. Table 6.2 lists the values of M_p, S_B, N_{dyn} and N_c used in investigations for a specific value of N_f (=256) as well as the number of overlapping scales (S_O). As expected, there was a noticeable dependence of the total time coverage on N_{dyn} (and M_p) as given in Table 6.2. For smaller values of N_{dyn} the dependence was much stronger and, for large values of N_{dyn} (≥32), the total time coverage approached that of the full fine simulation.

The variations in dCWM with respect to the amount of very fine-scale fluctuations and coarser-scale oscillatory trends were examined for different N_c (therefore different S_B), N_{dyn} and M_p. Recognizing the effectiveness of wavelets as scale filters, the contributions of different scales to the composition of each trajectory can be examined via the padding technique discussed previously. For instance, the mean trend of each trajectory can be filtered out by removing the coarsest-scale

Table 6.2 Values of various simulation parameters used for the case of $N_f = 256$

Total time steps = 65,536	N_{dyn}	M_p	N_c	S_B	Total time (t)
	128	2	512	1	365.2
	64	4	1024	2	350.0
$N_f = 256$	32	8	2048	3	341.3
$S_O = 5(8 - 3)$	16	16	4096	4	308.0
	8	32	8192	5	250.4
	4	64	16,384	6	180.9
	2	128	32,768	7	83.6
	1	256	65,536	8	27.5

contributions, enabling the quantification of the fluctuations corresponding to the finer scales, or in contrast, the fine-scale fluctuations could be filtered out, while retaining the larger-scale oscillatory contributions to the overall trajectory. The following steps were used to examine the effect of the various scales:

1. Interpolating all trajectories using splines to ensure equivalent time step sizes. This is done since each trajectory corresponds to different time coverages and consequently different time step sizes. All resulting trajectories are now resampled at the time step size as that of the '$N_{dyn} = 1$' case, and leads to a one-to-one comparison between each dCWM trajectory and the benchmark, for the time coverage of the given dCWM trajectory.
2. Next, the padding technique is invoked to scale separate the finer and coarser information available for each trajectory. For extracting 'coarse' information, all coefficients beyond a cutoff scale are set to zero, while the opposite is done to obtain 'fine' information. This allows the ability to clearly demarcate the different scale contributions.

As an example of the above method, consider Fig. 6.22a, b, which present the '$N_{dyn} = 128$' trajectory of species [A] when reconstructed selectively by inclusion of specific scales. Clearly, the resultant coarse trajectory (Fig. 6.22a) obtained by excluding all scales greater than 4 ($2^4 = 16$) is (almost) devoid of any fluctuations, while the frequency and magnitude of fluctuations increase with inclusion of additional finer scales. The exact opposite behavior is seen for the corresponding fine components (Fig. 6.22b), with maximum stationary fluctuations present when scale filter (i.e. number of excluded scales) equals 4. Thus, in order to characterize the quality of each dCWM trajectory, the respective fine components obtained at very small values of the scale filter (4–6), were examined and the 'energy' corresponding to the transferred fine fluctuations was evaluated by calculating the average intensity of the fluctuations for each case. The average intensity was in turn, evaluated by calculating the square root of the sum of squares of the magnitudes of fluctuations over the time domain of each trajectory and then averaging it over the total number of time steps.

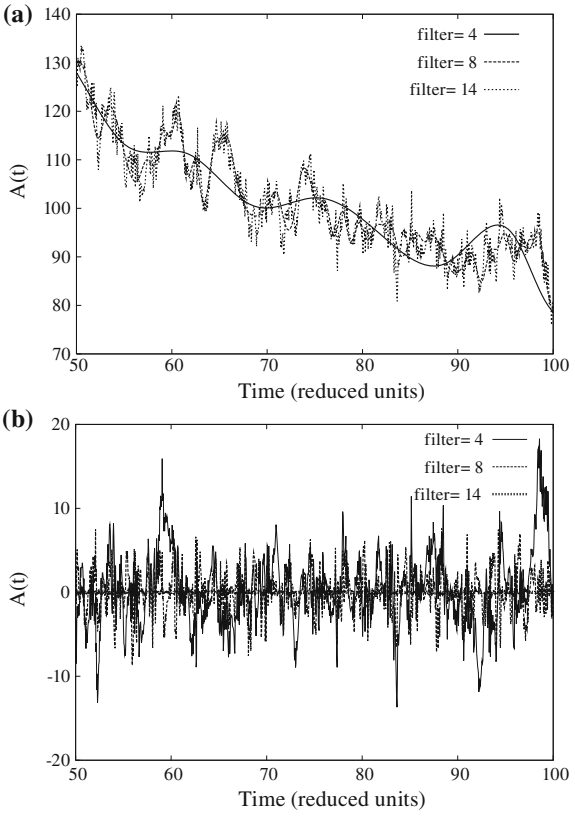

Fig. 6.22 Time variation in the (**a**) coarse and (**b**) fine components of species *A* concentration for different values of the scale filter for the $N_{dyn} = 2$ dCWM trajectory

Depending on the time coverage of the different dCWM trajectories, the corresponding average intensity of the benchmark trajectory was evaluated, thereby enabling the quantification of energy transfer expressed as a ratio of the corresponding dCWM intensity to the benchmark intensity (see Fig. 6.23a). In Fig. 6.23a, for trajectories with $N_{dyn} \geq 8$, the energy was evaluated over 524,288 (2^{19}) points resampled at the $N_{dyn} = 1$ time step size, while, 262144, 131072 and 65536 points were used for $N_{dyn} = 4$, 2, and 1 respectively. Also, note that for a given choice of N_{dyn} (and other related parameters), and a given value of the scale filter, the calculated energies were very similar for different values of the random number seed (5 different trajectories for each case were considered), and the standard deviation in the scale energy was less than 1.0 %. Figure 6.23b provides information on the energy contained in each scale for the benchmark trajectory evaluated over 524,288 points; this can be used for the purposes of estimation of scale energy for the different dCWM trajectories when used in conjunction with Fig. 6.23a.

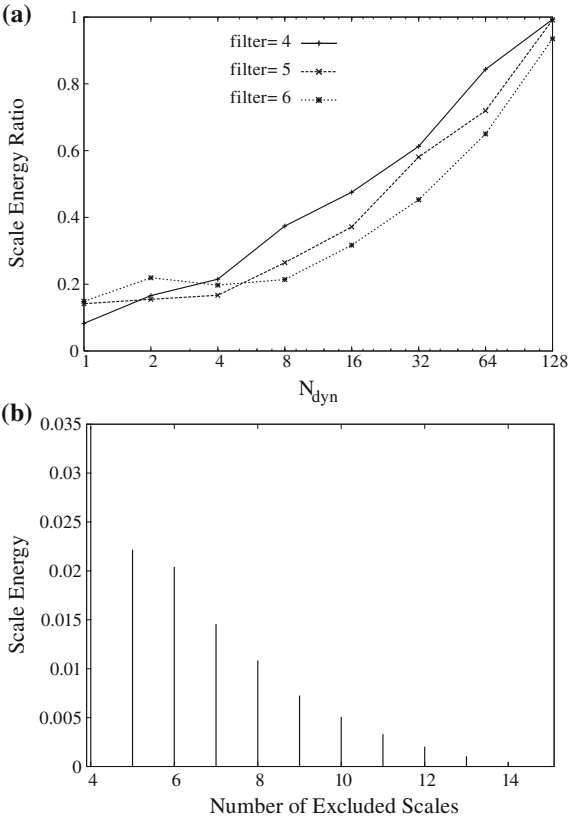

Fig. 6.23 **a** Variation in the relative energy of different dCWM trajectories as a function of N_{dyn} for three different values of the scale filter. **b** Scale-energy variation of the benchmark signal; here the number of excluded scales represents the value of the scale filter

As expected, Fig. 6.23a shows a strong correlation between the energy ratio and N_{dyn}, such that with increasing N_{dyn}, there is a parallel increase in the energy of the corresponding trajectories; in fact the energy of the $N_{dyn} = 128$ trajectory is remarkably similar to the benchmark for the three different scale filters given. Since the computational cost for the fine method is the most expensive, one can relate M_p (and N_{dyn}) in tandem with the ability of the corresponding trajectories to emulate the benchmark, to be a measure of the computational efficiency of the dCWM method. Obviously, the benchmark stochastic simulation is the most computationally expensive calculation, while at the other end is the non-dynamic CWM simulation where $N_{dyn} = 1$ ($M_p = 256$).

Figure 6.23a in conjunction with Table 6.2, show that 65 % of the benchmark energy is replicated with a time coverage of 90 % of the benchmark time for one eighth of the computational cost ($M_p = 8$, $N_{dyn} = 32$). With a computational savings of a factor of 4, i.e. $M_p = 4$, $N_{dyn} = 64$, 85 % of the benchmark energy is captured with a corresponding 92 % time coverage, while for $N_{dyn} = 2$ or equivalently for half the computational cost of the benchmark, there is almost 100 % efficiency.

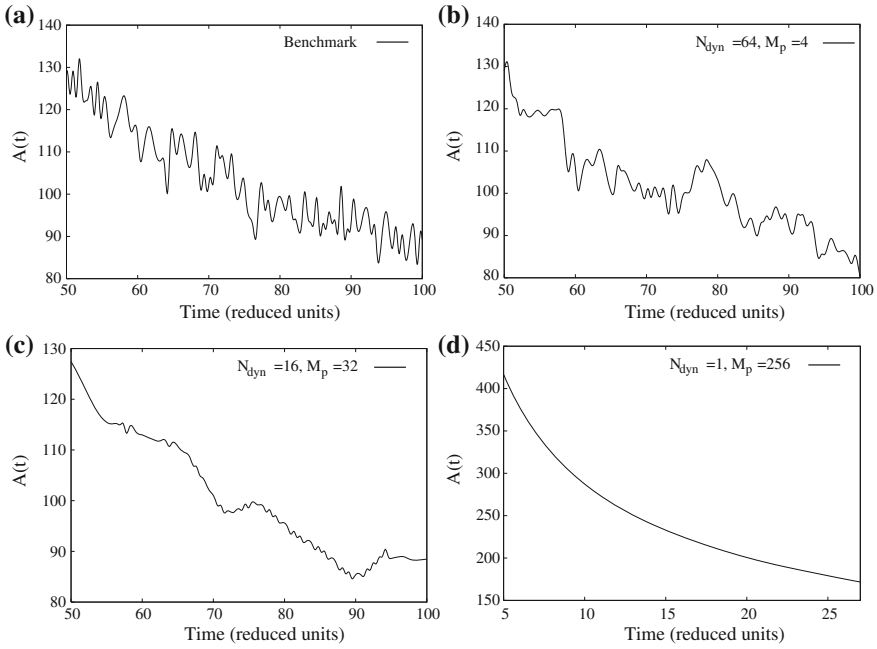

Fig. 6.24 Time variation in the coarse components of concentration of species *A* for different values of N_{dyn} and M_p for a fixed value of N_f = 256 as obtained for scale filter equal to 10

These results accentuate the fact that the quality of the trajectories depends intimately on the choice of N_{dyn}, S_B, and M_p; an increase in fine computations ($N_{dyn} \times N_f$) leads to a marked improvement in the quality of the trajectory as evidenced in Fig. 6.23a as well as in Fig. 6.24, which depicts select dCWM 'coarsened' trajectories devoid of the very fine 'white noise' components, and evaluated at scale filter = 10.

For large S_B (i.e. scales beyond overlap), the deterministic coarse behavior dominates, thereby suppressing the coarser-scale contributions of the fine method towards shaping the trends of the dCWM trajectory. An equally important reason for the dominance (suppression) of either method is the fact that there are $(M_p - 1) \times 6$ replacements of the fine method coefficients by the scaled coarse method coefficients at each scale, and at larger M_p, this effect is more pronounced leading to considerable loss in transfer of the fine-method information.

The above results clearly indicate that the dCWM method is capable of reproducing many aspects of the true behavior of the simulated system for a lower computational cost. The dCWM method can be readily extended in a non-trivial yet a straightforward way to model far more complex multiscale phenomena in a computationally efficient fashion by combining different levels of spatio-temporal description of the constituent degrees of freedom. In addition, combining dCWM with other multiscale methods such as parallel-in-time PIT [45] has led to further

reduction in computational overhead as demonstrated by the time-parallel CWM (tp-CWM) method. Here, the system is initially modeled by the coarse method; next, the coarse trajectory of the system is broken up into a series of smaller segments, with the initial conditions of each segment serving as starting points for the fine method. Computing the more accurate trajectories (by the fine method) from these starting points are then done in parallel, until the respective end points are reached. We refer the reader to [46] for more details on tp-CWM.

6.6 Example 4: A Cautionary Tale: Multiscale Models for Elastic Wave Propagation in Materials

In linear elasticity the fundamental properties such as stress, strain and elastic moduli are thermo-mechanical quantities. These quantities are defined such that they satisfy the thermodynamic and the long time limit. That is, these quantities represent averages over a large enough number of microscopic constituents and nearly infinite time. Calculating some of these quantities from atomistic models does not present significant difficulties as long as large enough systems and long enough times are used. This constitutes the basis for coarse graining that enables the extension of atomistic systems into the realm of continuous models with seamless coupling between length scales [3]. Problems in bridging continuum and atomistic regions may arise when the continuum or part of the continuum region is pushed outside the thermodynamic and long time limit. This is the case in many of the multiscale methodologies that examined crack propagation where the finite elements coupled to an MD region are reduced to "atomic" dimensions as discussed above. The spatial coupling between unphysically small FE and atoms implies also that the long time limit may not be satisfied. In addition, an elastic continuum does not obey the same physics over all possible wavelengths as that of a discrete atomic system. This physical mismatch is easily noted in the dispersion relations of both systems that overlap only in the long wavelength limit [47]. Based on the above observations, one can expect an elastic impedance mismatch between a continuum and atomic simulation when an attempt is made to couple them.

In this regard, we have undertaken a study to examine and quantify the impedance mismatch between an **elastic** continuum and an atomistic region as the continuum spatial and temporal scales are forced toward atomic scales. We have coupled dynamically an elastic continuum modeled with the finite difference time domain (FDTD) method and an atomistic system modeled with MD. The impedance mismatch at the interface between the MD and the FDTD systems is probed with an incoming elastic wave packet with broadband spectral characteristics centered on a predetermined central frequency. Reflection of part of the probe wave packet is a sign of impedance mismatch between the two systems. This reflection is characterized in the time and frequency domains over spatial and time scales

ranging from atomic scales, nanometers (nm) and picoseconds (ps), to the thermodynamic/long time limit scales, millimeters (mm) and seconds (s).

The FDTD method solves numerically the elastic wave equation in homogeneous or inhomogeneous media. The elastic wave equations are integrated by means of discretization in both the spatial and the time domains [48, 49]. More specifically, real space is discretized into a grid on which all the variables and parameters are defined. The main variables are the acoustic displacement and the stress field at every site on the grid. The relevant parameters of the system are the densities and elastic constants for each constitutive element. The relevant parameters of the simulation are the grid spacing and the size of the time step. Appropriate boundary conditions such as periodic boundary conditions or absorbing boundary conditions are applied.

The FDTD scheme discretizes the wave equation $\frac{\partial^2 u_i}{\partial t^2} = \frac{1}{\rho}\frac{\partial T_{ij}}{\partial x_j}$ in both the spatial and time domains and explicitly calculates the evolution of the displacement 'u' in the time domain (T_{ij} is the stress tensor, ρ is the density, and u_i is the displacement of the ith element). For the sake of simplicity, we limit the system to one-dimensional propagation. The FDTD region is discretized into N 1-D elements. We assume that the FDTD region to be infinitely stiff in the other two directions. The elastic wave equations are approximated using forward differences in both time and space. The displacement u_n of any element 'n' at each time step is a function of the stress gradient $\sigma_n'(t)$ across that element. The displacement and the stress evolution of the system is given by

$$\begin{aligned}
\varepsilon_n(t+\Delta t) &= \frac{u_{n+1}(t) - u_n(t)}{\Delta x} \\
\sigma_n(t+\Delta t) &= C_{11n}\varepsilon_n(t+\Delta t) \\
\sigma_n'(t+\Delta t) &= \frac{\sigma_{n+1}(t+\Delta t) - \sigma_n(t+\Delta t)}{\Delta x} \\
u_n(t+\Delta t) &= 2u_n(t) - u_n(t-\Delta t) + \frac{\Delta t^2}{\rho_n(t)}[\sigma_n(t)]
\end{aligned} \tag{6.11}$$

where Δx is the length of each element, Δt is the size of the FDTD time step, and ε_n, C_{11n}, and ρ_n are the strain, stiffness and the density of the nth element. Thus in this technique we can predict the displacement of every element knowing the stress on that element. We assume that the stress on any element is uniform. Absorbing boundary conditions [48] are implemented in order to prevent reflection from the end elements of the FDTD mesh. The following relations denote the boundary condition

$$\begin{aligned}
u_n(t+\Delta t) &= u_{n-1}(t) + \frac{c_0\Delta t - \Delta x}{c_0\Delta t + \Delta x}u_{n-1}(t+\Delta t) - u_n(t) \\
u_1(t+\Delta t) &= u_2(t) + \frac{c_0\Delta t - \Delta x}{c_0\Delta t + \Delta x}u_2(t+\Delta t) - u_1(t)
\end{aligned} \tag{6.12}$$

where $c_0 = \sqrt{\frac{C_{11}}{\rho}}$ and corresponds to the longitudinal velocity of the elastic wave through the medium.

While we assume that the reader is familiar with MD (see Chap. 2), for the sake of clarity we provide some basic concepts of MD. The essence of MD methods involves solving the N-body problem of classical mechanics [50]. In other words, it involves solving Newton's equations of motion for a given set of particles, the interactions between particles governed by an interatomic potential, enabling one to keep track of the evolution of the system in phase space. The equations of motions are solved using standard finite difference schemes at each time step of the simulation. An MD simulation can be carried out under a variety of constant thermodynamic conditions. Here we use the macroscopic conditions of constant strain, number of molecules and temperature. In our studies, the MD system is propagated through phase space by solving for the equations of motion at each time step using the Verlet integrator. Temperature is maintained constant via a momentum rescaling procedure. Periodic boundary conditions are employed.

The elastic system to be probed is chosen to be Argon. The elastic constant C_{11} was found from a series of MD simulations carried out under the following conditions: The model for the atomic system was a three-dimensional face centered cubic (FCC) crystal with periodic boundary conditions containing 500 particles interacting through a 6–12 Lennard-Jones potential with parameters chosen to simulate argon. The interatomic potential was truncated at a distance of 8.51 Å. The one-dimensional long-time limit stress (σ)-strain (ε) curve for that crystal (Fig. 6.25) was obtained at 46 K with the temperature maintained via a velocity-rescaling scheme. For these calculations, a strain was applied in one direction while maintaining the length of the other edges of the simulation cell rigid. The strain was applied in increments of 2×10^{-4} in the interval [−0.1 to 0.1] and the resulting stress was then calculated from a virial-like equation [51] by averaging over 5000 MD time steps. An MD time step (dt) equals 10.0394 fs. The curve was then fitted to a third degree polynomial (6.13).

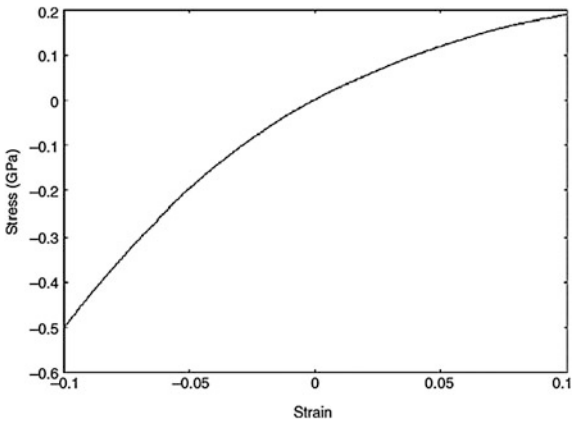

Fig. 6.25 Stress-strain curve for the LJ-MD system

$$\sigma = 4.304 \times 10^{10}\varepsilon^3 - 1.54 \times 10^{10}\varepsilon^2 + 3.045 \times 10^9\varepsilon \tag{6.13}$$

The handshaking between FDTD and MD is handled by replacing one FDTD by a 3-D MD cell, the MD cell parameters chosen to exactly match the 'model MD system'. As shown in (6.11), the equations of motion for the propagation of the wave through the medium involves solving for the displacement as well as the stress fields for every FDTD element at every FDTD time step. The stress and the strain for every element are assumed to be uniform over its length. Thus, when an FDTD element is replaced by an MD cell, the equivalent displacement and stress for the element is calculated by uniaxially straining the MD cell along the direction of the wave propagation (the strain is obtained from (6.11)). The condition of rigidity in the other two-directions is satisfied by keeping the length of the edges of the MD cell constant in these directions. The average value of the MD stress is evaluated for every FDTD time step with the final configuration of the MD atoms obtained at the previous FDTD time step serving as the initial state for the current MD calculation.

Since the continuum is assumed to be perfectly elastic, the elastic constant (C_{11}) of the FDTD region was chosen to be 3.045 GPa, which corresponds to the linear term of (6.12). This ensures a coarse serial coupling between the MD and FDTD region satisfying the consistent embedding requirement.

A preliminary study of the wave propagation characteristics indicated that the FDTD time step (Δt_{crit}) had to be smaller than $(\Delta x/2c_o)$ for a stable algorithm. Thus at every FDTD time step, the MD stress is calculated by averaging over $N_{md}(= \Delta t/dt)$ time steps, with $\Delta t \leq \Delta t_{crit}$. A reduction in Δt automatically leads to a decrease in the number of MD time steps over which stress is averaged (for every FDTD time step). It is therefore possible to push the time coupling between the two simulation techniques in such a way that one achieves in the smallest limit, a one to one correspondence between the two time steps. The FDTD/MD hybrid method, therefore, allows us to test a range of time scaling conditions from coarse graining to time matching between a continuum and an atomic system.

The number of FDTD elements is chosen to be 10,000. The length of each FDTD element (Δx) was chosen to equal the zero pressure box length of the MD cell (26.67 Å). The central wavelength of the wave packet was chosen to be an integral multiple of Δx, to ensure stability of the FDTD algorithm.

The probing wave packet is a one-dimensional wave packet and is of the form $a_0\cos(-kx)\exp\left(\frac{-(kx)^2}{2}\right)$, where k is the wave vector, and a_0 the maximum amplitude of the wave. The wave is propagated through the medium with an initial longitudinal velocity c_0. The signal's frequency spectrum is broadband and the central frequency of the wave packet ν equals c_0k. The impedance mismatch between the two systems (FDTD and MD) was probed as a function of the central frequency of the wave packet (ν). Simulations were carried out for two values of ν namely 0.393 and 3.93 GHz with a_0 equaling 50 and 5 Å respectively. Though, the maximum displacement a_0 equals 50 Å (which is greater than the length of the FDTD element), the strain on any element is only a fraction of its length (refer

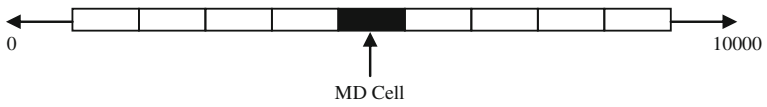

Fig. 6.26 An illustrative representation of the system consisting of 10,000 elements; the *open boxes* represent FDTD elements and the *darkened box* corresponds to the MD cell

(6.11) for the definition of strain). At every frequency, the size of the FDTD time step (Δt) was systematically decreased, and the impedance mismatch was studied as a function of N_{md}, with N_{md} equaling (94, 47, 23, 12, 8, 4).

The signal is initially centered about the 5000th element and is propagated along the positive X direction. The MD cell is located at element #6000 (see Fig. 6.26). The coupling is examined by analyzing the reflected signal at an element some distance away from the MD cell. This signal is compared and contrasted with the signal that is reflected in the case when the MD cell behaves as an FDTD element with a non-linear C_{11} as determined previously from the third-order (σ) versus (ε) relationship. The latter case will be referred to as the 'pseudo MD-FDTD coupling (PC)' while the former will be referred to as 'real-time MD-FDTD coupling (RTC)'. Discrete fast Fourier transforms (FFT) [52] are used in obtaining the frequency spectrum of all the signals.

Case 1: (v = 0.393 GHz)

Figure 6.27 represents the signal as it propagates (for the case when N_{md} = 94) as well as its frequency spectrum. Even though a cursory glance at Fig. 6.27 may appear to indicate to the naked eye that the signal propagates through the medium without any loss, there is a fraction of the initial signal that is reflected from the MD-FDTD interface. The reflected signals (both RTC and PC) were obtained as a function of N_{md} (Fig. 6.27). The magnitude of the reflected signals does not change significantly as a function of N_{md}. Therefore in this paper we discuss the signals for some representative cases (N_{md} = 94, 23 and 4). As obvious from Fig. 6.27, the magnitude of the 'PC' signal is always smaller than the 'RTC' signal, though the general shapes of the signals are identical.

The frequency spectra of these signals are shown in Fig. 6.28. One notices the fact that for every value of N_{md}, the lower end of the frequency spectrum for both the 'RTC' and the 'PC' signals are pretty similar, with the number of features (humps) in the signals increasing with decreasing N_{md}. In addition, for every N_{md}, the 'PC' signal has an upper cutoff at around 3×10^9 Hz, while the 'RTC' signal has a cutoff at around 150 GHz.

Case 2: (v = 3.930 GHz)

As in the previous case, though the signal appears to propagate without any visible loss (Fig. 6.29), there is a small amount of reflection. Similar to the previous case, the amount of reflection does not depend significantly on N_{md}. But unlike the previous case, the 'RTC' signals are not similar to the 'PC' signals. There is a clear mismatch in the frequency spectrum of the two signals, with the intensity of the

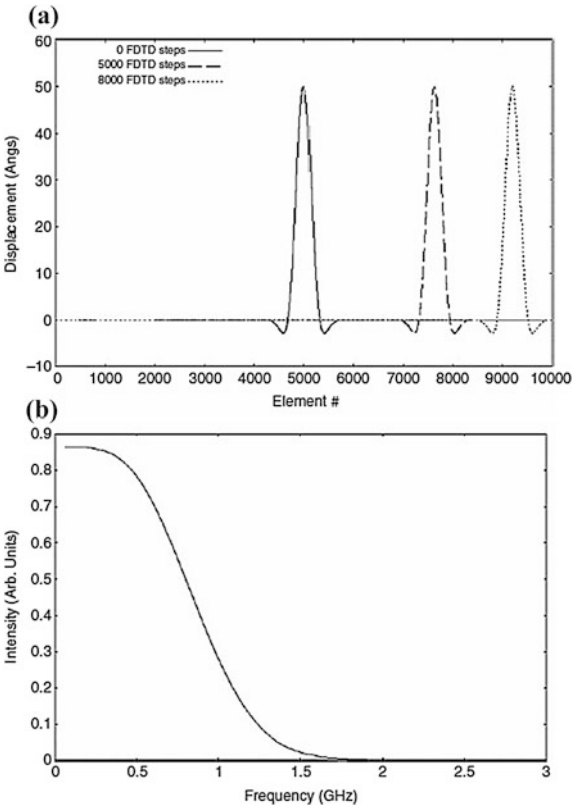

Fig. 6.27 **a** Time evolution of the wave packet @ 0.393 GHz for N_{md} = 94. **b** Frequency Spectrum of the wave packet

frequencies being significantly higher for the 'RTC' signals (Fig. 6.30). The cutoff for the 'PC' signals is around 30 GHz (Fig. 6.31), which is an order of magnitude greater than the cutoff for case 1. This corresponds exactly to the fact that ν for case 2 is an order of magnitude greater than ν for case 1. Interestingly, the cutoff for the 'RTC' signals is still around 150 GHz.

Based on the signal propagation characteristics for the above two cases, the following can be inferred: The amount of reflection (both RTC and PC signals) is independent of N_{md}. The 'PC' and the 'RTC' spectra are comparable at the lower probing frequency, and the mismatch significantly increases for the higher probing frequency. The 'RTC' and the 'PC' signals have distinct cutoffs, with the cutoff for the 'PC' signals being much smaller than that of the 'RTC' signals. This can be explained on the basis that the 'PC' signal represents the long-time limit of the coupling, where the high frequency (short wavelength) modes are averaged out, while the abrupt cutoff for the 'RTC' signal represents an upper-limit in the frequencies that can be supported by the FDTD system. The discretization of the continuum into small elements modifies its dispersion relation by introducing an upper limit on the frequencies (a Debye-like frequency) that can be resolved. This upper limit on frequency for traveling waves depends on the extent of discretization

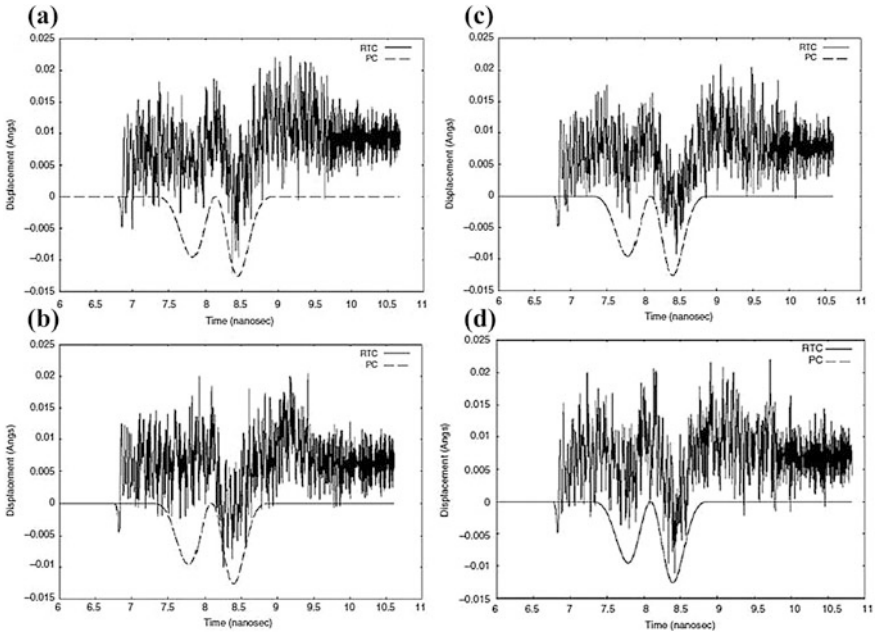

Fig. 6.28 Amplitude of the reflected signals @ 0.393 GHz when N_{md} = **a** 94, **b** 47, **c** 23, **d** 4

of the continuum i.e. the size of the element. The effect of this frequency cut-off will be illustrated further below.

Consider the case of a stand-alone MD simulation at zero external pressure. The internal stress of the MD cell will typically oscillate about zero such that the average stress equals zero if the simulation is run for many MD time steps. The frequency spectrum of the stress for the model stand-alone MD system used in our simulations was determined and is shown in Fig. 6.32. The frequency spectrum of the stand-alone MD cell spans frequencies up to several 10^{12} Hz. As a next step we carried out an FDTD-MD multiscale simulation (with the same conditions used for the wave propagation studies when N_{md} = 4), except for the fact that there was no probing wave packet. The frequency spectrum of the coupled MD cell was obtained and is shown in Fig. 6.33a. The spectrum shows a distinct peak at 160 GHz; the mode of vibration associated with a feedback resonance between the thermalized MD and the FDTD region. This resonant mode of vibration arises because of the way the two regions are coupled (refer to the definition of strain and stress on any element in (6.11)). Here the FDTD cells neighboring the MD cell respond to its stress fluctuation in the form of a non-zero displacement; this displacement in turn strains the MD cell with a subsequent change in internal stress. The resonant frequency is a characteristic of the size of the MD cell and corresponds to a wavelength of about 80 Å, which is about three times the size of an FDTD element. In addition, frequency spectrums of the displacements of elements away from the MD cell were also obtained (Fig. 6.33b). Figure 6.34 clearly shows the fact that as

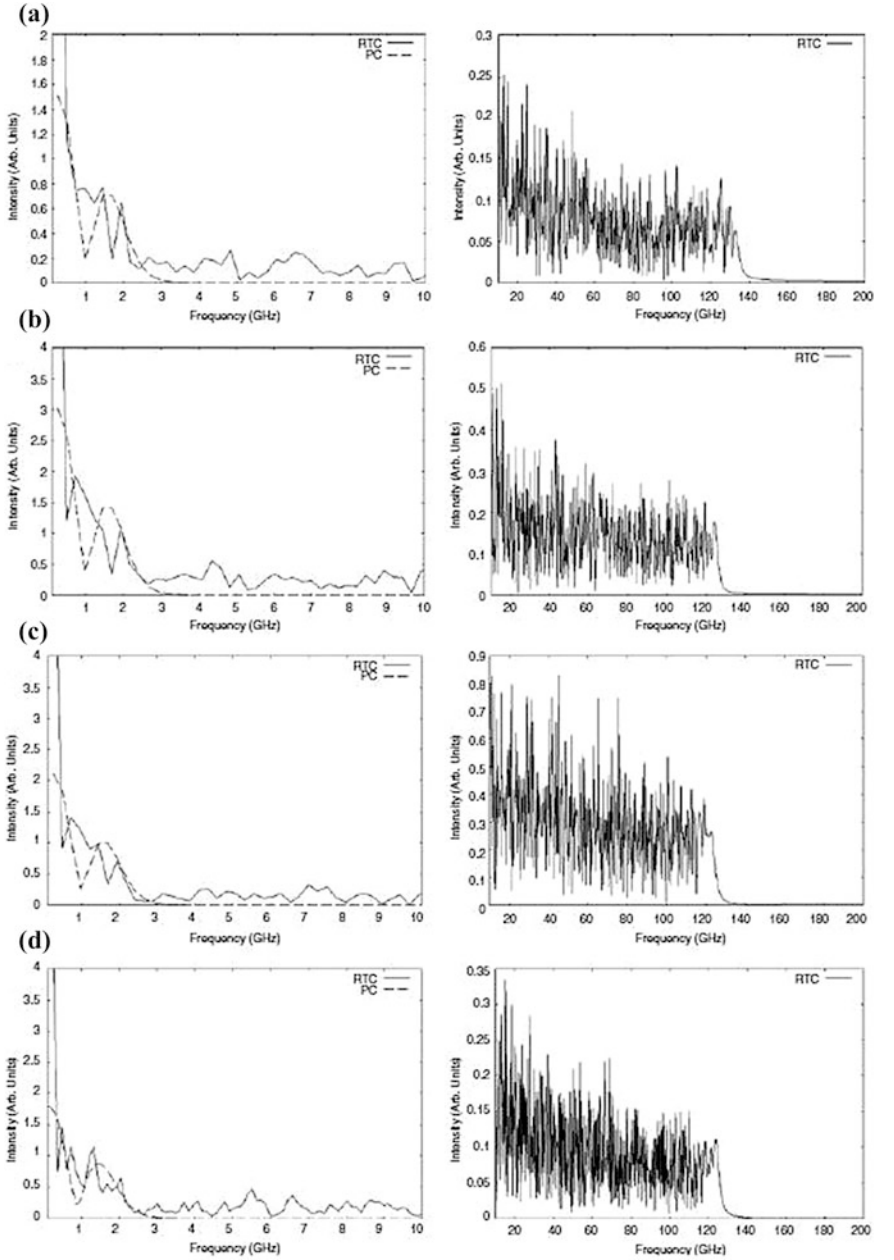

Fig. 6.29 Frequency spectrum of the reflected signals @ 0.393 GHz; **a** $N_{md} = 94$, **b** $N_{md} = 47$, **c** $N_{md} = 23$, **d** $N_{md} = 4$. For each case the figure on the *left* represents the low frequency range of the signal, while the figure on the *right* represents the high frequency range of the signal. Here '*RTC*' corresponds to a real time coupling between the FDTD and the MD region and '*PC*' corresponds to a pseudo coupling between the two regions

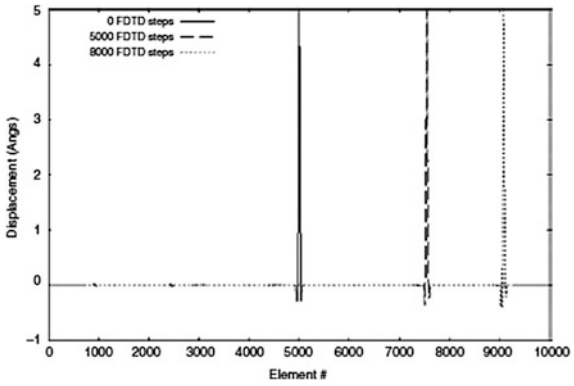

Fig. 6.30 Time evolution of the wave packet @ 3.930 GHz for $N_{md} = 94$

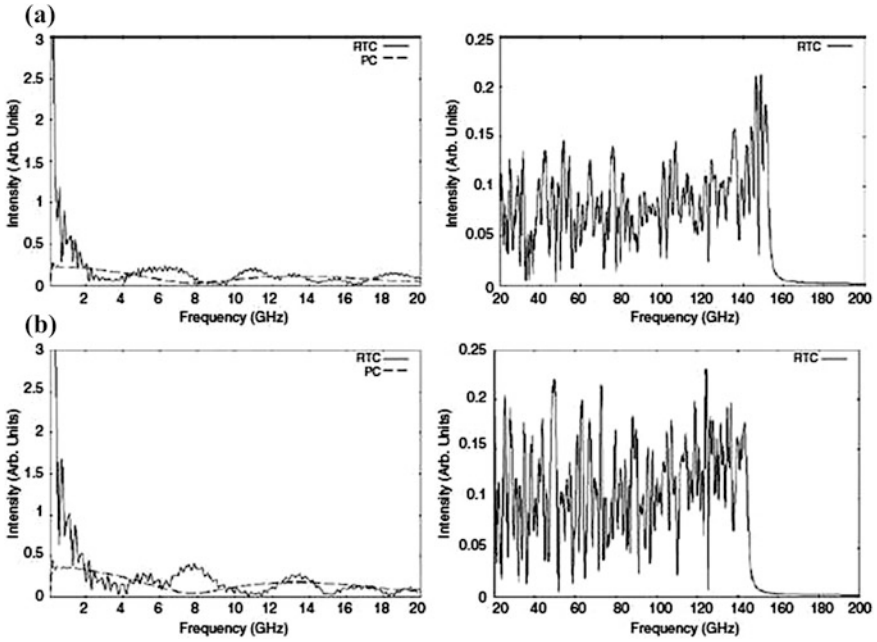

Fig. 6.31 Frequency spectrum of the reflected signals @ 3.930 GHz; **a** $N_{md} = 94$, **b** $N_{md} = 23$. For (**a**) and (**b**) the figure on the *left* represents the low frequency range of the signal, while the figure on the *right* represents the high frequency range of the signal. Here '*RTC*' corresponds to a real time coupling between the FDTD and the MD region and '*PC*' corresponds to a pseudo coupling between the two regions

once moves away from the MD cell, the cut-off moves to lower frequencies suggesting that the discretized FDTD system possesses a Debye-like frequency below 160 GHz. The resonant FDTD-MD vibrational mode appears therefore as a

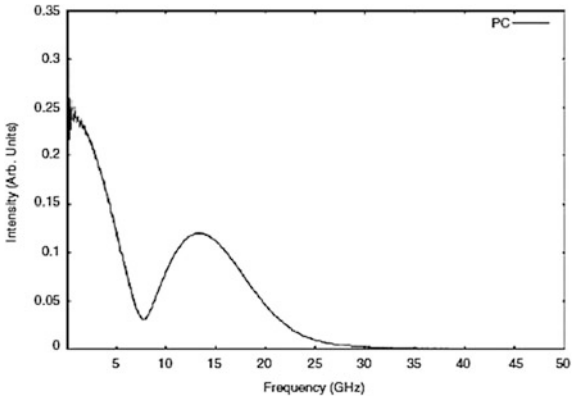

Fig. 6.32 The complete 'PC' frequency spectrum of the reflected signal @ 3.930 GHz for $N_{md} = 94$. 'PC' corresponds to a pseudo coupling between the FDTD and the MD regions. Since the 'PC' signals for every N_{md} are very similar, only a representative case is shown

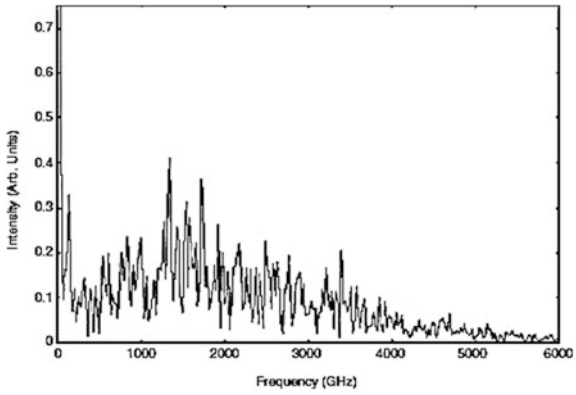

Fig. 6.33 Frequency spectrum of the stress for the model stand-alone MD system

'localized' mode in the vicinity of the MD cell which amplitude decays as one moves away from the MD cell.

Thus, it is clear that during the coupling of atomistic and continuum methods, there exists an elastic impedance mismatch between the two methods. This gives rise to an unavoidable reflection of the probing signal at the atomistic-continuum interface, which is minimized only if the parameters derived from the atomistic methods are obtained at the thermodynamic limit. Further intrinsic thermal effects associated with the atomistic region are always present, leading to localized resonant effects that are dependent on the size of the atomistic region.

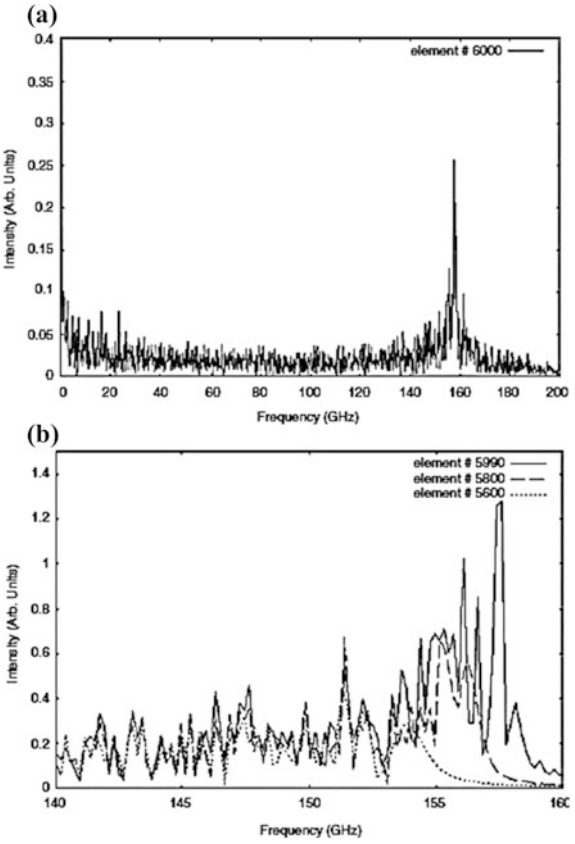

Fig. 6.34 **a** Frequency spectrum of the MD cell (element # 6000) when coupled with FDTD elements. **b** Frequency spectra of FDTD elements at various distances away from the MD cell. The *inset* of the respective figures refers to the location of the elements

In conclusion, using simple illustrative examples, we have examined the consistent embedding computational framework to couple scale-specific and theory-specific methods. In particular we have identified the necessary conditions that have to be satisfied in order to carry out multiscale multi-theory simulations to model properties of materials. Artificial interfaces arise when multi-theory models are spatially coupled in an explicit fashion, and precautionary measures are necessary to minimize the role of these interfaces, in order to ensure the fidelity of the multiscale simulation frameworks. In this context, the compound wavelet matrix (CWM) method offers an attractive route to couple multi-theory methods as long as there are overlapping spatial and temporal scales between the different methods invoked within the multiscale framework. When used with compatible simulation techniques, the consistent embedding framework can balance the demands of physical realism, system size, and computational resource to produce highly accurate simulations for large systems over long times.

References

1. E.B. Tadmor, R. Phillips, M. Ortiz, Langmuir **12**, 4529 (1996)
2. M. Mullins, M.A. Dokanish, Phil. Mag. A **46**, 771 (1982)
3. R.E. Rudd, J.Q. Broughton, Phys. Rev. B **58**, R5893 (1998)
4. H. Kitagawa, A. Nakatami, Y. Sibutani, Mat. Sci. & Eng. A **176**, 263 (1994)
5. J.Q. Broughton, F.F. Abraham, N. Bernstein, E. Kaxiras, Phys. Rev. B **60**, 2391 (1999)
6. S. Kohlhoff, P. Gumbsch, H.F. Fischmeister, Phil. Mag. A **64**, 851 (1991)
7. F.F. Abraham, J.Q. Broughton, N. Bernstein, E. Kaxiras, Comp. Phys. **12**, 538 (1998)
8. S. Ogata, E. Lidorikis, F. Shimojo, A. Nakano, P. Vasishta, R.K. Kalia, Comp. Phys. Comm. **138**, 143 (2001)
9. A. Mallik, D.E. Taylor, K. Runge, J.W. Dufty, Int. J. Quant. Chem. **100**, 1019–1025 (2004)
10. M.J.S. Dewar, W. Thiel, J. Am. Chem. Soc. **99**, 4899 (1977)
11. A. Mallik, K. Runge, H.-P. Cheng, J.W. Dufty, Mol. Simul. **31**, 695 (2005)
12. U.C. Singh, P.A. Kollman, J. Comp. Chem. **7**, 718 (1986)
13. D. Das, K.P. Eurenius, E.M. Billings, D.C. Chatfield, M. Hodos, B.R. Brooks, J. Chem. Phys. **117**, 10534 (2002)
14. M. Swart, Int. J. Quan. Chem. **91**, 177 (2003)
15. M. Eichinger, P. Tavan, J. Hutter, M. Parrinello, J. Chem. Phys. **110**, 10452 (1999)
16. Y. Zhang, T.-S. Lee, W. Yang, J. Chem Phys **110**, 46 (1999)
17. F. Maseras, K. Morokuma, J. Comp. Chem. **16**, 1170 (1995)
18. S. Humbel, S. Sieber, K. Morokuma, J. Chem. Phys. **16**, 1959 (1996)
19. M. Svensson, S. Humbel, R.D.J. Froese, T. Matsubara, S. Sieber, K. Morokuma, Phys. Chem. **100**, 19357 (1998)
20. S. Dapprish, I. Kormaromi, K.S. Byun, K. Morokuma, M.J. Frish, J. Mol. Struct: THEOCHEM **1**, 461 (1999)
21. R. Poteau, I. Ortega, F. Alary, A.R. Solis, J.-C. Barthelat, J.-P. Daudey, J. Phys. Chem. **105**, 198 (2001)
22. A. Arneodo, G. Grasseau, M. Holschneider, Phys. Rev. Lett. **61**, 2281 (1988)
23. L. Telesca, V. Lapenna, N. Alexis, Chaos. Solitons Fractals **22**, 741 (2004)
24. S.N. Zhao, X.Y. Xiong, X.H. Cai, F. Hu, Europhys. Lett. **69**, 81 (2005)
25. D.F. Martin, P. Colella, M. Anghel, F.J. Alexander, Comput. Sci. Eng. **7**, 24 (2005)
26. T. Yanai, G.I. Fann, Z. Gan, R.J. Harrison, G. Beylkin, J. Chem. Phys. **121**, 2866 (2004)
27. G. Frantziskonis, P.A. Deymier, Modeling Simul. Mater. Sci. Eng. **8**, 649–664 (2000)
28. G. Frantziskonis, Probab. Eng. Mech. **17**, 359 (2002)
29. G. Frantziskonis, P. Deymier, Phys. Rev. B **68**, 024105 (2003)
30. G. Frantziskonis, S.K. Mishra, S. Pannala, S. Simunovic, C.S. Daw, P. Nukala, R.O. Fox, P.A. Deymier, Int. J. Multiscale Comp. Eng. **4**, 755 (2006)
31. C.K. Chui, *An Introduction to Wavelets* (Academic Press, London, 1992)
32. I. Daubechies, Commun. Pure Appl. Math. **41**, 909 (1988)
33. I. Daubechies, *Ten Lectures on Wavelets CIAM* (Philadelphia, PA, 1992)
34. S.G. Mallat, I.E.E.E. Trans, Pattern Anal. Mach. Intell. **11**, 674 (1989)
35. A. Harbo-Cour, A. Jensen, *Ripples in Mathematics: Discrete Wavelet transform* (Springer, New York, 2000)
36. G.B. Olson, Science **277**, 1237 (1997)
37. P.S. Sahni, G.S. Grest, M.P. Anderson, D.J. Srolovitz, Phys. Rev. Lett. **50**, 263 (1983)
38. L.Q. Chen, Scripta Metall. Mater. **32**, 115 (1995)
39. A. Soares, A.C. Ferro, M.A. Fortes, Scripta Metall. **19**, 1491 (1985)
40. D.W. Heermann, *Computer Simulations Methods in Theoretical Physics* (Springer-Verlag, Berlin, 1986)
41. E.A. Holm, J.A. Glazier, D.J. Srolovitz, G.S. Grest, Phys. Rev. A **43**, 2662 (1991)
42. M.P. Anderson, D.J. Srolovitz, G.S. Grest, P.S. Sahni, Acta Metall. **32**, 783 (1984)
43. J. Gao, R.G. Thompson, Acta Mater. **44**, 4565 (1996)

44. D.T. Gillespie, J. Phys. Chem. **81**, 2340 (1977)
45. L. Baffico, S. Bernard, Y. Maday, G. Turinici, G. Zerah, Phys. Rev. E **66**, 057701 (2002)
46. G. Frantziskonis, K. Muralidharan, P. Deymier, S. Simunovic, P. Nukala, S. Pannala, J. Comput. Phys. **228**, 8085 (2009)
47. C. Kittel, *Introduction to Solid State Physics 5th Edition* (John Wiley and Sons, New York, 1976), p. 111
48. M.M. Sigalas, N. Garcia, J. Appl. Phys. **87**, 32122 (2000)
49. D. Garcia-Pablos, M.M. Sigilas, F.R.M. de Espinosa, M. Torres, M. Kafesaki, N. Garcia, Phys. Rev. Lett. **84**, 4339 (2000)
50. D.C. Rapaport, *The Art of Molecular Dynamics Simulation* (Cambridge University Press, Cambridge, 1995)
51. M. Parrinello, A. Rahman, J. Appl. Phys. **52**, 7182 (1981)
52. W.H. Press, S.A. Teulosky, W.T. Vetterling, B.P. Flannery, *Numerical Recipes in Fortran, The Art of Scientific Computing Second Edition* (Cambridge University Press, Cambridge, 1992), pp. 501–502

Index

P.A. Deymier et al. (eds.), *Multiscale Paradigms in Integrated Computational Materials Science and Engineering*, Springer Series in Materials Science 226,
DOI 10.1007/978-3-319-24529-4