

Low-Dimensional Representations of Transitions in Molecular Systems

Thesis by
Katalin A. Grubits

In Partial Fulfillment of the Requirements
for the Degree of
Doctor of Philosophy



California Institute of Technology
Pasadena, California

2008
(Defended June 4, 2008)

© 2008

Katalin A. Grubits

All Rights Reserved

Acknowledgements

I would like to thank my advisor Jerry Marsden for his support and guidance. His love of science and curiosity about nature is infectious and inspirational.

Thanks also go to my collaborators: Sorin Costiner, Michael Dellnitz, Philip du Toit, Frederic Gabern, Jerrold Marsden, Kathrin Padberg, Bianca Thiere, and Tomohiro Yanao. They have taught me a lot and I have found it a pleasure to work with them. Chapter 2 has been previously published as [23]. The work in Chapter 3 was done in collaboration with Tomohiro Yanao. Jimmy Fung provided help with the simulation code. Chapter 4 and parts of Chapter 5 will be published as [34]. Most of Section 5.1 will appear as [16].

On a personal note, thanks to my parents for their love and support, which reaches across the ocean separating us. I can not imagine my time at Caltech without Lotty. She has been there since the beginning, through everything; her generosity and compassion are without bound. Finally, thanks to my fiancé, Jon, who has made me happier than I ever thought possible.

Abstract

A major difficulty in modeling molecular systems is that the number of dimensions, even for a small system, is commonly too large for computation to be feasible. To overcome this challenge, a combination of lower-dimensional representations of the system and improved computational methods are required. In this thesis, we investigate techniques to achieve both of these aims via three model problems.

By exploiting an understanding of the mechanism and dynamics of reaction in the systems considered, we attain a low-dimensional description of the transition that captures the essential dynamics. For the ionization of a Rydberg atom we utilize concepts from dynamical systems theory that reveal the geometric structures in phase space that mediate the reaction. The gyration radius formalism captures the kinematic effects of the secondary particles in a coarse variable that reduces the number of dimensions of the model, thereby providing a simple description of our methane and oxygen dissociation example. These methods are applicable more generally and provide a coarse model of chemical reactions that can be combined with efficient computational tools, such as the set-oriented method employed in our Rydberg example, to efficiently compute reaction rates of previously difficult problems.

The third model problem considered is the self-assembly of particles into an ordered lattice configuration under the influence of an isotropic inter-particle potential. With the aim of characterizing the transition from a disordered to an ordered state, we develop metrics that assess the quality of the lattice with respect to the target lattice configuration. The five metrics presented use a single number to quantify the order within this large system of particles. We explore numerous applications of these quality assessment tools, in particular, finding the optimal potential for self-assembly. The very noisy, highly variable nature of our expensive-to-evaluate objective function prompted the development of a trend optimization algorithm that efficiently minimizes the objective function, using upper and

lower envelopes that are responsible for the robustness of the method and the solution. This trend optimization scheme is widely applicable to problems in other fields.

Contents

Acknowledgements	iii
Abstract	v
List of Figures	x
1 Introduction	1
2 Set-Oriented Computation of Transport Rates in 3-Degree-of-Freedom Systems: The Rydberg Atom in Crossed Fields	5
2.1 Background	7
2.2 Model	8
2.2.1 Half and Full Scattering Problem	8
2.2.2 The Hamiltonian System	9
2.2.3 Dynamics Near the Saddle Equilibrium Point	10
2.2.4 The Poincaré Section	11
2.3 Set-Oriented Methods	13
2.3.1 Computation of Tube Intersections	14
2.3.2 Transport Rates	18
2.3.3 Implementation	23
2.4 Examples	24
2.4.1 Full Scattering Problem for the 2- and 3-Degree-of-Freedom System ($\epsilon = 0.57765$)	24
2.4.2 Full and Half Scattering Problem for the 3-Degree-of-Freedom System ($\epsilon = 0.58$)	27
2.5 Conclusion	28

3	Energy Effects and Gyration Radius Dynamics in the Dissociation of Methane and Oxygen	31
3.1	Model	32
3.2	Kinetic Energy Bias Method	34
3.2.1	Theory	35
3.2.2	Effective Potential	36
3.2.3	Energy Distribution in the System	39
3.2.3.1	Ratio of Average Kinetic Energies	39
3.2.3.2	α Tubes	40
3.2.3.3	Equipartition of Energy	41
3.2.4	Discussion	41
3.3	Gyration Radius Method	42
3.3.1	Background	43
3.3.2	Theory	44
3.3.2.1	Collinear System	48
3.3.3	Systems with Different Total Energies	50
3.3.3.1	Distributions for Different Energies	50
3.3.3.2	Effective Potential	52
3.3.3.3	Average Lifetime Before Dissociation	55
3.3.4	Systems with Different Secondary Particle Masses	56
3.3.4.1	Distributions for Different Masses	57
3.3.4.2	Effective Potential	57
3.3.4.3	Average Lifetime Before Dissociation	60
3.3.5	The Internal Centrifugal Term in a Simpler System	64
3.4	Conclusion	65
4	Lattice Quality Assessment Tools	69
4.1	Types of Lattice Defects	70
4.2	Metrics for Assessing the Quality of Lattices	71
4.2.1	Defect Measure	71
4.2.1.1	Procedure for Identifying the Target Lattice and Computing the Defect Measure	72

4.2.1.2	Advantages of the Defect Measure	76
4.2.1.3	Identifying the Boundary of a Lattice	77
4.2.2	Geometric Defect Measure	80
4.2.3	Local Template Measure	82
4.2.4	Voronoi Metric	84
4.2.5	Cumulative Distribution Function Metric	87
4.3	Comparison of Quality Metrics	89
4.4	Conclusion	96

5 Applications of the Defect Measure and Other Lattice Quality Assessment

Tools		99
5.1	Generation of Potentials for the Self-Assembly of Particles	100
5.1.1	Background	101
5.1.1.1	Baseline Geometric Method	102
5.1.2	Design of Potentials for Self-Assembly	103
5.1.3	Optimization of Potential for Self-Assembly	109
5.1.4	Trend Optimization of Potential for Self-Assembly	111
5.1.4.1	Objective Function Evaluation	112
5.1.4.2	Objective Function Difficulties	112
5.1.4.3	A Brief Survey of Similar Optimization Techniques	113
5.1.4.4	General Trend Optimization Algorithm	116
5.1.4.5	Hierarchical Nature of the Trend Optimization Scheme	118
5.1.4.6	Initial Investigations	119
5.1.4.7	Comparison with Simulated Annealing	122
5.1.4.8	Results	122
5.1.5	Quantifying Robustness of Potentials	125
5.1.5.1	Equally Good Potentials	125
5.1.5.2	Acceptably Good Lattices	125
5.2	Construction of the Phase Diagram of Krypton Atoms on a Graphite Substrate	127
5.2.1	Low Temperature Region	128
5.2.2	High Temperature Region	131
5.3	Identification of Global Defects	134

5.4	Identification of Clusters to be treated as Rigid Bodies	136
5.5	Conclusions and Future Directions	137
6	Conclusions and Future Directions	141
A	Lattices Used for Comparison in Section 4.3	145
	Bibliography	145

List of Figures

2.1	Schematic of the energetically allowed regions and types of trajectories for the Rydberg atom	10
2.2	Typical trajectories	11
2.3	Chaotic sea and tube intersections, $\epsilon = 0.57765$	13
2.4	Chaotic sea and tube intersections, $\epsilon = 0.58$	13
2.5	First return time distribution for $\epsilon = 0.58$	15
2.6	First intersection of the stable and unstable tubes with the Poincaré section .	18
2.7	Boxes that contribute to the error between the computed and the actual value of the transport rate	21
2.8	Adaptive box covering for the Rydberg atom in crossed fields	25
2.9	Lower and upper bounds for the transport rate in the full scattering, 2-degree-of-freedom, $\epsilon = 0.57765$ problem	26
2.10	Transport rates for the full scattering, $\epsilon = 0.57765$ problem	27
2.11	Transport rates for the full and half scattering, 3-degree-of-freedom, $\epsilon = 0.58$ problem	28
3.1	Schematic of the model problem	33
3.2	Initial conditions and typical trajectories for the model dissociation system .	34
3.3	The effective potential calculated by the Kinetic Energy Bias method	36
3.4	Comparison of distribution in time and effective potential	37
3.5	Comparison of distribution in space and effective potential	38
3.6	Ratio of bath to primary particle kinetic energy	40
3.7	Visualization of the bottlenecks experienced by the dissociating particles . .	41
3.8	Kinetic energy of primary and secondary particles	42
3.9	Typical trajectories along the gyration radius	50

3.10	Distributions in time and space along r and a for different energies	51
3.11	Internal centrifugal force and potential force for different energies	53
3.12	Potential due to the internal centrifugal force and the potential force	54
3.13	Effective potential for different energies	54
3.14	Comparison of effective potential and distribution in space for different energies	55
3.15	Distribution of times of last bounce, t_{last} , of the system before dissociating for different energies.	56
3.16	Distributions in time and space along r and a for different masses	58
3.17	Internal centrifugal force and potential force for different masses	58
3.18	Potential due to the internal centrifugal force and potential force for different masses	59
3.19	Effective potential of the system for different masses.	59
3.20	Comparison of the effective potential and distribution in space for different masses	60
3.21	Dissociation times for different masses	61
3.22	Variation in reactivity with mass and spring constant	62
3.23	Variation in reactivity with k/m_s	63
3.24	Distributions, internal centrifugal force, and internal centrifugal potential for different energies in a system with two secondary particles.	64
4.1	Nearest neighbors circle	74
4.2	Types of defects used in computing the Defect Measure.	75
4.3	Boundary particle in a perfect triangular lattice	78
4.4	The convex hull of a set of points in the plane	78
4.5	Boundary points, as identified by the non-convex hull	79
4.6	Construction used in computing the Geometric Defect Measure for particle p .	80
4.7	Illustration of the Local Template Measure	83
4.8	Voronoi tessellation of a set of points in the plane	84
4.9	Cumulative Distribution Function Metric	88
4.10	Comparison of quality assessment metrics	90
4.11	Best lattices	91
4.12	Comparison of quality assessment metrics (extended)	93

4.13	Worst lattices	95
4.14	CDF Metric's two best lattices	95
5.1	(a) Perfect honeycomb lattice. (b) Potentials for the self-assembly of particles into a honeycomb lattice	104
5.2	Defect Measure versus parameter a_2 in Equation 5.2	105
5.3	Example lattices from Figure 5.2	106
5.4	Geometric Defect Measure versus parameter a_2 in Equation 5.2	107
5.5	The two components of the Geometric Defect Measure	107
5.6	Search over the parameters in the self-assembly potential Equation 5.2 using the Defect Measure to assess lattice quality	108
5.7	The Defect Measure over the parameters a_0, a_1, a_2 in Equation 5.2	109
5.8	Defect Measure versus time for a single simulation.	111
5.9	Schematic of the trend and upper and lower envelopes for a noisy objective function with a simple trend	120
5.10	(a) Defect Measure versus CPU time for trend optimization. (b) Comparison of trend optimization and simulated annealing	123
5.11	Potentials for the self-assembly of particles into a honeycomb lattice	124
5.12	Defect Measure versus initial density of particles	126
5.13	Schematic illustration of the temperature quench method	128
5.14	A typical snapshot of krypton atoms on a graphite substrate at 70K	129
5.15	Particles and Voronoi cells used for the fluid density over estimate	131
5.16	Motion of krypton atoms around their adsorption sites on a graphite substrate	132
5.17	Distribution of Defect Measure values used to determine fluid and solid states	133
5.18	Phase diagram for krypton atoms on a graphite substrate found by using the Defect Measure	134
5.19	Distribution of the area of Voronoi cells	135
5.20	Identification of global and local defects using the Voronoi metric	136

Chapter 1

Introduction

A major difficulty in modeling molecular systems is that the number of dimensions for even a small system is commonly too large for computation to be feasible. To overcome this challenge, a combination of lower-dimensional representations of the system and improved computational methods are required. Understanding the dynamics of a transition enables us to focus on mechanisms that mediate transport and to develop coarse models that have fewer dimensions than the original system, yet contain the essential dynamics.

These lower-dimensional models of chemical reaction dynamics can be used to calculate transition rates when coupled with efficient methods of computation. An efficient computational method will exploit global features of the domain, and locally refine areas of interest, rather than doing a brute force computation at every point in the domain.

Together, these approaches make the calculation of reaction rates in realistic chemical systems tractable. The predictions of computational reduced-order models can be compared with experiments to gain further understanding of the mechanism of reaction, thereby opening the door to the design of new reaction pathways and products.

In this thesis we develop low-dimensional representations of a number of chemical systems as well as implement efficient algorithms for computation. These tools are presented in the context of model problems that are sophisticated enough to capture phenomena that are, or are expected to soon be, experimentally verifiable.

In Chapter 2 we calculate reaction rates by coupling dynamical systems theory and a set-oriented method for computation. The number of dimensions to be modeled is reduced by introducing a surface of section through which all trajectories must pass and exploiting knowledge of the geometric structures in phase space that mediate the reaction. Transition

rates are calculated by looking at a global description of transport in the system. This is possible through the set-oriented approach, which constructs a box covering of the surface of section and computes many short trajectories in each box, rather than following individual trajectories for a long time. This is an efficient computational method for finding global structures and calculating transport rates; it uses local refinements of the box covering to ensure the accuracy of the computed rates. The method is demonstrated with the ionization of a Rydberg atom in crossed electric and magnetic fields, an example for which traditional transition state theory fails.

Chapter 3 investigates a system in which (primary) particles undergo a transition under the influence of other particles (secondary particles). The energy distribution between the reactive and bath modes is demonstrated to be important to an understanding of the system. A coarse variable, called the gyration radius, describes the dynamics of reaction in a succinct and elegant way. An effective potential computed along this gyration radius takes into account the kinematic effects of the secondary particles and is consistent with the observed probability distributions. This effective potential includes an internal centrifugal term that is shown to govern the mechanism of reaction for some system parameters. The internal centrifugal force couples the kinematics of the bath particles with the gyration radius, which encapsulates the mass distribution of the system. The system we consider is a simplified model of the dissociation of methane and oxygen. The coarse representation of the system by the gyration radius and accompanying effective potential can be combined with the set-oriented computational method to calculate dissociation rates for realistic chemical reactions.

In the second half of the thesis we move away from chemical reactions with a small number of constituents and instead focus on a large number of particles in the plane and their transition from a disordered state to an ordered state. In Chapter 4 we develop and discuss five metrics for assessing the quality of ordered lattice configurations. These quality assessment tools are a simple representation of the system that can capture phase transitions that involve a change in order. They are compared and their differences highlighted, noting the situations in which each is applicable. The best metric to use depends on whether the alignment of particles, the local density, or the correct inter-particle distance is the most important.

Chapter 5 explores some of the many applications of the Defect Measure and other lattice

quality assessment tools. The problems examined here are the generation of inter-particle potentials for the self-assembly of particles, examining the robustness of such potentials, finding the phase diagram of krypton atoms on a graphite substrate, and identifying global defects and clusters from local information. A computationally efficient method, called trend optimization, for optimizing the parameters of self-assembly potentials is presented and contrasted with simulated annealing. The trend optimization scheme is ideal for finding minima of very noisy (or discontinuous), highly variable, expensive-to-evaluate objective functions and quantifying the robustness of solutions. It is also very fast and robust. It is the combination of being able to represent the quality of a lattice with one number and this efficient trend optimization method that enables us to find the optimal inter-particle potential for self-assembling systems.

Chapter 2

Set-Oriented Computation of Transport Rates in 3-Degree-of-Freedom Systems: The Rydberg Atom in Crossed Fields

One of the primary goals of chemical physics is the calculation of the rate at which a reaction proceeds. Transition State Theory (TST) (see, for example, [84]), also known as Rice-Ramsperger-Kassel-Marcus (RRKM) theory (see, for example, [33]) is widely used in the chemistry community to calculate these rates. While successful in many situations, this statistical theory is inadequate in others, and in those cases it can have an error of a few orders of magnitude when compared with experimental results [18].

TST identifies a transition state for the system under consideration: This is a set of states through which the reactants must pass in order to become products of the reaction. These transition states may be in phase space rather than configuration space, but TST assumes that the regions in phase space connected by this transition state are structureless in the sense that motion within them is purely statistical [62]. However, in the examples where TST fails, this assumption breaks down, and indeed, the structure of phase space must be accounted for when calculating reaction rates [31].

By developing reaction island theory, De Leon, Mehta, and Topper have shown that cylindrical manifolds in phase space mediate 2-degree-of-freedom chemical reactions [19, 20]. Uzer, Jaffé, and co-workers have isolated some of the important geometrical aspects of the phase space structure for higher degree-of-freedom systems [44, 85]. We note that Koon

et al. [51] emphasized the importance of heteroclinic networks and the associated cylindrical manifolds (tubes) when considering dynamical channels, and Contopoulos and Efstathiou [15] used escape rates from a surface of section to identify regions that govern the transport between parts of the phase space.

Gabern et al. [31] have calculated reaction rates in chemical systems with three degrees of freedom using dynamical systems tools and Monte Carlo methods. By taking into account the invariant manifold tubes that mediate the dynamics of a reaction, these rates were calculated for a system with non-statistical dynamics. A major difficulty that was overcome by using a Monte Carlo method was the calculation of the volume of the overlap of the invariant manifold tubes.

This work uses a new approach, based on set-oriented methods (see, for example, [26]), to identify the structures in phase space that mediate chemical reactions and to calculate the associated reaction rates.

The set-oriented approach focuses on a global description of the dynamics on a coarse level and covers the relevant region of phase space by appropriately sized boxes. By considering a transfer operator associated with the underlying map, one is able to describe the evolution of an initial distribution under the dynamics. Via a partition of some interesting region in phase space, this operator can be discretized, yielding a stochastic matrix. The transport rates between different regions in phase space can then be computed using this matrix of transition probabilities. This global analysis is more efficient and can provide more information than the calculation of many individual trajectories.

The primary differences between the approach presented in this chapter and that of [31] are that the set-oriented method

1. does not use normal forms to find the invariant manifold tubes but rather uses information about the time trajectories take to return to a Poincaré section,
2. does not use Monte Carlo methods for the calculation of volumes, as the necessary information is naturally given by the box volumes and the matrix of transition probabilities, and
3. does not use long-term simulations but rather short-term simulations for a large number of globally distributed initial particles.

Despite the large differences in methodology and computational tools, the results of the set-oriented approach and that of [31] are in good agreement, which gives us confidence in both methods.

We take the ionization of a Rydberg atom in crossed electric and magnetic fields as our example. Both the planar problem and the three-dimensional problem are considered, with the half scattering and full scattering rates being calculated. The power and the potential of the set-oriented approach in dealing with high-dimensional systems is thereby demonstrated.

In the following section, the physical background of the example that we consider is presented, followed by a detailed description of the model. Section 2.3 elucidates the set-oriented method as it relates to the calculation of reaction rates. In Section 2.4 the results are presented and discussed, followed in Section 2.5 by conclusions and future directions.

2.1 Background

The Rydberg atom is a hydrogen-like atom in that it has one valence electron. Highly excited Rydberg atoms have enough energy such that the valence electron is far away from the nucleus and its dynamics can be treated classically, to a good approximation. Introducing external perpendicular electric and magnetic fields breaks the symmetry of the problem so that the escaping electron will do so in a particular direction. The escape of the electron from the field of the nucleus (and surrounding inner electrons) is known as ionization. The electron moves off to infinity and there is no possibility of return. This process is an example of a unimolecular reaction or dissociation.

The highly excited Rydberg atom is an interesting example not only because of its relation to other problems in chemical physics but also because of applications in diverse areas ranging from lasers to quantum computing [29, 87, 1]. They are also of interest as they are at the overlapping region between classical and quantum mechanics, where the correspondence principle applies [73]. In addition to their theoretical interest, such atoms in crossed fields arise naturally in some astrophysical plasmas.

Rydberg atoms are a compelling test bed as they have a theoretical richness while also being experimentally accessible. Such atoms have been used to study the onset of classical chaos and to develop semi-classical models of quantum resonances [57, 17]. They are well

suited to experiments as the internal field strengths of the atom are comparable to the external field strengths that are attainable in the laboratory [72]. Thus it is possible to study the strong-field regime.

Raithel, Walther, and co-workers have studied Rydberg atoms in a number of arrangements, including the crossed fields arrangement. They have calculated ionization rates as a function of excitation energy for different values of the electric and magnetic fields [73]. Advances in experimental methods now allow the excitation of a Rydberg atom to a known energy level [39]. Thus, the techniques are available for experimentally calculating the ionization rates that are computed in this chapter. It is hoped that the explicit experimental connection is achieved in the near future.

2.2 Model

2.2.1 Half and Full Scattering Problem

In a unimolecular dissociation reaction, the reactant is the bound state and the product the unbound state. To pass from a bound state to an unbound state, the system must go through the transition state. Such reactions have come to be known as *half scattering problems* [44]. Once in the unbound state, there is no possibility of return for the electron. The *full scattering problem* involves moving through the transition state from an unbound state to a bound state and then back through the transition state to an unbound state, that is, the *capture* of the electron followed by *ionization* of the same electron. The example discussed in Section 2.4 calculates rates of reaction for both the half scattering and full scattering problems.

The reaction will proceed only if the system has enough energy to overcome the energy barrier between reactants and products. For an energy at which the reaction can proceed, the energy in the system must find its way into a reactive mode for the reaction to occur. It is this process which determines the rate of the reaction.

2.2.2 The Hamiltonian System

The dynamics of the electron is described classically by the following 3-degree-of-freedom Hamiltonian in co-ordinates that have been scaled by the cyclotron frequency:

$$H = \frac{1}{2}(p_x^2 + p_y^2 + p_z^2) - \frac{1}{r} + \frac{1}{2}(xp_y - yp_x) + \frac{1}{8}(x^2 + y^2) - \epsilon x, \quad (2.1)$$

where $r = \sqrt{x^2 + y^2 + z^2}$ is the distance between the electron and the center of the nuclear core. The cyclotron frequency, ω_c , is given by $\omega_c = eB/m$ where e is the electron charge, B is the magnetic field strength and m is the mass of the electron. The scaled electric field strength, ϵ , is defined by $\epsilon = \omega_c^{-4/3}E$ where E is the applied electric field strength (see for example [44]).

The Legendre transformation gives us the velocities

$$\dot{x} = p_x - \frac{y}{2}, \quad \dot{y} = p_y + \frac{x}{2}, \quad \dot{z} = p_z.$$

The Jacobi constant (first integral) is given by

$$C(x, y, z, \dot{x}, \dot{y}, \dot{z}) = -(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) + 2\Omega(x, y, z) = -2E(x, y, z, \dot{x}, \dot{y}, \dot{z})$$

where

$$E(x, y, z, \dot{x}, \dot{y}, \dot{z}) = \frac{1}{2}(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) - \Omega(x, y, z),$$

is the energy function. The effective potential function is $\Omega(x, y, z) = \epsilon x + \frac{1}{r}$. The 2-degree-of-freedom system is obtained by setting $z = \dot{z} = 0$ in the equations above.

The Stark saddle point occurs at

$$\begin{aligned} x &= \frac{1}{\sqrt{\epsilon}}, & y &= 0, & z &= 0, \\ \dot{x} &= 0, & \dot{y} &= 0, & \dot{z} &= 0. \end{aligned}$$

The Hill's region is the region of configuration space in which the electron is energetically forbidden to go, and for $\dot{x} = \dot{y} = \dot{z} = 0$ is given by

$$M(\epsilon, C) = \{(x, y, z) \in \mathbb{R}^3 \mid \Omega(x, y, z) \geq C/2\}.$$

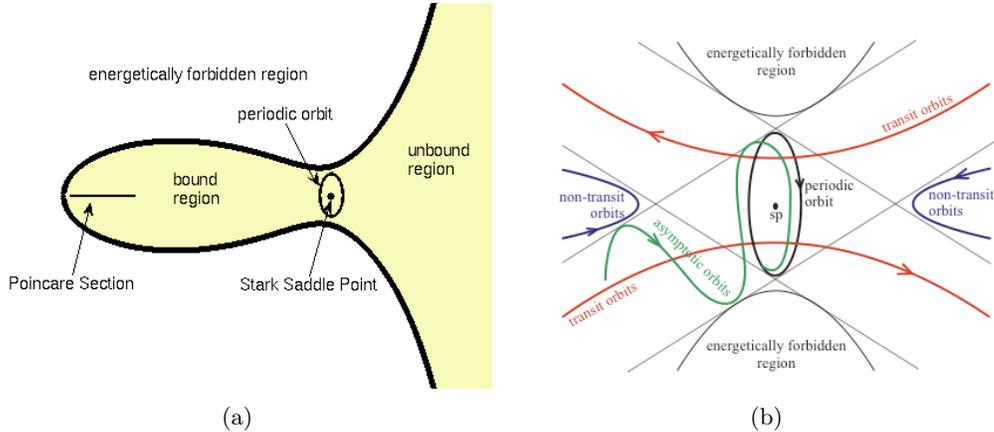


Figure 2.1: (a) xy projection of the Hill’s region (schematic). (b) The possible types of trajectories in the equilibrium region are shown in the xy projection (schematic). There are three different types of orbits — asymptotic, transit, and non-transit orbits (see [51]).

Figure 2.1(a) shows one of the possible cases of the Hill’s region for the Rydberg atom projected onto the xy -plane. For some values of the energy (and hence C), the energetically forbidden region will be such that there is no way to go from a bound state to an unbound state or vice versa. That is, the neck region in the figure will close. Thus, for scattering problems we must have sufficient energy for the Hill’s region to look qualitatively as shown in Figure 2.1(a).

2.2.3 Dynamics Near the Saddle Equilibrium Point

For the computation of transition probabilities, we need to identify regions in phase space that correspond to transport regions. In Transition State Theory the phase space associated with the reaction is traditionally assumed to be structureless [62]. Jaffé et al. [44], as well as Gabern et al. [31], have shown that this is not true for the problem of a Rydberg atom in crossed electric and magnetic fields. Their work builds on the work of Conley [14], McGehee [63], and Koon et al. [51], which have shown this to also be true for the restricted three body problem in celestial mechanics.

For a system with n degrees of freedom, there is an invariant deformed $(2n - 3)$ -sphere that is the normally hyperbolic invariant manifold (NHIM), near the rank-one saddle equilibrium point sp . Orbits asymptotic to this sphere form the stable and unstable manifolds of the sphere. These manifolds (also called “tubes”) are the key features of phase space

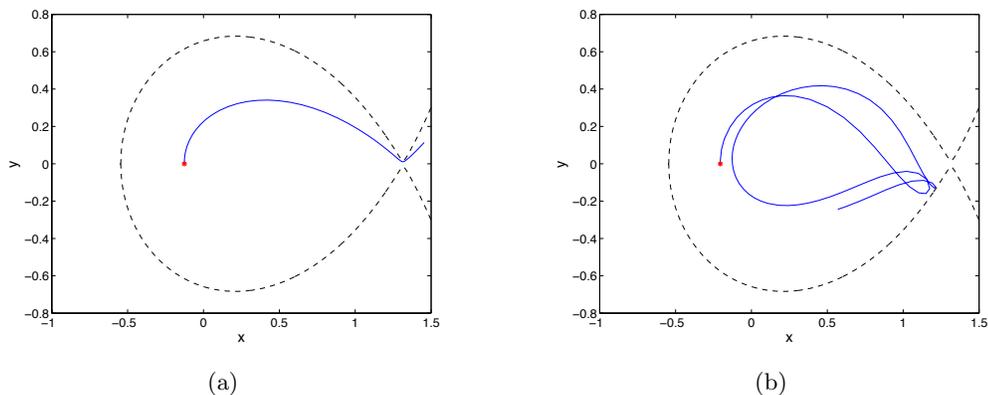


Figure 2.2: Typical (a) transit and (b) nontransit trajectories. The xy projection of the Hill's region is shown as the dotted line. Compare with Figure 2.1.

that mediate transport through the transition state. As these tubes are $(2n - 2)$ dimensional objects in a $(2n - 1)$ dimensional energy surface, they divide the possible orbits into two categories: those that will pass through the transition state and those that will not. Orbits inside the stable manifold in the interior of the atom will pass through the transition state, that is, particles on these orbits will react. Figure 2.1(b) shows the possible types of trajectories in the region near the equilibrium point sp .

2.2.4 The Poincaré Section

To reduce the dimensionality, an appropriately chosen $(2n - 2)$ -dimensional Poincaré section is taken in the $(2n - 1)$ -dimensional energy surface. In our context of the 2- and 3-degree-of-freedom system we choose a Poincaré section Σ given by the conditions

$$y = 0, \quad x < 0, \quad \dot{y} > 0. \quad (2.2)$$

All of the essential dynamics are captured by the Poincaré section as trajectories will cross the Poincaré section only once in every loop that they make about the nuclear core at the origin. Thus we focus our attention on the dynamics on the Poincaré section.

In Figure 2.2 the xy -projection of typical (a) transit and (b) nontransit trajectories are shown. The starting point of the trajectory in (a) lies in the interior of the stable manifold tube and leaves the Poincaré section directly, whereas the starting point in (b) lies somewhere in the chaotic sea and comes back to Σ several times.

The first intersection of the unstable manifold tube with the Poincaré section contains those orbits that have just passed through the transition state from the unbound to the bound state. The forward mapping of these orbits under the Poincaré return map designates successive intersections of the unstable manifold tube with the Poincaré section. The first intersection of the stable manifold tube with the Poincaré section contains orbits that are about to pass through the transition state from the bound state to the unbound state. The m -th pre-image of this intersection under the Poincaré return map designates orbits that will pass through the transition state after m iterations. Thus in order to calculate rates of reaction for the half scattering problem, it is sufficient to find the transport probability into these intersections of the stable tube with the Poincaré section. For the computation of transport rates in the full scattering problem it is necessary to calculate transition probabilities between intersections of the stable and unstable manifold tubes with the Poincaré section. The methodology for these computations is explained in Section 2.3.

All of our computations were with a fixed energy level of $E = -1.52$, which corresponds to a Jacobi constant of $C = 3.04$. With this value of energy, a reaction will be able to proceed if the electric field parameter is greater than $\epsilon = 0.5776$. If the scaled electric field is less than this value, the neck region between the bound and unbound states will be closed.

Figures 2.3 and 2.4 show the chaotic sea for the 2-degree-of-freedom Rydberg atom together with intersections of the stable (blue) and unstable (red) tube boundaries. The electric field parameter is $\epsilon = 0.57765$ for Figure 2.3, which is just above the critical value. Figure 2.3(a) shows the first six intersections of the stable (blue) and unstable (red) tube boundaries with Σ , and Figure 2.3(b) focuses on the region of interest. These tube intersections are very thin in comparison to the intersections of the tube boundaries for $\epsilon = 0.58$ shown in Figure 2.4.

The black dots in these diagrams represent trajectories crossing Σ . The same number of iterates and the same initial conditions were used for both values of ϵ . In both diagrams, the inside of the first intersection of the unstable tube with Σ is white because particles of this region will be mapped out of this region under one iteration of the map and no particles of the initial distribution will be mapped into this region. For an electric field parameter $\epsilon = 0.57765$, if a particle's trajectory begins in the unstable (red) tube, it will take five iterations before it could possibly be in the stable (blue) tube. Thus in the full scattering problem, once an electron has been captured, it will make five loops about the nuclear core

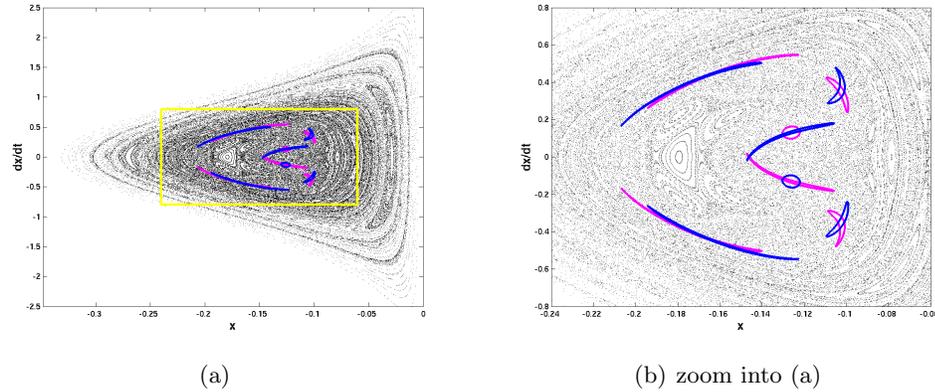


Figure 2.3: Chaotic sea for the 2-degree-of-freedom Rydberg atom is shown with the first six intersections of the unstable (red) and stable (blue) tube boundaries with the Poincaré section Σ under consideration. For this electric field parameter of $\epsilon = 0.57765$, the neck between the bound and unbound region is only open a little.

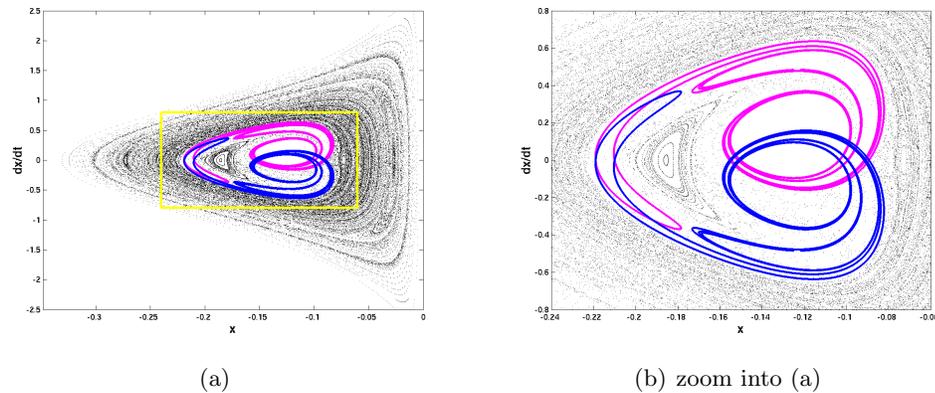


Figure 2.4: Chaotic sea for the 2-degree-of-freedom Rydberg atom for an electric field parameter of $\epsilon = 0.58$. The first three intersections of the unstable (red) and stable (blue) tube boundaries with the Poincaré section Σ under consideration are shown in (a), whereas (b) is a zoom into the interesting region of (a).

before it could possibly leave the atom. For an electric field parameter $\epsilon = 0.58$, the first unstable tube intersection with Σ already overlaps the first stable tube intersection.

2.3 Set-Oriented Methods

In this section we describe the general methodology for the computation of transition probabilities. We first introduce a method for the identification of the regions we are interested

in. We then discuss a technique for the computation of transport rates and probabilities. It makes use of an appropriate discretization of a transfer operator. Both of these methods are based on the set-oriented approach (see, for instance, [24, 25, 26]).

2.3.1 Computation of Tube Intersections

As mentioned earlier, to compute the transition rate for the half and full scattering problems, one needs to identify the intersections of the stable and unstable manifolds with the Poincaré section Σ . One possible way is described in Gabern et al. [31] and the references therein. The authors use a normal form method for the computation of the stable and unstable manifolds and their intersection with Σ .

We follow a different approach to compute the intersections, and build on the concepts of [67]. They use an algorithm for a decomposition of the phase space into those invariant sets on which the corresponding dynamical system is ergodic. Based on these ideas, we develop a multilevel approach for the decomposition of the set of interest.

First Return Time. Consider the system $\dot{x} = g(x)$ with $x \in \mathbb{R}^d$ and a smooth function $g : \mathbb{R}^d \rightarrow \mathbb{R}^d$. Then the vector field g generates a flow $\varphi^t : \mathbb{R}^d \rightarrow \mathbb{R}^d$ with a smooth function φ defined for all $x \in \mathbb{R}^d$ and t in some interval $I \in \mathbb{R}$. Consider a local compact cross section $\Sigma \subset \mathbb{R}^d$ which is transverse to the flow φ , and each point $q \in \Sigma$ has to be valid in the system g . Recall that the *Poincaré map* $F : U \rightarrow \Sigma$ for a point $q \in U$ is defined by $F(q) = \varphi^{\tilde{\tau}(q)}(q)$, where $U \subseteq \Sigma$ and $\tilde{\tau}(q)$ is the time taken for the orbit $\varphi^{\tilde{\tau}(q)}(q)$ which starts at q to first return to Σ . We call $\tilde{\tau}(q)$ the *first return time* (see, for example, [35]).

We make use of the return time to divide the section Σ into different regions. Therefore, we need to define $\tilde{\tau}(q)$ for all $q \in \Sigma$ even if points do not come back to Σ . If $U = \Sigma$ then all points of the Poincaré section Σ will come back to it by definition and $\tilde{\tau}(q)$ exists for all $q \in \Sigma$. If $U \subset \Sigma$ then there are points in $\Sigma \setminus U$ for which the Poincaré map F is not defined. For our analysis, it is necessary that all points in Σ are assigned a time. Therefore, we define

$$\tau(q) := \begin{cases} \tilde{\tau}(q) & : q \in U \\ \infty & : q \in \Sigma \setminus U. \end{cases} \quad (2.3)$$

We use definition (2.3) for the computation of the first stable and unstable tube intersec-

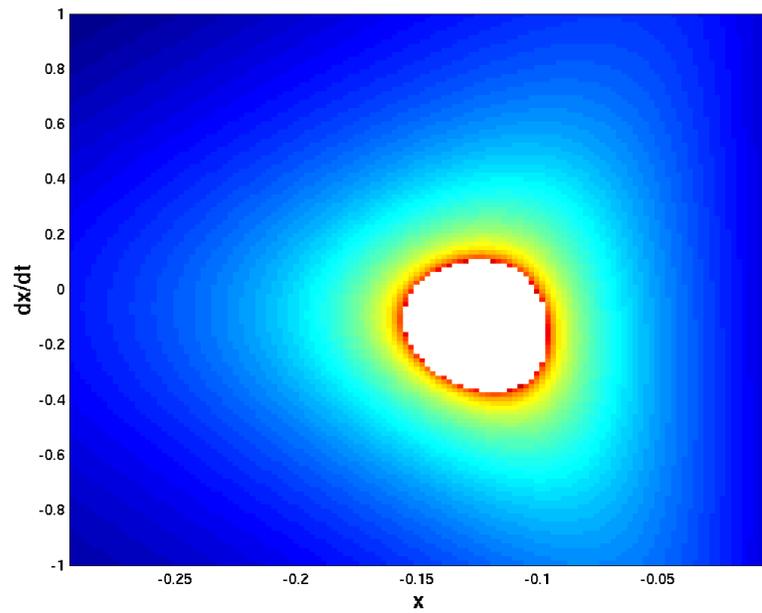


Figure 2.5: First return time distribution of the rectangle $X = [-0.295, -0.005] \times [-1.0, 1.0]$ for an electric field parameter $\epsilon = 0.58$. The white region in the middle indicates an infinite return time, meaning points in this region do not come back to the Poincaré section under consideration, and the other colors correspond to a finite return time decreasing from the inner to the outer region.

tions with Σ . Figure 2.5 shows the first return time distribution for the 2-degree-of-freedom Rydberg atom in crossed electric and magnetic fields for an electric field parameter $\epsilon = 0.58$. For this we took the rectangle $X = [-0.295, -0.005] \times [-1.0, 1.0]$ as Σ and divided it into 16384 small boxes. The color of the boxes corresponds to the average return time with respect to initial conditions in the respective box. The white region corresponds to the interior and the boundary of the stable tube (compare with Figure 2.4) and indicates an infinite return time. Besides this, the other colors show a finite return time decreasing from the inner to the outer region.

In Section 2.2.3 we introduced asymptotic, transit, and nontransit orbits, which we will denote by \mathcal{O}_{as} , \mathcal{O}_{tr} , and \mathcal{O}_{ntr} , respectively. These are orbits on the boundary, inside, and outside of the invariant manifolds, respectively. Uniqueness of solutions ensures that an orbit cannot change between these groups [31, 51].

Recall that there is no possibility of return for the valence electron after it crosses from the bound to the unbound state. This means that for the system under consideration, particles that leave the Poincaré section through the interior of the first intersection of the stable manifold with Σ will never come back to Σ . The same applies to particles on the boundary of this intersection. Therefore, in terms of return times, the sets \mathcal{O}_{as} , \mathcal{O}_{tr} , and \mathcal{O}_{ntr} are given by

$$\mathcal{O}_{as} = \{x \in \Sigma \mid \exists \epsilon > 0 \text{ and } \exists y, z \in V_\epsilon(x) \text{ with } \tau(y) = \infty \text{ and } \tau(z) < \infty\},$$

$$\mathcal{O}_{tr} = \{x \in \Sigma \mid \exists \epsilon > 0 \text{ such that } \forall y \in V_\epsilon(x), \tau(y) = \infty\},$$

$$\mathcal{O}_{ntr} = \{x \in \Sigma \mid \exists \epsilon > 0 \text{ such that } \forall y \in V_\epsilon(x), \tau(y) < \infty\},$$

where $V_\epsilon(x)$ denotes an ϵ -neighborhood of x .

With these theoretical considerations we are now able to devise an algorithm, which is based on the ideas of [24, 25] and provides a method for the approximation of \mathcal{O}_{as} .

Set-Oriented Subdivision Algorithm. The set-oriented *subdivision algorithm* generates a sequence $\mathcal{B}_0, \mathcal{B}_1, \dots$ of finite collections of compact subsets of \mathbb{R}^n such that the diameter $\text{diam}(\mathcal{B}_k) = \max_{B \in \mathcal{B}_k} \text{diam}(B)$ converges to zero for $k \rightarrow \infty$. Given an initial collection \mathcal{B}_0 , we obtain \mathcal{B}_k from \mathcal{B}_{k-1} for $k = 1, 2, \dots$ by

(i) *Subdivision:*

Construct a new collection $\hat{\mathcal{B}}_k$ such that

$$\bigcup_{B \in \hat{\mathcal{B}}_k} B = \bigcup_{B \in \mathcal{B}_{k-1}} B \text{ and}$$

$$\text{diam}(\hat{\mathcal{B}}_k) \leq \theta_k \text{diam}(\mathcal{B}_{k-1}) \text{ where } 0 < \theta_{\min} \leq \theta_k \leq \theta_{\max} < 1.$$

(ii) *Selection:* Define a new collection \mathcal{B}_k by

$$\mathcal{B}_k = \{B \in \hat{\mathcal{B}}_k \mid \exists x, y \in B \text{ with } \tau(x) = \infty \text{ and } \tau(y) < \infty\}.$$

Remark. By construction we have

$$\text{diam}(\mathcal{B}_k) \leq \theta_{\max}^k \text{diam}(\mathcal{B}_0) \rightarrow 0 \text{ for } k \rightarrow \infty.$$

We denote by Σ_k the collection of compact subsets obtained after k subdivision steps, $\Sigma_0 = \Sigma$. These Σ_k s define a nested sequence of compact sets, i.e., $\Sigma_{k+1} \subset \Sigma_k$. For each l we have $\Sigma_l = \bigcap_{k=0}^l \Sigma_k$, and we may view

$$\Sigma_\infty = \bigcap_{k=0}^{\infty} \Sigma_k$$

as the limit of the Σ_k s.

This algorithm converges to

$$\mathcal{O}_{as} = \Sigma_\infty.$$

Remark. To obtain the sets corresponding to the unstable manifold one needs to proceed backwards in time.

For the 3-degree-of-freedom system and a parameter value of $\epsilon = 0.58$, Figure 2.6 shows the $x\dot{x}$ - and $z\dot{z}$ -projections of the first stable (blue) and first unstable (red) tube intersections.

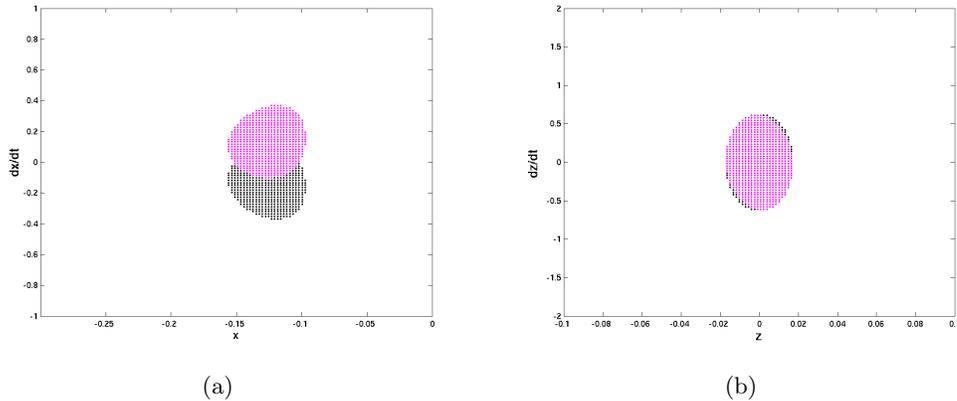


Figure 2.6: First intersection of the stable (blue) and unstable (red) tube with the Poincaré section in (a) $x\dot{x}$ - and (b) $z\dot{z}$ -projections for a parameter value $\epsilon = 0.58$.

Higher Return Times. The concept of the computation of the first tube intersection with the Poincaré section can easily be extended to the computation of further intersections. The n -th return time to Σ is denoted by $\tau^n(q)$ for $q \in \Sigma$. Figures 2.3 and 2.4 show further intersections of the stable (blue) and unstable (red) tube boundaries with Σ for two different parameter values. These computations were carried out using the above subdivision algorithm.

Now we have identified and approximated the regions of interest—for the following transport computations we only need the first intersections of the stable and unstable manifold with the Poincaré section. In the next subsection we show how the transition rates between these sets can be computed.

2.3.2 Transport Rates

The set-oriented approach provides a convenient framework for the computation of transport rates between regions of interest. In the following, we briefly describe a method that relies on an appropriate discretization of a transfer operator—the Perron-Frobenius operator. For a detailed description we refer to [27, 70].

Transfer Operator. Let

$$f : M \rightarrow M, \quad x_{k+1} = f(x_k), \quad k \in \mathbb{Z},$$

be a map and $R_1, \dots, R_l \subset M$ a partition of M into l regions. We are interested in the transport rates

$$T_{i,j}(n) = m(f^{-n}(R_j) \cap R_i),$$

where m denotes the Lebesgue measure, that is, the mass or volume of material transported from some region R_i to R_j in n steps.

Generally, the evolution of measures ν on M can be described in terms of the *transfer operator* (or *Perron–Frobenius operator*) associated with f . This is a linear operator $P : \mathcal{M} \rightarrow \mathcal{M}$,

$$(P\nu)(A) = \nu(f^{-1}(A)), \quad A \text{ measurable,}$$

on the space \mathcal{M} of signed measures on M .

This operator concept relates to the transport quantities in the following way:

Corollary 1. *Let $m_i \in \mathcal{M}$ be the measure $m_i(A) = m(A \cap R_i) = \int_A \chi_{R_i} dm$, where χ_{R_i} denotes the indicator function on the region R_i . Then*

$$T_{i,j}(n) = (P^n m_i)(R_j).$$

(Here P^n refers to the n -fold application of the transfer operator P .)

Since an analytic expression for this operator will usually not be available, we need to derive a finite-dimensional approximation to it.

Discretization of Transfer Operators. As a finite dimensional space $\mathcal{M}_{\mathcal{B}}$ of measures on M we consider the space of absolutely continuous measures with density $h \in \Delta_{\mathcal{B}} := \text{span}\{\chi_B \mid B \in \mathcal{B}\}$, i.e., one which is piecewise constant on the elements of the partition (box covering) \mathcal{B} . Let $Q_{\mathcal{B}} : L^1 \rightarrow \Delta_{\mathcal{B}}$ be the projection

$$Q_{\mathcal{B}}h = \sum_{B \in \mathcal{B}} \frac{1}{m(B)} \int_B h dm \chi_B.$$

Then for every set A that is the union of partition elements we have

$$\int_A Q_{\mathcal{B}}h dm = \int_A h dm.$$

Hence a discretization of the transfer operator P with respect to the box collection \mathcal{B} ,

consisting of b boxes, is given in terms of a transition matrix $P_{\mathcal{B}} := (p_{ij})$ with

$$p_{ij} = \frac{m(f^{-1}(B_i) \cap B_j)}{m(B_j)}, \quad i, j = 1, \dots, b.$$

So the entry p_{ij} gives the (conditional) probability that a particle is mapped from box B_j to B_i within one iterate of f .

Approximation of Transport Rates. For a measurable set A let

$$\underline{A} = \bigcup_{B \in \mathcal{B}: B \subset A} B \quad \text{and} \quad \overline{A} = \bigcup_{B \in \mathcal{B}: B \cap A \neq \emptyset} B.$$

We obtain the following estimate on the error between the true transport rate $T_{i,j}(n)$ and its approximation using powers of the transition matrix $P_{\mathcal{B}}$. To abbreviate notation, let $\underline{e}_R, \overline{e}_R, \underline{u}_R$ and $\overline{u}_R \in \mathbb{R}^b$ be defined by

$$(\underline{e}_R)_i = \begin{cases} 1, & \text{if } B_i \subset R \\ 0, & \text{else} \end{cases}, \quad (\overline{e}_R)_i = \begin{cases} 1, & \text{if } B_i \cap R \neq \emptyset \\ 0, & \text{else} \end{cases}$$

and

$$(\underline{u}_R)_i = \begin{cases} m(B_i), & \text{if } B_i \subset R \\ 0, & \text{else} \end{cases}, \quad (\overline{u}_R)_i = \begin{cases} m(B_i), & \text{if } B_i \cap R \neq \emptyset \\ 0, & \text{else} \end{cases}$$

where $i = 1, \dots, b$.

Lemma 1. *Let $R_i, R_j \subset M$,*

$$S_0 = R_j, \quad S_{k+1} = f^{-1}(\overline{S}_k), \quad k = 0, \dots, n-1$$

and

$$s_0 = R_j, \quad s_{k+1} = f^{-1}(\underline{s}_k), \quad k = 0, \dots, n-1.$$

Then

$$\begin{aligned}
& \left| T_{i,j}(n) - \underline{e}_{R_j}^T P_{\mathcal{B}}^n \underline{u}_{R_i} \right| \\
& \leq \underline{e}_{R_j}^T P_{\mathcal{B}}^n (\overline{u}_{R_i} - \underline{u}_{R_i}) + (\overline{e}_{R_j} - \underline{e}_{R_j})^T P_{\mathcal{B}}^n \overline{u}_{R_i} \\
& + \max \left\{ m \left(f^{-n}(R_j \setminus \underline{R_j}) \cap R_i \right), m \left(f^{-n}(\overline{R_j} \setminus R_j) \cap R_i \right) \right\} \\
& + \max \left\{ m \left((S_n \setminus f^{-n}(\underline{R_j})) \cap R_i \right), m \left((f^{-n}(\overline{R_j}) \setminus s_n) \cap R_i \right) \right\}.
\end{aligned}$$

For a proof of this statement we refer to [70]. This result is an improvement of a similar estimate in [27]. The main difference is that in this statement the error stays bounded if n goes to infinity. Furthermore, this estimate gives a bound on the error between the true transport rate $T_{i,j}(n)$ and the one computed via the transition matrix $P_{\mathcal{B}}$. Especially those elements of the fine partition \mathcal{B} contribute to the error which either intersect the boundaries or which contain pre-images of the boundary of R_j , see Figure 2.7 for an illustration. A direct consequence of Lemma 1 is that in order to ensure a certain degree of accuracy of the transport rates, these particular boxes need to be refined.

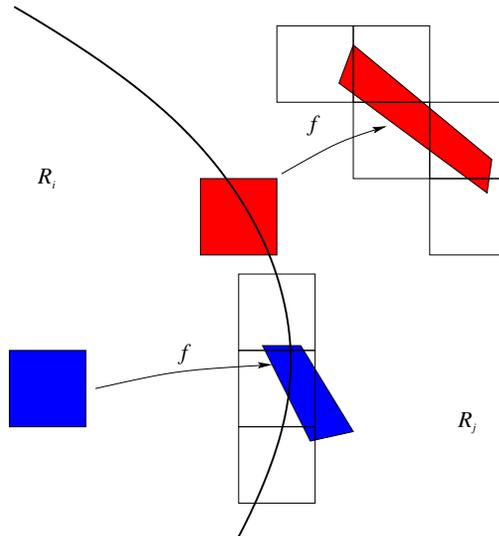


Figure 2.7: Two box transitions that contribute to the error between the computed and the actual value of the transport rate from a region R_i into region R_j after one iterate. This figure is taken from [27].

Convergence. Using Lemma 1 one can prove convergence for the approximate transport rate as the box covering is refined (see [27, 70]).

Adaptive Refinement of the Box Covering. As shown above, the boxes that contribute considerably to the error are those that either map onto the boundary of the target set or whose pre-image lies on the boundary of the source set. Unlike the situation in [27], one usually does not observe the desired transport within one iteration of the map, but only after a longer time span. Therefore, we use the following algorithm, discussed in [70], for the refinement of the transport boxes.

Adaptive Algorithm. Let $R_i, R_j \subset M$ and $n \in \mathbb{N}$. Let \mathcal{B} be a box covering of M , let $N := \lceil \frac{n}{2} \rceil$ and let $P_{\mathcal{B}}$ be the transition matrix as defined above. Determine the boundary boxes

$$\begin{aligned} b_{R_i} &:= \overline{R_i} \setminus \underline{R_i} \\ b_{R_j} &:= \overline{R_j} \setminus \underline{R_j} \end{aligned}$$

and compute

$$\begin{aligned} \underline{T}_{i,j}(n) &:= \underline{e_{R_j}}^T P_{\mathcal{B}}^n \underline{u_{R_i}} \\ \overline{T}_{i,j}(n) &:= \overline{e_{R_j}}^T P_{\mathcal{B}}^n \overline{u_{R_i}}, \end{aligned}$$

the *numerical* lower and upper bound on the transport rate $T_{i,j}(n)$, respectively. Choose $J \in \mathbb{N}$.

For $j = 1, \dots, J$

1. Mark all boxes B for which

$$f_k(B) \cap b_{R_j} \neq \emptyset \text{ for } k \in \{1, \dots, N\}$$

or

$$f^{-k}(B) \cap b_{R_i} \neq \emptyset \text{ for } k \in \{1, \dots, N\}.$$

(This information is coded in the transition matrix.)

2. Subdivide marked boxes.
3. Compute $P_{\mathcal{B}}$.
4. Determine b_{R_i} , b_{R_j} , $\underline{T_{i,j}(n)}$, and $\overline{T_{i,j}(n)}$.

The algorithm produces an adaptive covering, refining those boxes in particular that contribute to the error in computing the transport rates. Moreover, the algorithm gives an upper and lower bound to the transport rate, at least up to the error estimated in Lemma 1. Note that the numerical effort to compute the approximate transport rates essentially consists in n matrix-vector-multiplications—where the matrix $P_{\mathcal{B}}$ is sparse.

Transport Probabilities. In many applications one is interested in transport probabilities rather than in the transported volume. The transport probability as a function of the number of iterations n is given by

$$q_{i,j}(n) = \frac{T_{i,j}(n)}{m(R_i)},$$

that is, the fraction of particles in R_i that gets transported to R_j in n steps.

An approximation $\tilde{q}_{i,j}(n)$ to this quantity can be obtained using the upper and lower bounds on the transport rates and taking an average in the following way:

$$\tilde{q}_{i,j}(n) = \frac{T_{i,j}(n) + \overline{T_{i,j}(n)}}{m(\underline{R}_i) + m(\overline{R}_i)}.$$

Note that the quantities $\tilde{q}_{i,j}(n)$ can be computed from the box covering and the transition matrix, whereas in our setting the true transport probabilities $q_{i,j}(n)$ are theoretical values. Convergence of $\tilde{q}_{i,j}(n)$ to $q_{i,j}(n)$ follows from the results above when the box covering is appropriately refined.

2.3.3 Implementation

The algorithms described above are implemented in the dynamical systems software package GAIO (*Global Analysis of Invariant Objects*, see [22]). The box collections \mathcal{B}_k are realized by generalized rectangles of the form

$$B(c, r) = \{y \in \mathbb{R}^d \mid |y_i - c_i| \leq r_i \text{ for } i = 1, \dots, d\},$$

where $c \in \mathbb{R}^d$ denotes the center and $r \in \mathbb{R}^d$ the radius of the rectangle (box). For our computations we use a finite number of test points in each box, such as a regular grid or Monte Carlo points; see, for instance, [24] or [48] for a discussion on the choice of test points. In GAIO, the boxes are stored in a binary tree, where the children of a box at depth k are constructed by bisecting the box in alternate coordinate directions.

Note that the methods described above can be used in parallel to speed up the computation time.

2.4 Examples

We demonstrate the strength of our methods by computing ionization probabilities for the full and half scattering problems of the Rydberg atom in crossed electric and magnetic fields. We choose an energy of $E = -1.52$.

First we consider the full scattering problem of the 2-degree-of-freedom system for an electric field parameter $\epsilon = 0.57765$. We compare the results of the computation with the respective results for the 3-degree-of-freedom system. Then we analyze the 3-degree-of-freedom system for $\epsilon = 0.58$, allowing a comparison with [31]. Finally, we use the results from the previous computations to consider the half scattering problem.

2.4.1 Full Scattering Problem for the 2- and 3-Degree-of-Freedom System ($\epsilon = 0.57765$)

For the 2-degree-of-freedom system we consider the rectangle $X = [-0.295, -0.005] \times [-1.0, 1.0]$ on the Poincaré section Σ . We start with a partition of X on depth 8. By applying the return time algorithm in forward and backward time, we can identify and approximate the first stable and unstable tube intersections, respectively. As a result, we obtain a covering of X on depth 8, with the boxes covering the boundary of the tube intersections on depth 18. This covering consists of 736 boxes. We denote by \underline{R}_1 the set of boxes in the interior of the unstable tube intersection and by \overline{R}_1 the boxes covering the interior and the boundary. The sets \underline{R}_2 and \overline{R}_2 correspond to the stable tube intersection. Note that we are not given R_1 and R_2 explicitly because we can only approximate these sets on the box level, yielding $\underline{R}_1, \overline{R}_1, \underline{R}_2, \overline{R}_2$.

We then apply $J = 5$ steps of the adaptive refinement algorithm with 25 grid points

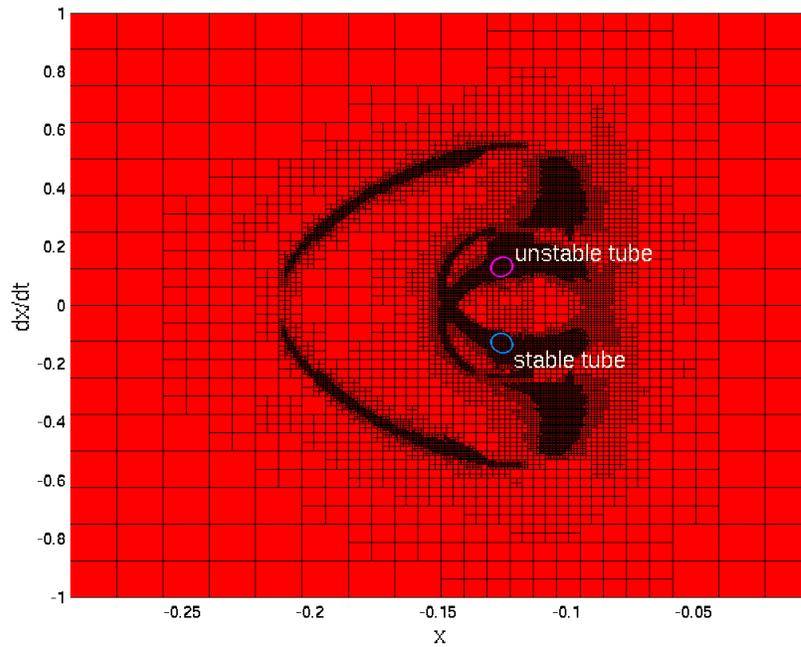


Figure 2.8: Adaptive box covering for the Rydberg atom in crossed fields. In the 2-degree-of-freedom system for $\epsilon = 0.57765$ those boxes are refined that contribute to the error in the computation of the transport rates. The unstable (light red) and stable (blue) tube intersections are superimposed.

per box. We choose $N = 5$ because we want to consider at least $n = 10$ iterations of the Poincaré map for our transport calculations. In each step, we subdivide in both coordinate directions at once. As the boundary is on depth 18, there is no gain in considering boxes on finer levels because while the computational effort increases, we would not get any new information. The resulting box covering (18670 boxes), with those boxes contributing to the error in the transport rate having being refined, is on depth 8/18; see Figure 2.8.

In Figure 2.9 we show the numerical lower and upper bounds on the transport rates, $T_{1,2}(n)$ and $\overline{T_{1,2}(n)}$, respectively, for $n = 1, \dots, 15$. Observe that the scattering profile is structured. The approximate scattering probabilities $\tilde{q}_{1,2}(n)$ are shown in Figure 2.10(a). The electron scattering probability is about 22% for $n = 5$ loops around the nuclear core. It is zero or almost zero for all other n apart from $n = 10$ and $n = 11$, where we observe small probabilities.

To check the results, we computed these probabilities using as many as 900 grid points

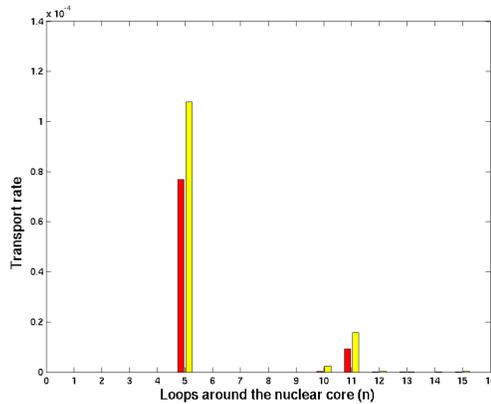


Figure 2.9: Full scattering problem for the 2-degree-of-freedom Rydberg atom in crossed fields for $\epsilon = 0.57765$. Approximations of the lower bound $\underline{T_{1,2}(n)}$ (blue) and the upper bound $\overline{T_{1,2}(n)}$ (red) on the transport rate for $n = 1, \dots, 15$ are shown.

per box, obtaining almost identical results. So for the given accuracy of the sets of interest, we can be sure that the results are correct.

We compare the results in the planar Rydberg system with those obtained in the 3-degree of freedom problem. In the 3-degree of freedom system we have the coordinates $x, y, z, \dot{x}, \dot{y}, \dot{z}$. Fixing a constant energy and a Poincaré section defined by (2.2), our remaining coordinates are x, z, \dot{x}, \dot{z} . Therefore, the initial box needs to be four-dimensional. For the following computations we chose $X = [-0.3, 0] \times [-0.1, 0.1] \times [-1.0, 1.0] \times [-2.0, 2.0]$.

We start with a box covering on depth 16 and apply the return time algorithm in forward and backward time which yields a covering of the boundaries of R_1 and R_2 on depth 36. The resulting box collection consists of 139276 boxes. We then apply $J = 7$ steps of the adaptive refinement algorithm, choosing $N = 5$ and 100 Monte Carlo points per box. In each step we subdivide in two coordinate directions at once and obtain a covering of 2056672 boxes, again on depth 16/36. The approximate electron scattering probabilities $\tilde{q}_{1,2}(n)$ for the full scattering problem are shown in Figure 2.10(b). Note that the scattering profile has the same qualitative characteristics as for the 2-degree-of-freedom system. Yet, the probabilities are lower than in the planar case. A reason for this might be that the volumes of the tubes are smaller, while the relative box sizes are considerably bigger than in the planar case. If the volume of the tube is comparatively small, as in our example, we need to use a box covering on a much deeper level to decrease the error between the upper and

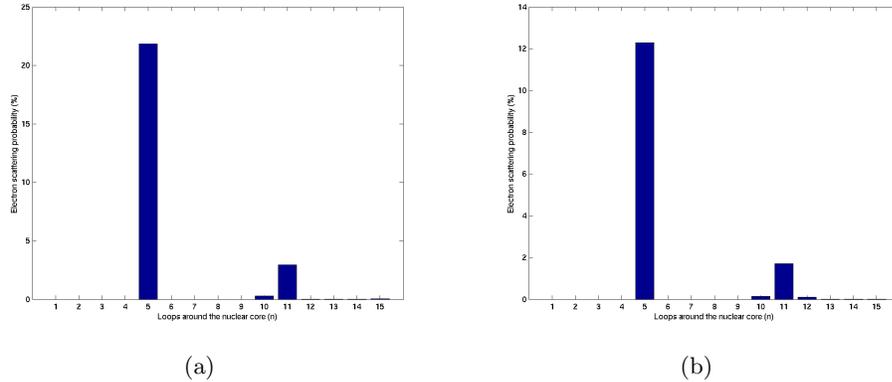


Figure 2.10: Full scattering problem for the Rydberg atom in crossed fields for $\epsilon = 0.57765$. Approximate transport probabilities $\tilde{q}_{1,2}(n)$ for $n = 1, \dots, 15$ in (a) the 2-degree-of-freedom system and (b) the 3-degree-of-freedom system.

lower bounds of the transport rates. However, by doing this we obtain a covering that is hardly manageable because it consists of a huge number of boxes.

To verify our results for this parameter value we computed the transport probabilities in the 3-degree of freedom system using as many as 1000 Monte Carlo points per box. This computation confirmed our results. Furthermore, in the 2-degree-of-freedom system our results agree very well with calculations done for this value of ϵ by the authors of [31] (personal communication).

2.4.2 Full and Half Scattering Problem for the 3-Degree-of-Freedom System ($\epsilon = 0.58$)

Choosing $\epsilon = 0.58$, the first intersections of the stable and unstable tubes with the Poincaré surface overlap. For the computation of the electron scattering probabilities in the 3-degree-of-freedom system we consider a partition of X as defined above. The box covering consists of 2155528 boxes on depth $16/28$, with the transport boxes refined using $J = 4$ steps of the adaptive algorithm as described above.

The scattering probabilities $\tilde{q}_{1,2}(n)$ with $n = 1, \dots, 10$ are shown in Figure 2.11(a). These results compare well with the scattering probabilities obtained by Gabern et al. [31], who analyzed the system using the same parameters.

In addition, we can re-use the box covering and the transition matrix already computed for dealing with the half scattering problem.

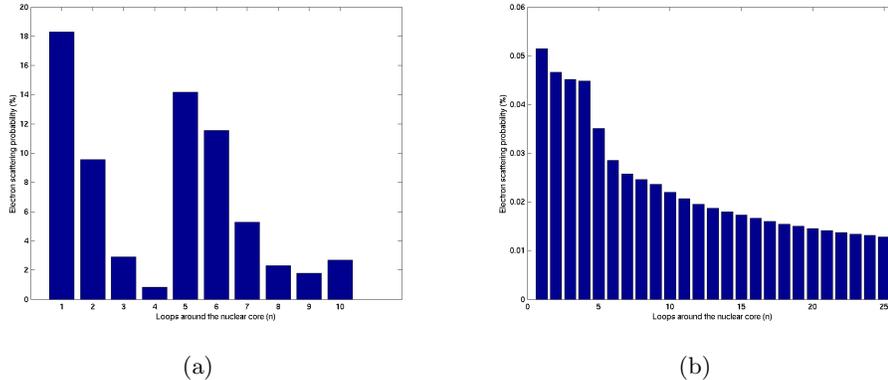


Figure 2.11: The 3-degree-of-freedom Rydberg atom in crossed fields for $\epsilon = 0.58$. Approximation of transport probabilities in (a) the full scattering problem and (b) the half scattering problem.

Here we define $\underline{R}_3 = X \setminus \overline{R}_2$ and $\overline{R}_3 = X \setminus \underline{R}_2$, that is, we consider the transport of particles from every region outside the stable tube R_2 into R_2 . Note that by this construction, \overline{R}_2 and \overline{R}_3 have a non-empty intersection, containing the boundary boxes of R_2 . So $\underline{T}_{3,2}(n)$ and $\overline{T}_{3,2}(n)$ can only give very coarse estimates on the transport rate because the boundary boxes are taken into account twice. Therefore we compute an approximation of the half scattering probability by

$$\hat{q}_{3,2}(n) = \frac{e_2^T P_{\mathcal{B}}^n \overline{u}_3 + \overline{e}_2^T P_{\mathcal{B}}^n \underline{u}_3}{m(\underline{R}_3) + m(\overline{R}_3)}.$$

This represents an average of the transport from \overline{R}_3 to \underline{R}_2 and \underline{R}_3 to \overline{R}_2 . The results are presented in Figure 2.11(b). Note that for higher iterates one observes an exponential decay of the electron scattering probabilities. Jaffé et al. [44] used a similar parameter value for the computation of so-called survival probabilities for the 2-degree-of-freedom half scattering problem in their paper. Even if it is not exactly the same value (they used an electric field parameter of $\epsilon = 0.6$) the shapes of the probabilities for both ϵ values look qualitatively the same.

2.5 Conclusion

This chapter has presented a set-oriented method for computing transport rates. We considered a suitable Poincaré section and introduced a new method for the computation of

tube intersections with this section using the Poincaré first return time. Based on these intersections we have the necessary information to find the regions between which transport will occur. We can use an adaptive algorithm for the computation of the transport rates relevant for the present situation (see [27] and [70]). It focuses on a global description of the dynamics using a box covering of the interesting region and a matrix of transition probabilities between these boxes for the calculation of the transport rates.

These techniques were demonstrated in the 2- and 3-degree-of-freedom systems for the Rydberg atom in crossed electric and magnetic fields. The generalization to higher dimensions is straightforward, with the limitation being the time taken to do the computations, as well as the memory required. For this reason, the calculation of transition rates for larger systems requires a reduced-order model of the system on which to use the set-oriented computational method.

In contrast to [31], the set-oriented approach does not require normal form techniques for the computation of tube intersections and does not use a Monte Carlo approach for the computation of the reaction rates. However, there is agreement between the results of the two approaches.

One possible next step in this line of research is to experimentally verify the numerical results presented in this chapter. The techniques for calculating the relevant transition rates are available but these observations have not yet been made. In such an experiment, there would be a spread of energies of the incoming electrons and also a variation in the electric field parameter ϵ . Thus, results of physical observations would not exactly match those in Figure 2.10 and Figure 2.11, but should be qualitatively the same. That is, an experiment should look for a non-exponential structure in the ionization rate that resembles those calculated here. For the results presented here to be directly comparable with experimental results, we would need to average over both the energy and the electric field.

The methods presented in this chapter represent a good starting point for further investigations using dynamical systems and geometric observations combined with set-oriented methods and statistics. To the best of our knowledge the results presented here and those of [31] represent the first successful calculation of reaction rates in a 3-degree-of-freedom chemical system.

An ongoing priority is to make set-oriented calculations in higher-dimensional problems more computationally efficient. Since even small molecular systems commonly have too

many dimensions for computation to be feasible, methods to reduce the number of variables needed to describe the coarse dynamics of a high-dimensional system are needed. The aim is to investigate high-dimensional multiscale problems by combining the set-oriented method with an appropriate procedure for distinguishing between optimal coarse and fine variables. An interesting possibility for coarse variables is the Perron-Frobenius eigenfunctions themselves [49, 68].

We would like to apply these methods to the computation of reaction rates and transition probabilities for more complex molecules. To do so will surely require some form of model reduction and the associated identification of suitable reaction coordinates, with the methods presented here applied to the coarse-level dynamics. In the following chapter we investigate a good candidate for a coarse variable that describes a transition. The gyration radii of a system could be used to describe the dynamics and the set-oriented method applied to these coarse variables to calculate transition rates.

Chapter 3

Energy Effects and Gyration Radius Dynamics in the Dissociation of Methane and Oxygen

In molecular dynamics, the effects of the bath particles on the motion of the primary particles are known to be important, yet are difficult to quantify. The primary particles are the particles undergoing the reaction and the bath particles are particles that exert an influence on these primary particles. Analyses of systems where the bath particles exert a significant influence most often consider only the potential acting between particles. However, these potentials are modified by the kinematics, or in other words, the motion, of the particles and thus, a realistic description of the system would quantify this effect.

Considering only the potentials acting between particles is insufficient in many molecular systems to describe the dynamics of transition [90]. The goal is to calculate the effective potential that governs the motion of the primary particles of a system. This effective potential will take into account the kinematics of the bath particles, or secondary particles (these terms will be used interchangeably). Another important goal is to find appropriate coarse variables to describe the dynamics of the transition. A good set of coarse variables will capture the essential dynamics and produce a lower-dimensional description of the system.

The system that we investigate is a dissociation reaction in one dimension, involving two primary particles and four secondary particles. We calculate two different effective

potentials. These effective potentials describe the dissociation behaviour of the system very well, whereas simply looking at the potential between the dissociating particles does not. The first exploits the kinetic energy distribution within the system to derive an entropic force that describes the dynamics of the primary particles as they dissociate. The reaction co-ordinate in this case is simply the distance between the dissociating primary particles.

The second method for calculating the effective potential makes use of the gyration radius of the system as a coarse variable. The gyration radius formalism developed by Yanao and collaborators [89] considers the kinematic influence of the secondary particles. The elegance of this method is its simple description of the dynamics. Previously, they had been applied by Yanao and co-workers to study conformational changes in small molecular systems [88], but since they essentially capture the mass distribution of the system, they are ideally suited to dissociation reactions. This is the first application of the formalism to the dynamics of dissociation reactions. We shall see that the gyration radius not only provides a good coarse variable for describing the transition, but that the effective potential calculated along this reaction co-ordinate explains the observed dissociation behaviour.

In the next section we introduce the model dissociation reaction that we will study. The Kinetic Energy Bias method for calculating the effective potential is presented in Section 3.2. Following this, we explain the Gyration Radius method, which recasts the system dynamics in terms of a coarse variable that takes the effect of the bath particles into account. We then derive an effective potential for the system using this formalism. The probability distribution and average lifetime of our model system is studied with respect to changes in energy and changes in the masses of the secondary particles. At the end of the section, a simpler system in which the influence of the bath particles' kinematics is seen more clearly is presented.

3.1 Model

The system that we investigate is motivated by one of the reactions that occur in the burning of methane, CH_4 , and oxygen, O . We study the one-dimensional system that has two primary particles (the carbon atom (C) and the oxygen atom) with a Lennard-Jones

potential between them, which is given by

$$V_{LJ} = 0.01 \left(\left(\frac{0.1}{r} \right)^{12} - \left(\frac{0.1}{r} \right)^6 \right). \quad (3.1)$$

We work in non-dimensional units. The hydrogen atoms (H) are attached to the carbon atom by spring potentials, with spring constant $k = 0.01$. A schematic of the system is shown in Figure 3.1. The potential energy function of the system is

$$V = 0.01 \left(\left(\frac{0.1}{|x_1 - x_2|} \right)^{12} - \left(\frac{0.1}{|x_1 - x_2|} \right)^6 \right) + \frac{1}{2}k|x_3 - x_1|^2 + \frac{1}{2}k|x_4 - x_1|^2 + \frac{1}{2}k|x_5 - x_1|^2 + \frac{1}{2}k|x_6 - x_1|^2. \quad (3.2)$$

The total energy of the system is a constant and of a magnitude that is sufficient for dissociation to occur. In our simulations, the primary particles have no initial velocity and the initial conditions of the hydrogen atoms, here the bath particles, are chosen randomly so as to conserve energy. Thus, for dissociation to occur, energy must be redistributed from the bath modes to the reactive mode, that is, from the H atoms to the C and O atoms. We work in the center of mass and center of momentum frame.

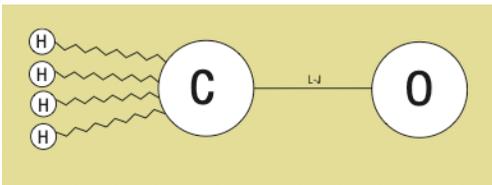


Figure 3.1: Schematic of the model problem

The masses of the four secondary particles are the same. These masses were changed for the different sets of simulations. The masses of the primary particles are set to unity. The different values of the secondary particle masses that we consider are $m_s = 0.1, 0.01, 0.001$ for $s = 3, \dots, 6$.

For the Kinetic Energy Bias method of Section 3.2, the total energy of the system was set to $E = 0.0036$. The initial conditions of the system are shown in Figure 3.2(a). The Lennard-Jones potential between the primary particles is shown in black and the total energy of the system is shown as the dashed red line. The initial separation of the primary

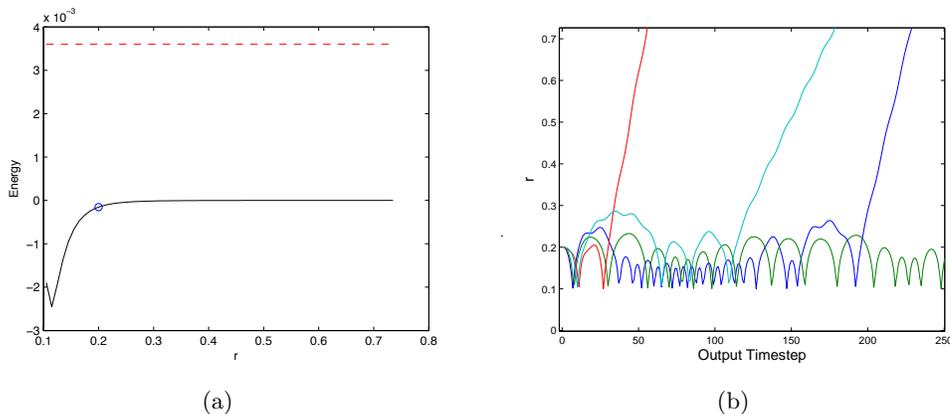


Figure 3.2: (a) Initial conditions. The Lennard-Jones potential between the primary particles is shown in black and the total energy of the system is the red dashed line. The initial separation of the primary particles is indicated by the blue circle. For dissociation to occur, energy from the bath modes must be transferred to the reactive mode, that is, to the primary particles. (b) Typical trajectories. The separation between the primary particles is shown on the vertical axis. This separation oscillates until dissociation occurs. Some trajectories, such as the green one, do not lead to dissociation.

particles is indicated by the blue circle. For dissociation to occur, energy from the bath particles must be transferred to the primary particles. The minimum of the Lennard-Jones potential occurs at $r = 0.11225$ where the potential is $V_{min} = -2.5 \times 10^{-3}$.

Figure 3.2(b) shows typical trajectories with the separation between primary particles, r , on the vertical axis and time on the horizontal axis. Simulations were run for much longer than shown here (8 times longer). Trajectories such as the green one continued to bounce (that is, the separation between the primary particles oscillated) until the end of the simulation without leading to dissociation. The other trajectories show the type of bouncing that occurs before dissociation.

For the Gyration Radius method of Section 3.3, the dissociation reaction was investigated both for a variety of bath particle masses and different total energies.

3.2 Kinetic Energy Bias Method

Knowing that for dissociation to occur in our system, energy must be transferred to the reactive mode, suggests that an examination of the energy distribution in the system is important for understanding the observed dynamics.

3.2.1 Theory

For an n -atom system in the center of mass and momentum frame, there are $n - 1$ degrees of freedom. The law of equipartition of energy states that the average kinetic energy per degree of freedom in a system is given by

$$\langle K \rangle_i = \frac{E - V}{n - 1} \quad \text{for } i = 1, \dots, n - 1 \quad (3.3)$$

where E is the total energy of the system and V is the potential energy. If equipartition of energy does not hold then there will be a bias in the kinetic energy balance and this bias can be characterized by α :

$$\langle K \rangle_i = \frac{E - V}{n - 1} \times \alpha_i \quad \text{for } i = 1, \dots, n - 1. \quad (3.4)$$

A reaction co-ordinate q_R is used to describe the reaction pathway. For our dissociation reaction, we take the distance between the primary particles, r , to be the reaction co-ordinate, with the additional restriction that along the reaction co-ordinate, the potential energy of the bath particles is very small. This is because the energy must be in the reactive mode for the dissociation to occur.

By considering the equation of motion for the reaction coordinate, we can derive the average force

$$\langle \text{Force} \rangle = -\frac{\partial V}{\partial q_R} + \frac{E - V(q_R)}{n - 1} \frac{\partial}{\partial q_R} \log \left(\frac{\alpha_2 \times \dots \times \alpha_{n-1}}{\alpha_R} \right). \quad (3.5)$$

The second term is an *entropic force term* where $\frac{E-V}{n-1}$, the average kinetic energy per degree of freedom, is the temperature. The motivation for the form of this term is the statistical mechanical force due to entropy. The α_i represent the available volume in phase space. We refer to the first term as the *potential force term* since it is the force due to the potential term. Data from reactive trajectories (trajectories leading to dissociation) at points where the potential energy of the bath particles is less than 5% of the total system energy was used to compute the average force at each point along the reaction co-ordinate. These data points are along the reaction co-ordinate as defined above.

The effective potential V_{eff} is the work done against this averaged force:

$$V_{\text{eff}} = - \int_{\text{path}} \langle \text{Force} \rangle dq_R. \quad (3.6)$$

This effective potential quantifies the kinematic effects due to the breakdown of the equipartition law via the dependence of the average force on the α_i .

3.2.2 Effective Potential

The effective potential given by Equation 3.6 was calculated numerically by running 100,000 simulations. Figure 3.3 shows the effective potential along the reaction coordinate (separation of primary particles) compared to the Lennard-Jones potential between the primary particles. Results are shown for secondary particle masses $m_s = 0.01$ and $m_s = 0.1$, with total energy $E = 0.0036$. The masses considered did not all yield effective potentials with the same number of wells but they all had at least a secondary well (a well apart from that of the primary Lennard-Jones potential).

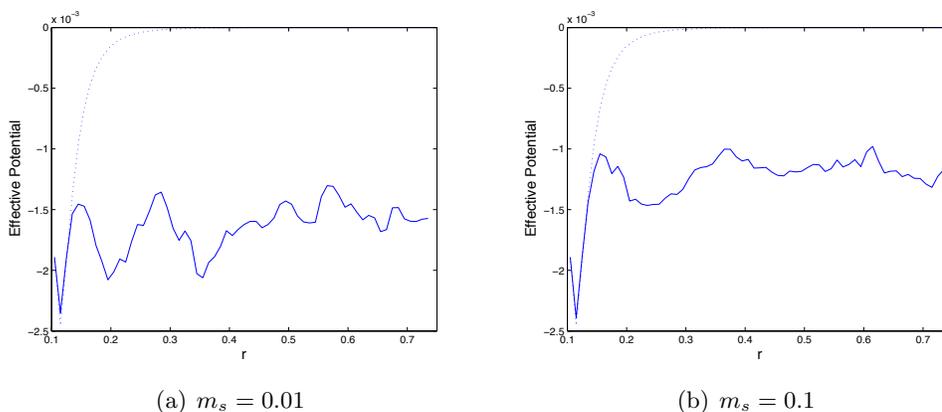


Figure 3.3: The effective potential as calculated from Eq. 3.6 is shown as a solid line. The horizontal axis is the reaction co-ordinate, which is the separation of the primary particles. The Lennard-Jones potential between the primary particles is shown as the dotted line. (a) secondary particle masses $m_s = 0.01$ and (b) secondary particle masses $m_s = 0.1$.

We see that the effective potential calculated by taking into account the break-down of the equipartition law is significantly different from the Lennard-Jones potential. It has the same primary well, but in addition, there are clearly-defined secondary and tertiary wells. The effects of these additional wells should be observable in the dissociation behaviour of

the system.

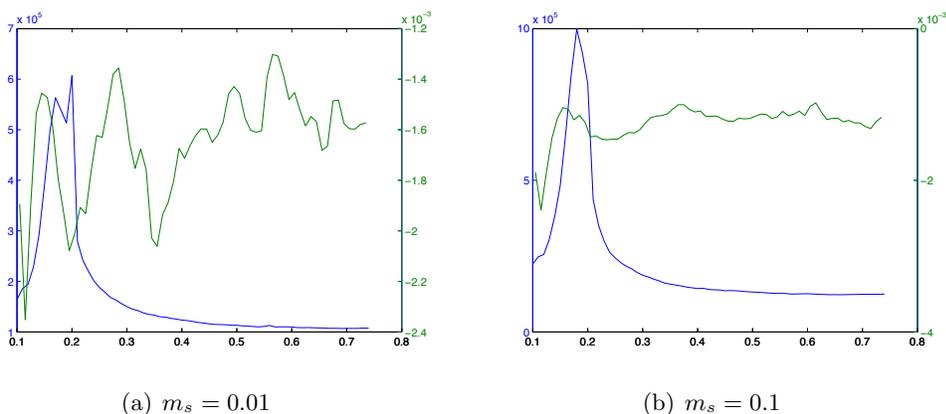


Figure 3.4: The distribution (in time) for reactive trajectories versus r is shown on the left vertical axis in blue and the effective potential is shown on the right vertical axis in green.

Figure 3.4 shows the effective potentials of Figure 3.3 along with the distribution of the reactive trajectories. These distributions plot the separation between the primary particles at each time step. This is the probability distribution in time of the system, i.e., peaks correspond to the primary particle separation most likely to be seen. This indicates that the system spends most of its time before dissociation on the downward segment of the effective potential leading into the secondary well. This is consistent with the system bouncing back and forth between the primary and secondary wells of the effective potential. This behavior can also be seen in Figure 3.2(b): the primary particle separation oscillates between $r = 0.125$ and $r = 0.26$.

The distributions in Figure 3.4 are dominated by data points from the part of the trajectories before dissociation and by the effects of velocity. To explain the latter, we observe that in the regions of the peaks in the distribution, the primary particles have less kinetic energy and thus traverse these regions more slowly. This can be seen by observing the kinetic energy distribution in Figure 3.8, where we see that the primary particles have the least kinetic energy at $r = 0.2$. A distribution in time (such as in Figure 3.4) is more likely to find the primary particles with a separation in this region. There is no peak in the distribution corresponding to the third well in the effective potential because in this region the primary particles have a larger velocity.

To account for these effects governing the distributions in Figure 3.4, we compute a

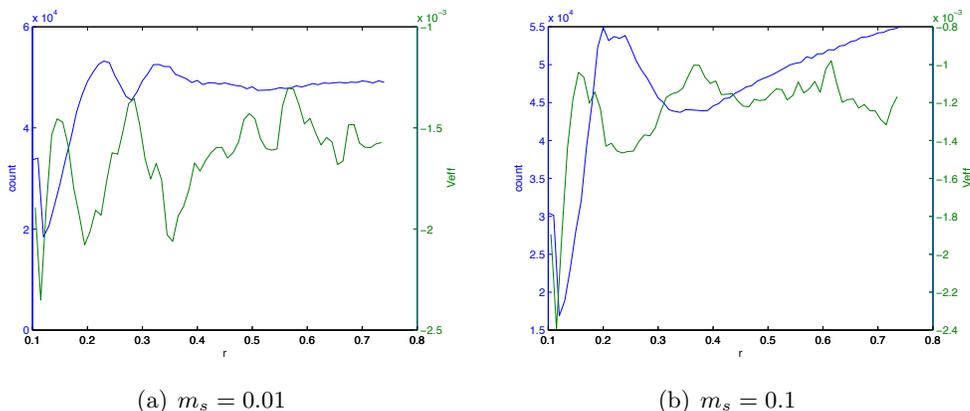


Figure 3.5: Distribution in space. The distribution of reactive trajectories, from the last bounce onwards, along the reaction coordinate, taking into account differences in dissociation velocity, is shown in blue on the left vertical axis. The effective potential is shown in green on the right vertical axis.

different distribution that may be considered to be a distribution in space (Figure 3.5). We consider only reactive trajectories from the time of their last bounce onwards, that is, trajectories that are on their way to dissociation. This last bounce is defined to occur at the last time that the primary particle separation is less than $r = 0.125$ before dissociation.¹ It was found (as can be seen in Figure 3.2(b)) that there is typically a last bounce with $r < 0.125$ immediately before dissociation occurs.

The distribution in space is found as follows. For any series of consecutive output time steps, if the primary particle separations are in the same bin in r , then only one of them is counted. We note that the output time step was chosen to be small enough such that every bin along r is visited consecutively, once the system is beyond the primary well. The difference between the distributions in Figure 3.4 and Figure 3.5 is that Figure 3.4 corresponds to where you are likely to find the system when you look at it and Figure 3.5 corresponds to what you would see if you drew the trajectories and plotted the number of unique visits to bins along r .

Figure 3.5 reveals that the effective potential calculated using the Kinetic Energy Bias method successfully explains the dissociation dynamics of the system. We see that in the process of dissociation, the system is temporarily trapped in the wells of the effective po-

¹This value for the last bounce is for a secondary particle mass of $m = 0.01$. Different values were used for each different mass.

tential. This behavior can not be explained by the Lennard-Jones potential alone. In terms of the primary particle separation, the trapped motion corresponds to certain separations about which the carbon and oxygen atoms oscillate on their way to dissociating.

We note that in Figure 3.5(b) the distribution increases at larger values of r beyond the secondary well. This is because as the primary particles dissociate, they oscillate in such a way as to cross the boundary of the bins along r a number of times. These small scale oscillations (as opposed to the larger ones across many bins in r caused by trapping in the secondary well) were more prominent for the larger bath masses. This binning issue may be overcome by changing the bin size, but this would require a bin size that would obscure the features we are interested in at smaller values of r .

3.2.3 Energy Distribution in the System

To further understand the mechanism behind the observed dissociation behaviour, we investigate the energy distribution in the system along the reaction pathway. We saw in Figure 3.5 that the dissociating particles are temporarily trapped in the secondary and tertiary wells of the effective potential. Since the effective potential was calculated from considering the breakdown of the law of equipartition of energy, we expect that the explanation for the trapping phenomenon will involve the transfer of energy between the modes in the system.

3.2.3.1 Ratio of Average Kinetic Energies

Figure 3.6 shows the ratio of the average kinetic energies of the primary and secondary particles for data points along the reaction co-ordinate. Peaks in this ratio occur in the wells of the effective potential. These wells determine the dissociation dynamics of the primary particles: the primary particles are caught in one of these wells when they do not have enough kinetic energy to continue to dissociation. For a fixed energy, this occurs when the bath particles have more kinetic energy. For dissociation to continue, this kinetic energy of the bath particles must be redistributed to the primary particles.

The average kinetic energy of the bath particles is higher than the average kinetic energy of the primary particles once the barrier between the primary and secondary wells of the effective potential has been crossed. In the primary well, the average kinetic energy of the primary particles is greater than that of the bath particles. This is to be expected, as we

are considering data along the reaction co-ordinate from the last bounce onwards, and in order for dissociation to occur, the primary particles must have sufficient energy to climb out of the primary well. This well is deeper than the other wells and so requires much more of the energy to be in the reactive mode.

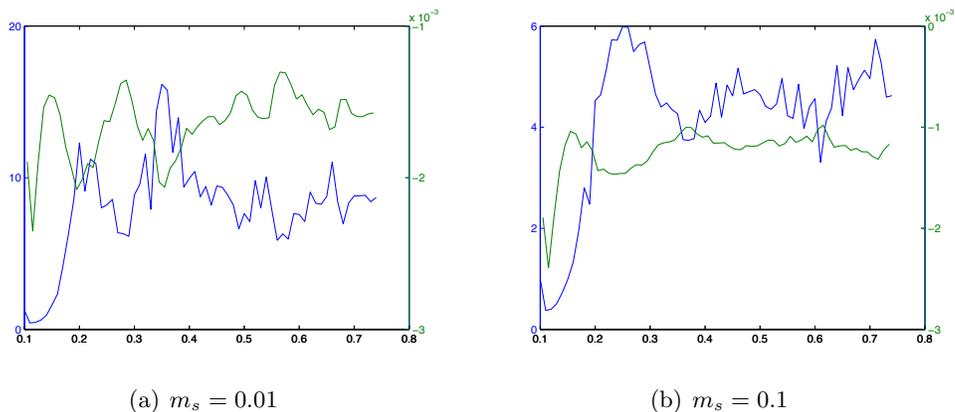


Figure 3.6: Ratio of bath kinetic energy to primary particles' kinetic energy versus r . The kinetic energy ratio is plotted on the left vertical axis in blue and the effective potential is plotted on the right vertical axis in green.

3.2.3.2 α Tubes

For α , which characterizes the breakdown of the law of equipartition of energy, as defined in Equation 3.4, we can construct tubes, called α -tubes, with radius given by $\alpha_{\text{bath}}^4 / \alpha_{\text{primary}}$. These tubes are shown in Figure 3.7. The α -tubes provide a visualization of the bottlenecks experienced by the dissociating particles. We see that the tubes are narrower in regions where there is a local maximum in the effective potential. Since the dissociation reaction proceeds through this tube, a narrower region of the tube indicates that it will be more difficult to get through this region of the tube and this corresponds to having to overcome a local barrier in the effective potential. The α -tubes are very narrow in the primary well of the effective potential because the depth of this well compared to the successive wells is much greater and therefore it is more difficult to overcome the barrier between the primary and secondary well. In other words, it is difficult to get the dissociation reaction started. Note that initially the primary particles have a separation of $r = 0.2$, so our system starts outside of the primary well and to the right of the narrowest region of the α -tubes.

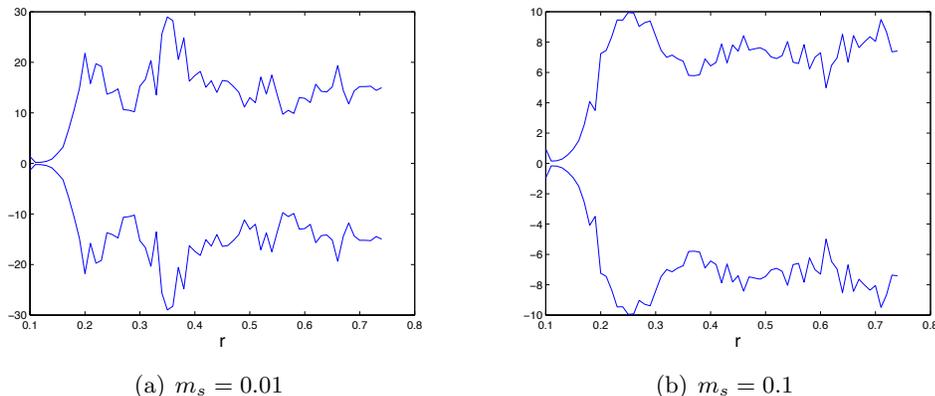


Figure 3.7: Visualization of the bottlenecks experienced by the dissociating particles: α -tube versus r . The radius of the tube is given by $\alpha_{\text{bath}}^4/\alpha_{\text{primary}}$.

3.2.3.3 Equipartition of Energy

Figure 3.8 shows the average kinetic energy for the primary and secondary particles. If the energy were equipartitioned then the plots for the primary and secondary particles would be the same. As they are not, equipartition of energy does not hold. We note that this is because in a dissociation reaction, the phase space is not bounded.

In contrast to previous figures, these plots were produced using all trajectories in the simulations. The sharp peak in the average kinetic energy of the secondary particles seen in the $m_s = 0.1$ case, close to $r = 0.2$, coincides with the secondary well in the effective potential. The minimum of the average kinetic energy of the primary particles corresponds to the peak in the distributions of Figure 3.4.

3.2.4 Discussion

The results shown in Section 3.2.2 and similar results for other mass parameters highlight the need to take into account kinematic effects in addition to the potentials between particles in molecular dynamics calculations. These kinematic effects are captured by an effective potential, which can be calculated by quantifying the break-down of the law of equipartition of energy and accounting for the entropic force that this leads to. The effective potential has been shown to agree with the observed behavior of the system.

An investigation of the energy distribution in the system in Section 3.2.3 confirmed that the kinetic energy in the system is not equipartitioned. The ratio of kinetic energies of the

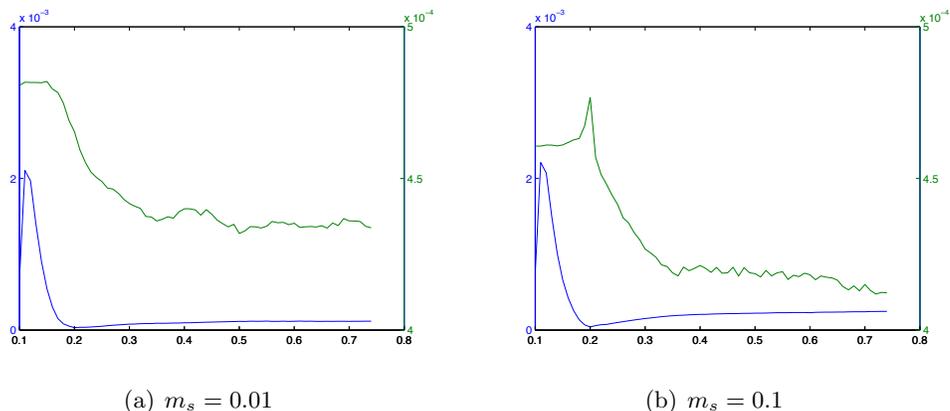


Figure 3.8: Average kinetic energy of primary (left, in blue) and secondary (right, in green) particles versus r . As these average energies are not the same, equipartition of energy does not hold. Note that the scales on the vertical axes are different.

bath and primary particles revealed that when the primary particles are trapped in the wells of the effective potential, the bath particles have much more kinetic energy. Thus the trapped motion can be explained by the primary particles losing kinetic energy to the bath modes and having to get this energy back in order to continue to dissociation. This phenomenon can be visualised as bottlenecks along the reaction co-ordinate, as shown by the α -tubes.

Although the Kinetic Energy Bias method has successfully explained the dynamics, a more elegant approach would provide a more sophisticated reaction co-ordinate than the distance between the primary particles. This reaction co-ordinate should take into account the kinematic effects observed to play a crucial role here. The effective potential calculated along this reaction co-ordinate should ideally be simpler and smoother than the effective potentials found in this section. The following section explores the use of the gyration radius of the system as the reaction co-ordinate.

3.3 Gyration Radius Method

The gyration radii, being related to the moments of inertia of a system, naturally incorporate the mass distribution of the system. As such, they are ideally suited for the dissociation reaction studied here and for conformational changes from symmetric to elongated shapes as studied by Yanao et al. [88]. By calculating an effective potential along the gyration

radius of the system (one gyration radius since our model is collinear), we can quantify the work that the dissociating system must do in order to transition to the unbound state. We are interested in the intrinsic equations of motion of the system, that is, in the internal shape space, and thus have eliminated overall translations and rotations of the system.

We begin with a general, isolated n -atom system and turn our attention to our specific model problem in Section 3.3.2.1. A more detailed exposition can be found in [88] and the references therein.

3.3.1 Background

In order to study the internal shape space of a system, it is important to first correctly separate the rotational and internal motion [89]. These two motions can be strongly coupled via the effects of a change in the mass distribution on the total angular momentum. Transition state theory typically assumes that they may be simply separated and that studying the dynamics on the potential energy surface is sufficient. However, as in the “falling cat” example, a molecule can continuously change its shape while keeping the total angular momentum equal to zero, and end up with a different orientation to the one it began with even though the molecule has returned to its initial shape. The body frame of the molecule (the frame with respect to which shape changes are viewed) changes as the molecule changes shape, even in a system with zero total angular momentum. Thus the body frame must be found in a careful manner in order to account for the internal changes.

The separation of the motions is correctly done using reduction theory from geometric mechanics [61, 60] and the associated gauge theory for systems with rotational symmetry [59, 36, 79, 54]. Since the body frame is not unique, the internal space has an associated gauge field. Having separated the overall translational and rotational motion in this way, we are left with a $(3n - 6)$ -dimensional internal space. In this internal space, the metric tensor is in general non-Euclidean and determined only by the mass and shape of the system. It is independent of the potential energy surface and the choice of body-fixed frame. The mass-related kinematic effects that we investigate arise from the non-Euclidean nature of this internal space.

We note that for collinear dynamics, as in our model, internal motion and rotation do not couple, but the body frame changes as the shape of the system changes, giving rise to a non-Euclidean internal space. In what follows, we will keep the discussion general until

Section 3.3.2.1, when we confine our attention to collinear systems.

An appropriate co-ordinate system in the internal space is given by the principal-axis hyperspherical co-ordinates. This formalism was developed by Chapuisat [12] and Kuppermann [52], based on early work by Eckart [28]. The principal axes of the instantaneous moment of inertia tensor provide a time-dependent body frame. The gyration radii of the system are the mass-weighted lengths of the molecule along each principal axis. The principal-axis hyperspherical co-ordinates are the three gyration radii and the $(3n - 9)$ hyperangles. A change in the latter generates a cyclic deformation of the molecule’s shape, referred to as a kinematic rotation [89]. These are different from the external rotations of the system in configuration space.

The coupling of the gyration radii with the internal kinematic rotation, via the non-Euclidean metric of the internal space, induces an “internal centrifugal force” in the space spanned by the gyration radii. This centrifugal force is an *internal* force and can be generated even if the system has zero total angular momentum. It is considered a centrifugal force since it has a quadratic dependence on the angular velocities of the kinematic rotations. It can have a significant influence on chemical reaction dynamics and has been extensively studied in the case of isomerization of atomic clusters [88]. The inclusion of this internal centrifugal force yields the effective potential for our model system. The force due to the potential in the system tends to confine the particles, acting against dissociation. The internal centrifugal force on the other hand works to elongate molecular systems and increase the mass distribution. In other words, the effect of the motion of the bath particles, via the internal centrifugal force, is to drive the system towards dissociation. The effective potential captures this competition between the potential and internal centrifugal forces.

3.3.2 Theory

The exposition of the formalism for finding the gyration radii of an isolated n -atom system follows the presentation in Yanao and Takatsuka [89], which is based on the work of Littlejohn and Reinsch [54]. In Section 3.3.2.1 we limit our attention to our collinear model system.

Let \mathbf{r}_{sj} be the position of the j th particle, where $j = 1, \dots, n$, with respect to some origin in a space-fixed frame. The overall translational degrees of freedom can be eliminated by using mass-weighted Jacobi co-ordinates $\{\rho_{s1}, \dots, \rho_{s(n-1)}\}$, where the subscript s refers to a

vector with respect to a space-fixed frame.

A body-fixed frame will be represented as a 3×3 proper rotation matrix $\mathbf{R} \in \text{SO}(3)$, whose columns are the orthonormal axes of the frame. We fix the origin of the frame to be at the center of mass of the system. This body frame, \mathbf{R} , specifies the orientation of the system. The relation between the mass-weighted Jacobi vectors above and the Jacobi vectors with respect to this frame, $\{\rho_i\}$, is given by

$$\rho_{si} = \mathbf{R} \rho_i(\{q^\mu\}) \quad i = 1, \dots, n-1, \quad (3.7)$$

where the q^μ are the $3n - 6$ internal co-ordinates of the system.

Differentiating with respect to time yields

$$\dot{\rho}_{si} = \dot{\mathbf{R}} \rho_i + \mathbf{R} \frac{\partial \rho_i}{\partial q^\mu} \dot{q}^\mu \quad (3.8)$$

where the Einstein summation convention (sum over repeated upper and lower indices) has been adopted over the indices $\mu = 1, \dots, (3n - 6)$ and will be used for all Greek indices.

The angular momentum with respect to the body frame can be written as

$$\begin{aligned} \mathbf{L} &= \mathbf{R}^T \mathbf{L}_s \\ &= \mathbf{R}^T \sum_{i=1}^{n-1} \rho_{si} \times \dot{\rho}_{si} \\ &= \sum_{i=1}^{n-1} \rho_i \times (\omega \times \rho_i) + \sum_{i=1}^{n-1} \rho_i \times \frac{\partial \rho_i}{\partial q^\mu} \dot{q}^\mu, \end{aligned} \quad (3.9)$$

where superscript T denotes the transpose and ω is the angular velocity vector of the body frame with respect to the frame itself.

The angular momentum can be rewritten in terms of the moment of inertia tensor \mathbf{M} , and the gauge potential \mathbf{A}_μ :

$$\mathbf{L} = \mathbf{M}(\omega + \mathbf{A}_\mu \dot{q}^\mu), \quad (3.10)$$

with

$$\mathbf{M}_{\alpha\beta} = \sum_{i=1}^{n-1} ((\rho_i \cdot \rho_i) \delta_{\alpha\beta} - \rho_{i\alpha} \rho_{i\beta}); \quad (3.11)$$

where $\delta_{\alpha\beta}$ is the Kronecker delta and

$$\mathbf{A}_\mu = \mathbf{M}^{-1} \left(\sum_{i=1}^{n-1} \rho_i \times \frac{\partial \rho_i}{\partial q^\mu} \right). \quad (3.12)$$

The kinetic energy with respect to the body-fixed frame is

$$K = \frac{1}{2} \sum_{i=1}^{n-1} \dot{\rho}_i \cdot \dot{\rho}_i \quad (3.13)$$

$$= \frac{1}{2} (\omega^T \mathbf{M} \omega) + (\omega^T \mathbf{M} \mathbf{A}_\mu) \dot{q}^\mu + \frac{1}{2} h_{\mu\nu} \dot{q}^\mu \dot{q}^\nu \quad (3.14)$$

where $h_{\mu\nu}$ is a pseudo-metric defined by

$$h_{\mu\nu} = \sum_{i=1}^{n-1} \frac{\partial \rho_i}{\partial q^\mu} \cdot \frac{\partial \rho_i}{\partial q^\nu}. \quad (3.15)$$

The first term on the right-hand side of Equation 3.14 is the rotational kinetic energy, the second term is the Coriolis term, and the third is the vibrational kinetic energy.

None of the terms in Equation 3.14 are separately gauge invariant, however Littlejohn and Reinsch [54] derived a rearrangement of the kinetic energy in which each term is individually gauge invariant:

$$K = \frac{1}{2} (\omega + \mathbf{A}_\mu \dot{q}^\mu)^T \mathbf{M} (\omega + \mathbf{A}_\mu \dot{q}^\mu) + \frac{1}{2} g_{\mu\nu} \dot{q}^\mu \dot{q}^\nu, \quad (3.16)$$

where $g_{\mu\nu}$ is a gauge-invariant metric tensor on the internal space defined as

$$g_{\mu\nu} = h_{\mu\nu} - \mathbf{A}_\mu^T \mathbf{M} \mathbf{A}_\nu. \quad (3.17)$$

This metric tensor is, in general, non-Euclidean and leads to the kinematic effects that we seek to explore.

If the total angular momentum of the system \mathbf{L} (Equation 3.10) is zero, the kinetic energy simplifies and the Lagrangian of the system becomes

$$\mathcal{L} = \frac{1}{2} g_{\mu\nu} \dot{q}^\mu \dot{q}^\nu - V(\{q^\mu\}), \quad (3.18)$$

where we restrict our attention to the case in which the potential energy, V , of the system

only depends on the internal co-ordinates, $\{q^\mu\}$.

The equations of motion are given by the Euler-Lagrange equations:

$$g_{\mu\nu}(\ddot{q}^\nu + \Gamma_{\kappa\lambda}^\nu \dot{q}^\kappa \dot{q}^\lambda) = -\frac{\partial V}{\partial q^\mu}. \quad (3.19)$$

The Christoffel symbols $\Gamma_{\kappa\lambda}^\nu$ are

$$\Gamma_{\kappa\lambda}^\nu = \frac{1}{2}g^{\nu\mu} \left(\frac{\partial g_{\mu\kappa}}{\partial q^\lambda} + \frac{\partial g_{\mu\lambda}}{\partial q^\kappa} - \frac{\partial g_{\kappa\lambda}}{\partial q^\mu} \right). \quad (3.20)$$

The metric force arises due to the non-vanishing Christoffel symbol yielding a force proportional to the velocities of the internal co-ordinates.

We now introduce the principal-axis hyperspherical co-ordinates, starting from the mass-weighted Jacobi vectors with respect to the body frame:

$$\rho_i = \sqrt{\mu_i} \left(\frac{\sum_{k=1}^i m_k \mathbf{r}_k}{\sum_{k=1}^i m_k} - \mathbf{r}_{i+1} \right) \quad i = 1, \dots, n-1, \quad (3.21)$$

where the μ_i are the reduced masses

$$\mu_i = \frac{m_{i+1} \sum_{k=1}^i m_k}{\sum_{k=1}^{i+1} m_k} \quad i = 1, \dots, n-1. \quad (3.22)$$

For an n -atom system in 3 dimensions, the matrix \mathbf{W} is defined to be the $3 \times (n-1)$ matrix with columns given by the Jacobi vectors,

$$\mathbf{W} \equiv (\rho_1 \dots \rho_{n-1}). \quad (3.23)$$

By the singular value decomposition theorem,

$$\mathbf{W} = \mathbf{N}\mathbf{U}^T \quad (3.24)$$

where \mathbf{N} is a $3 \times (n-1)$ matrix whose diagonal entries a_1, a_2, a_3 are the singular values of \mathbf{W} and the other elements are zero. The matrix \mathbf{U} is an $(n-1) \times (n-1)$ orthogonal matrix with the i th row given by the unit eigenvector \mathbf{u}_i^T . The square of the singular values are

the non-zero eigenvalues of the matrix $\mathbf{W}^T\mathbf{W}$ satisfying

$$\mathbf{W}^T\mathbf{W}\mathbf{u}_k = a_k^2\mathbf{u}_k \quad k = 1, 2, 3. \quad (3.25)$$

The matrix $\mathbf{W}^T\mathbf{W}$ has $(n-4)$ zero eigenvalues with corresponding eigenvectors $\mathbf{u}_4, \dots, \mathbf{u}_{n-1}$.

The a_k are the gyration radii of the system and the \mathbf{u}_i characterize the bath modes.

The gyration radii are related to the principal moments of inertia by

$$\mathbf{M}_1 = a_2^2 + a_3^2, \quad \mathbf{M}_2 = a_1^2 + a_3^2, \quad \mathbf{M}_3 = a_1^2 + a_2^2. \quad (3.26)$$

This formalism has been applied to the isomerization of three, four, and six atom clusters [88, 89]. We now use it to study our model collinear dissociation system, after rephrasing the formalism explicitly for collinear n -atom systems.

3.3.2.1 Collinear System

For an n -atom collinear system, such as the methane and oxygen dissociation model presented in Section 3.1, the overall translational degrees of freedom of the system are eliminated by introducing mass-weighted Jacobi vectors (scalars in our case) as defined in Equation 3.21.

The matrix of Jacobi vectors, \mathbf{W} , is a $1 \times (n-1)$ matrix that can be decomposed according to the singular value decomposition theorem as

$$\mathbf{W} = (\rho_1 \dots \rho_{n-1}) = (a_1 \ 0 \dots 0) \begin{pmatrix} \mathbf{u}_1^T \\ \mathbf{u}_2^T \\ \vdots \\ \mathbf{u}_{n-1}^T \end{pmatrix}, \quad (3.27)$$

where a_1 is the gyration radius and the \mathbf{u}_i are $(n-1)$ -dimensional orthogonal unit vectors.

The kinetic energy of the collinear n -atom system can be expressed as

$$\begin{aligned} K &= \frac{1}{2} \sum_{i=1}^{n-1} \dot{\rho}_i \cdot \dot{\rho}_i \\ &= \frac{1}{2} \dot{a}_1^2 + \frac{1}{2} \dot{a}_1^2 \dot{\mathbf{u}}_1^T \dot{\mathbf{u}}_1. \end{aligned} \quad (3.28)$$

The Lagrangian of the system, $\mathcal{L} = K - V$, yields the following Euler-Lagrange equation for the gyration radius

$$\ddot{a}_1 = a_1 \dot{\mathbf{u}}_1^T \dot{\mathbf{u}}_1 - \frac{\partial V}{\partial a_1}. \quad (3.29)$$

The first term on the right-hand side is an internal centrifugal force and will be called the *centrifugal force term*. It arises from the dynamical coupling of the gyration radius with the bath modes. The second term on the right-hand side will be called the *potential force term* since it is derived from the potential of the system.

The effective potential, V_{eff} , governing the reaction is given by the line integral of the averaged field of the force terms along the gyration radius:

$$V_{\text{eff}} = - \int \left(\langle a_1 \dot{\mathbf{u}}_1^T \dot{\mathbf{u}}_1 \rangle - \left\langle \frac{\partial V}{\partial a_1} \right\rangle \right) da_1. \quad (3.30)$$

Since there is only one gyration radius in our collinear system, we henceforth drop the subscript on a_1 . The potential force term can be computed by observing that

$$\frac{\partial V}{\partial a} = \sum_{j=1}^n \frac{\partial V}{\partial x_j} \frac{\partial x_j}{\partial a}. \quad (3.31)$$

The $\partial V/\partial x_j$ terms can be calculated from the potential energy function of the system, Equation 3.2. The $\partial x_j/\partial a$ terms are given simply by x_j/a . This can be seen as follows. The Cartesian co-ordinates $\{x_j\}$ are related to the mass-weighted Jacobi co-ordinates $\{\rho_i\}$ by the linear transformation

$$\begin{pmatrix} x_1 \\ \vdots \\ \vdots \\ x_n \end{pmatrix} = \mathbf{S}(\{m_j\}) \begin{pmatrix} \rho_1 \\ \vdots \\ \rho_{n-1} \\ R_{CM} \end{pmatrix} = \mathbf{S}(\{m_j\}) \begin{pmatrix} u_{(1,1)} \\ \vdots \\ u_{(1,n-1)} \\ 0 \end{pmatrix} \times a \quad (3.32)$$

where $\mathbf{S}(\{m_j\})$ is an $n \times n$ matrix whose components are functions of the masses of the particles and R_{CM} is the center of mass of the system. The subscripts on $u_{(1,j)}$ denote the $n-1$ components of the vector \mathbf{u}_1 . To obtain the second equality, we set $R_{CM} = 0$ without loss of generality, and use Equation 3.27. Since the matrix $\mathbf{S}(\{m_j\})$ and the vector \mathbf{u}_1 are independent of a , Equation 3.32 can be differentiated with respect to a to yield the desired

result,

$$\begin{pmatrix} \partial x_1 / \partial a \\ \vdots \\ \vdots \\ \partial x_n / \partial a \end{pmatrix} = S(\{m_j\}) \begin{pmatrix} u_{(1,1)} \\ \vdots \\ u_{(1,n-1)} \\ 0 \end{pmatrix} = \begin{pmatrix} x_1/a \\ \vdots \\ \vdots \\ x_n/a \end{pmatrix}. \quad (3.33)$$

3.3.3 Systems with Different Total Energies

In this section we investigate the effects of changing the total system energy on the dissociation of our model system in which the masses of the secondary particles are $m_s = 0.01$.

3.3.3.1 Distributions for Different Energies

The minimum of the potential energy, V , in Equation 3.2, occurs at $a = 0.08290$. Typical trajectories along the gyration radius are shown in Figure 3.9. A bounce was defined (for the mass and energy parameters of Figure 3.9) as occurring when a decreased until $a < 0.165$ and then subsequently increased. Dissociation is considered to have occurred once $a > 0.5$ as no trajectories were observed to bounce again after this. Bounces were not observed to exceed an a value of 0.4.

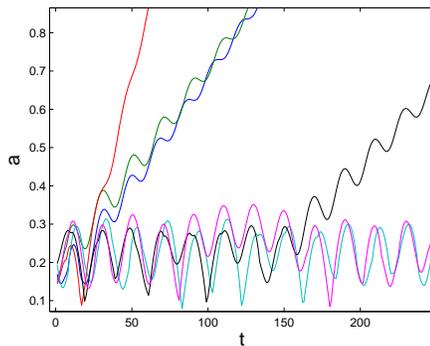


Figure 3.9: Typical trajectories. The gyration radius, a , is plotted against the time step.

We computed the two different types of distributions explained in Section 3.2.2, the distribution in time and the distribution in space, along both the primary particle separation, r , and the gyration radius, a . Total energies of $E = 0.001, 0.003, 0.005, 0.007$ were investigated. These distributions are shown in Figure 3.10. They include both reactive and non-reactive trajectories.

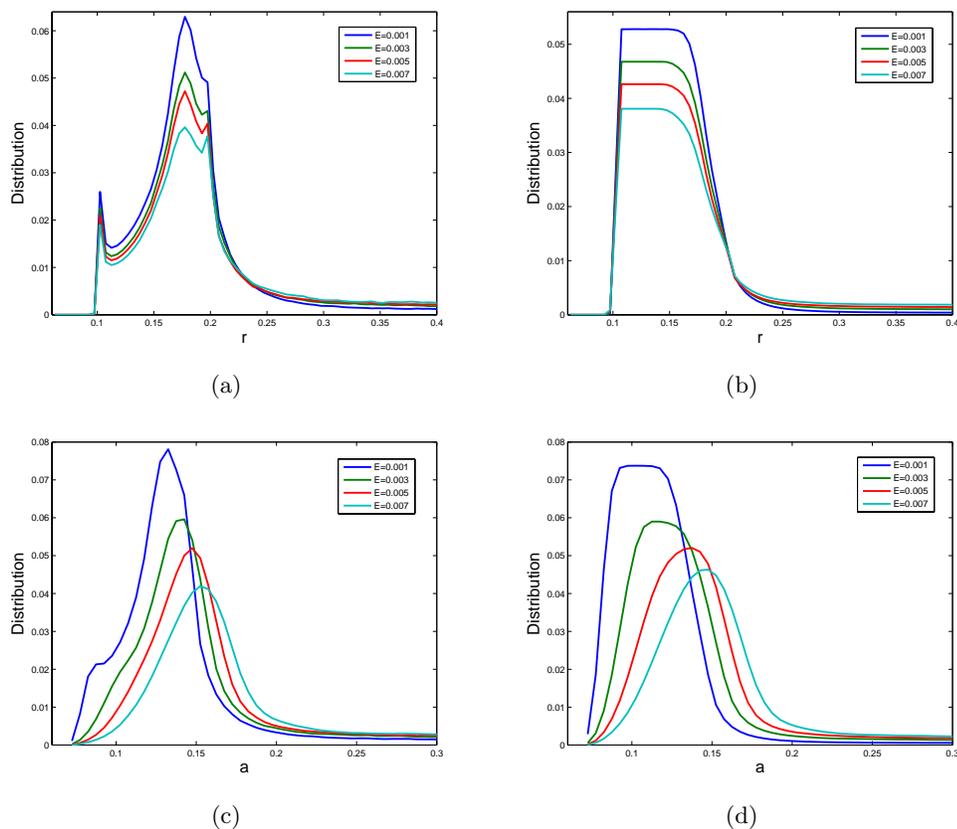


Figure 3.10: (a) Distribution in time along r ; (b) distribution in space along r ; (c) distribution in time along a ; (d) distribution in space along a . See text for explanation.

The distributions along r , (Figure 3.10 a,b), do not change very much in their width as the energy is increased. This is to be expected since this primary inter-particle distance is governed by the Lennard-Jones potential. The differences between distributions in time and space can also clearly be seen, indicating that the primary particles remain in the well of the Lennard-Jones potential, bouncing back and forth.²

In contrast to the distributions along r , the distributions along a , (Figure 3.10 c,d), moved to significantly greater values of a for increasing energy. This is consistent with the fact that the gyration radius is a measure of the extent of the whole system, not just the

²We note that the distribution in space was computed slightly differently to that in Section 3.2.2 since the output time step was not such that it would guarantee that every bin is visited consecutively. Instead, trajectories were assumed to move in a straight line between the positions in successive output time intervals. The bin number of data from successive output time intervals was found and the count in each of the bins between these two bin numbers was increased by one. If the bin number of successive data points were the same then the count in that bin was increased by only one. This indicates where trajectories traverse rather than where they spend most of their time.

primary particles. At higher energies, the extra energy manifests itself as both potential and kinetic energy. The extra potential energy causes the particles to be further apart, thus accounting for the greater a values. The extra kinetic energy leads to faster oscillations but this can not be seen from these distributions, particularly in (d) where the effects of velocity have been eliminated.

It is the form of the distributions in Figure 3.10(c) and (d) that we seek to explain by calculating the effective potential along the gyration radius.

3.3.3.2 Effective Potential

The internal centrifugal force and potential force of Equation 3.29 are shown in Figure 3.11. The local minimum of the internal centrifugal force and the minimum of the potential force move towards increasing values of a for larger energies. Both forces go to zero as a gets larger, which is expected for a dissociating system.

Figure 3.12 shows the potential due to the force terms in Figure 3.11. The potential due to the internal centrifugal term indeed acts like a centrifugal potential pushing the system apart. This effect is greater for larger energies, as is expected. The potential due to the potential term along the gyration radius has similarities with the Lennard-Jones potential since this is the dominant contribution to the potential in Equation 3.2. For larger energies, the potential well gets deeper and wider. The minimum of the potentials in Figure 3.12(b) are all close to $a = 0.1$.

Figure 3.12 illustrates that the effective potential is a balance between two opposing types of forces. The internal centrifugal potential acts to increase the mass distribution in the system, that is, it tends to elongate the system and thereby push it to dissociation. The potential due to the potential term is a confining potential that aims to keep the system compact, that is, keep it from dissociating. The balance between these opposing tendencies depends on the system parameters.

The sum of the potentials due to the internal centrifugal term and the potential term yield the effective potential of the system (Figure 3.13). The effective potentials for the different energies have a Lennard-Jones-like shape with a hard repulsive potential at low values of a and the force going to zero at large values of a . There is one potential well, but the location, depth, and width of the potential well changes with the energy of the system. The higher the energy, the wider the potential well and the further it is towards

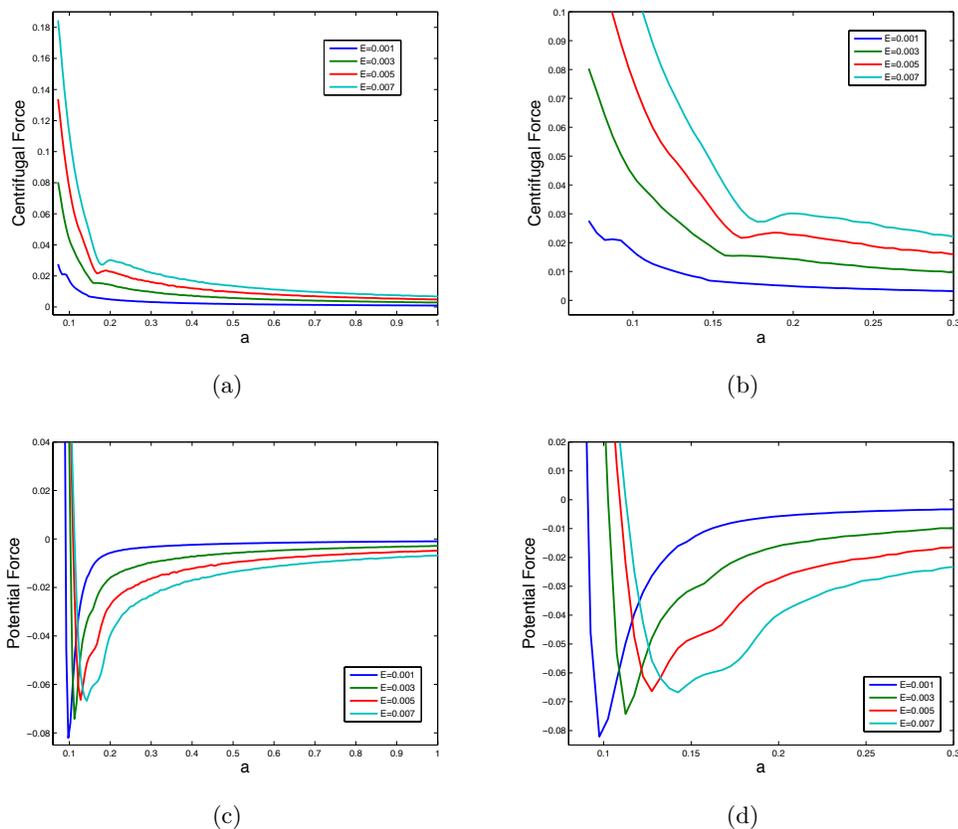


Figure 3.11: (a) Internal centrifugal force along the gyration radius a for different energies; (b) zoom into (a). (c) Potential force along the gyration radius a for different energies; (d) zoom into (c).

larger values of a .

As noted earlier, the potentials due to the potential term all have their minimum in approximately the same location. The internal centrifugal potential has the effect of pushing the system to higher values of a and this is more noticeable for higher energies. It is this internal centrifugal term that has spread out the minimum of the effective potentials in Figure 3.13.

It is notable that the largest and smallest energies considered have the deepest well. For $E = 0.001$ the internal centrifugal term's contribution to the effective potential is small, so the latter is dominated by the potential due to the potential term. The potentials in Figure 3.12(b) for the other energies are all very similar in structure and location. For $E = 0.007$ the internal centrifugal force is quite large and significantly affects the location and shape of the well in the effective potential. It is interesting to note that there is a

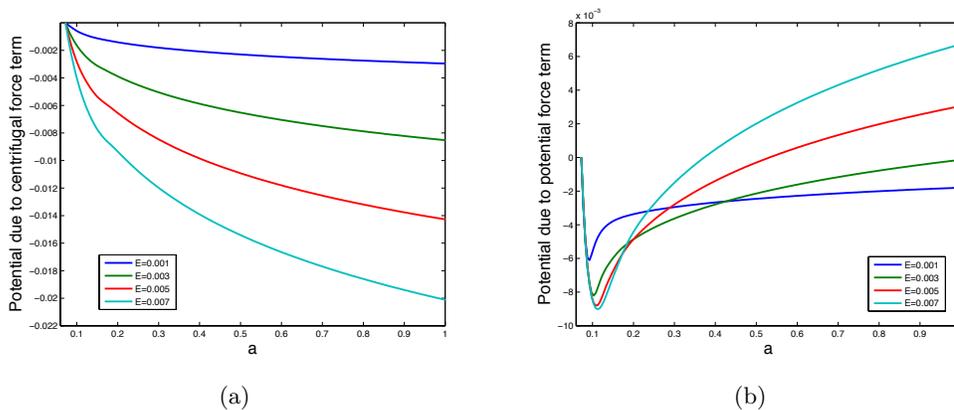


Figure 3.12: (a) Potential due to the internal centrifugal force. (b) Potential due to the potential force.

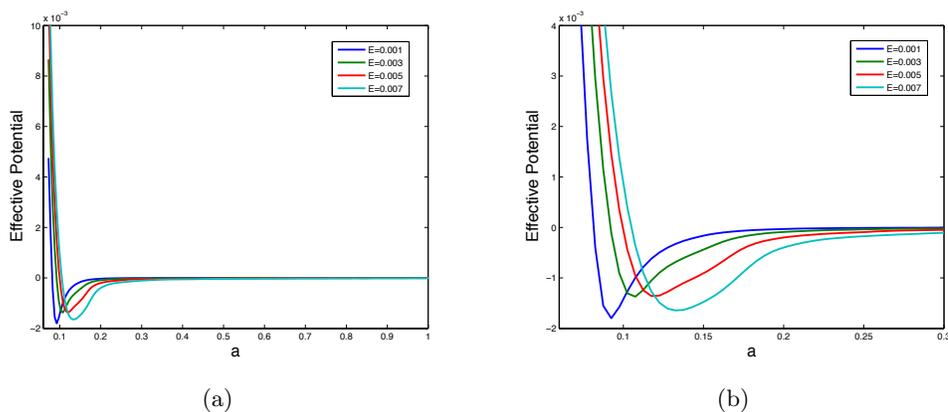


Figure 3.13: (a) Effective potential of the system for different energies; (b) zoom into (a).

pronounced dip in the centrifugal force for $E = 0.007$ in Figure 3.11(b). This is only just noticeable as a slight flattening of the potential in Figure 3.12(a) at the same location near $a = 0.18$. Comparison of these figures reveals that between $E = 0.003$ and $E = 0.005$ the dominant influence on the effective potential changes from the potential term to the internal centrifugal term.

The test of whether the effective potential can explain the dynamics of the dissociation is if it can account for the observed distributions in Figure 3.10(c) and (d). Figure 3.14 shows two of the effective potentials of Figure 3.13 along with the corresponding distribution from Figure 3.10. The effective potentials are shown on the right vertical axis in green and the distributions (in space, that is, taking into account the effects of velocity) on the left

vertical axis in blue. Energies of $E = 0.003$ and $E = 0.007$ are shown; the other energies gave qualitatively similar results.

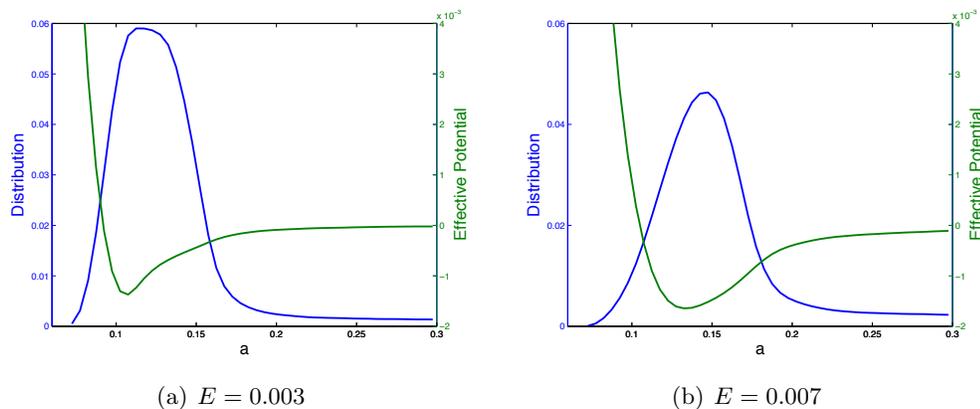


Figure 3.14: Comparison of effective potential and distribution in space for different energies, along the gyration radius. The effective potential is shown on the right vertical axis in green and the distribution on the left vertical axis in blue.

We see that the effective potential along the gyration radius successfully accounts for the observed distribution. The system is trapped in the well of the effective potential before dissociation.

The simple form of the effective potential and the distribution highlight the fact that the gyration radius is a good reaction co-ordinate to use to describe the dissociation dynamics of the system. A comparison with the effective potentials and distributions found using the Kinetic Energy Bias method (Figure 3.3) reveals that the Gyration Radius method is a more elegant way of explaining the reaction pathway, and that it naturally encompasses the dissociation mechanism governing the primary particles. The Gyration Radius effective potentials have only one well rather than two or three, and are smoother. In this way they are a cleaner description of the dissociation.

3.3.3.3 Average Lifetime Before Dissociation

The time at which dissociation occurs is defined here to be the time of the last bounce of the system (see Section 3.2.2). The time of dissociation, t_{last} , is the last time that the gyration radius was below a particular value (depending on the mass and energy parameters) before dissociation.

Figure 3.15 shows the distribution of t_{last} values for different energies. The distributions

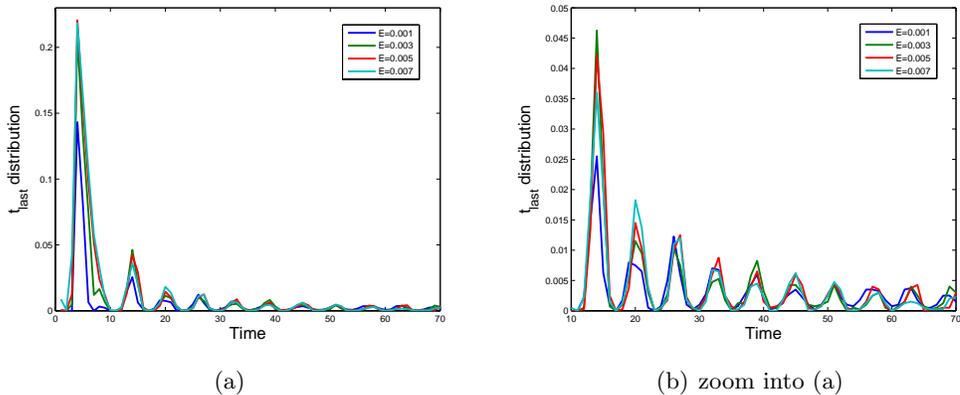


Figure 3.15: Distribution of times of last bounce, t_{last} , of the system before dissociating for different energies.

have clean peaks at regular intervals. These peaks occur at the same dissociation times for all of the energies, up to $t = 70$. For $14 < t < 70$, the successive most common dissociation times (peaks in the distribution) were 2π apart. The period of oscillation of the secondary particle spring potentials is 2π since $k = 0.01$ and the secondary particle mass is $m_s = 0.01$ (period = $2\pi/\sqrt{k/m}$).

The table below shows the percentage of simulations leading to dissociation, that is, the percentage of reactive trajectories, for different energies. It also shows the average lifetime before dissociation. As the energy increases, a greater number of simulations lead to dissociation and the average lifetime decreases. A comparison with the effective potentials in Figure 3.13 reveals that this is unexpected since the potential well is deeper for $E = 0.007$ than for $E = 0.005$. We note that the values in Table 3.1 are based on 4,000 simulations. A different set of 4,000 simulations lead to the percentage of reactive trajectories changing by 5%. Since the percentage of reactive trajectories for energies of $E = 0.005$ and $E = 0.007$ are within 5% of each other, we believe that this unexpected result is an artifact of the particular set of simulations used.

3.3.4 Systems with Different Secondary Particle Masses

In this section we investigate the effects of changing the masses of the bath particles on the dissociation of our model system with total system energy $E = 0.003$. Secondary particle masses of $m_s = 0.001, 0.01, 0.1$ are considered.

Energy	% reactive trajectories	Average t_{last}
0.001	58.5	32.6
0.003	78.6	20.5
0.005	88.7	17.3
0.007	92.5	14.5

Table 3.1: Percentage of simulations leading to dissociation and the average lifetime before dissociation for different energies.

3.3.4.1 Distributions for Different Masses

The distributions in time and space along both r and the gyration radius, a , for different masses are shown in Figure 3.16.³ We see that the distributions along r change only a little in their width as the secondary particle mass changes; however, for higher masses the primary inter-particle distance oscillates higher in the Lennard-Jones potential well. This can be seen from both Figure 3.16(a) and (b). Figure 3.16(a) shows that the primary particles spend most of their time with an inter-particle distance that puts them on the edge of the Lennard-Jones well, where their velocity is slower. Figure 3.16(b) shows this by the width of the broad, flat region of the distribution.

Figure 3.16(c) and (d) show that as the mass increases, the system moves to larger values of a and that it oscillates between a broader range of a values. Since the gyration radius characterizes the mass distribution of the system, it is expected that systems with a higher total mass will have larger a values.

3.3.4.2 Effective Potential

Figure 3.17 shows the internal centrifugal and potential force terms from Equation 3.29. It is interesting to note that the local minimum in the centrifugal force gets more pronounced as the mass is increased (as well as moving to larger a values). The locations of these local minima for the two higher masses are in the vicinity of peaks in the distribution in Figure 3.16(c). Figure 3.17(b) indicates that between $m_s = 0.01$ and $m_s = 0.1$ the potential force changes significantly in its structure along the gyration radius. For the highest mass, the potential well is very wide and shallow.

In Figure 3.18, where we see the potential due to the internal centrifugal term and

³Section 3.3.3.1 explains the differences between distributions in time and space.

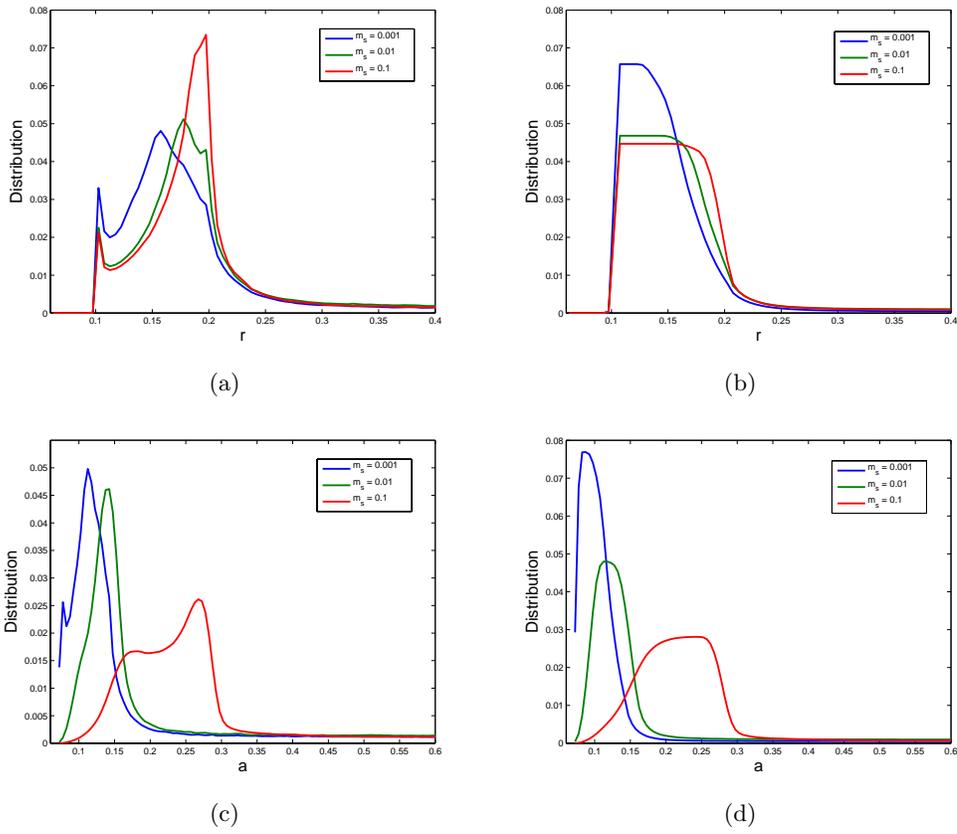


Figure 3.16: Distributions for different masses. (a) Distribution in time along r ; (b) distribution in space along r ; (c) distribution in time along a ; (d) distribution in space along a . See text for explanation.

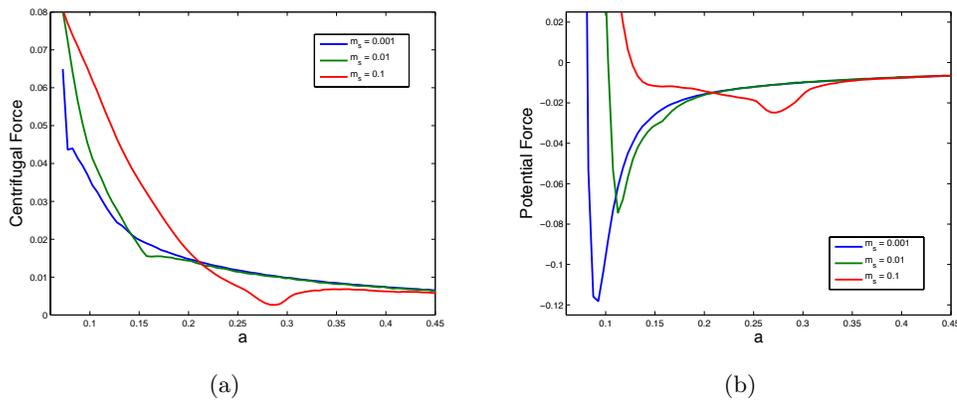


Figure 3.17: (a) Internal centrifugal force along the gyration radius a for different masses. (b) Potential force along the gyration radius a for different masses.

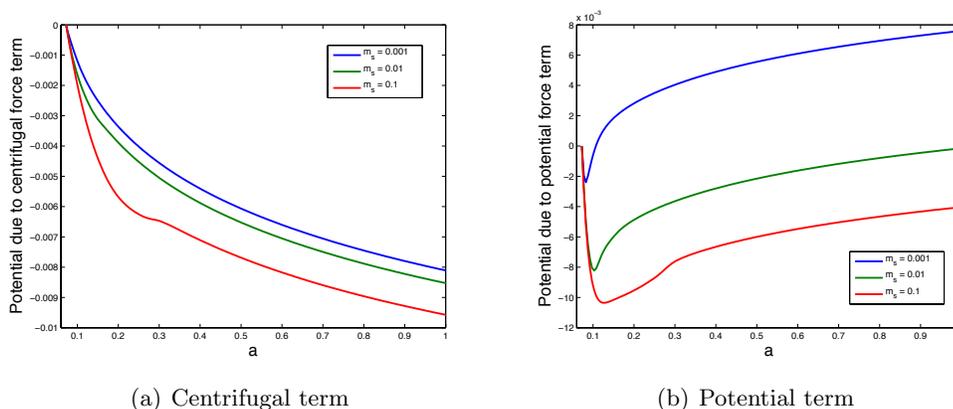


Figure 3.18: (a) Potential due to the internal centrifugal force. (b) Potential due to the potential force.

the potential term, the system with highest mass can be seen to experience a qualitatively different total potential. The potential due to the internal centrifugal term exhibits a pronounced flat region and the potential due to the potential term has a differently shaped well. These features are evident in the effective potential for the system, shown in Figure 3.19. For $m_s = 0.1$ the well is significantly broader. Thus there seems to be a change in the mechanism of dissociation between systems with both particle masses of $m_s = 0.01$ and those with $m_s = 0.1$. Part of this change is accounted for by the increasing influence of the internal centrifugal term and by the nature of this term. In particular, the location of the minimum of the effective potential for $m_s = 0.1$ can be seen to have moved to a larger value of a from the minimum in the potential due to the potential term. This is not true for the other m_s values considered.

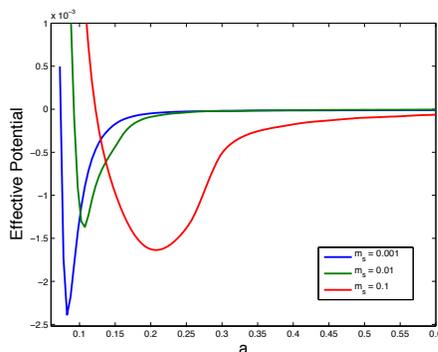


Figure 3.19: Effective potential of the system for different masses.

Based on these observations, and those in Section 3.3.3.2, we postulate that the internal centrifugal term begins to dominate the effective potential once a noticeable flat region develops in the potential due to the internal centrifugal term. Once this flat region develops, its location corresponds to the outer edge of the well in the effective potential along the gyration radius. If the flat region has not developed, the major features of the effective potential are determined by the potential term.

Figure 3.20 compares the effective potential with the distributions in space along the gyration radius. Results for $m_s = 0.001$ and $m_s = 0.1$ are shown (the result for $m_s = 0.01$ can be seen in Figure 3.14(a)). We see that the effective potential successfully accounts for the observed distributions. Furthermore, the effective potential has a simple well structure and is smooth. This was not the case for the effective potential found by the Kinetic Energy Bias method. The fact that the Gyration Radius method has such a clean description of the dynamics indicates that the dissociation behavior can be explained by accounting for the non-Euclidean nature of the internal space. The gyration radius is a good coarse variable to use to describe the dissociation dynamics as it captures these effects.

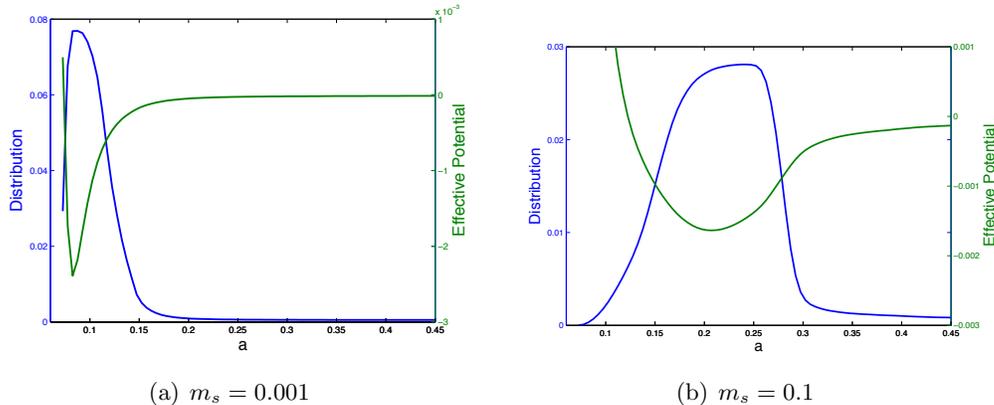


Figure 3.20: Comparison of the effective potential and distribution in space for different masses, along the gyration radius. The effective potential is shown on the right vertical axis in green and the distribution on the left vertical axis in blue.

3.3.4.3 Average Lifetime Before Dissociation

The distributions of dissociation times, as given by t_{last} (defined in Section 3.3.3.3), are shown in Figure 3.21. The most common dissociation times occur at regular intervals

for any particular value of m_s ; however, the distance between peaks is different for the different masses. The times between successive peaks (excluding the first peak) are shown in Table 3.2. Noting that the spring constant is $k = 0.01$, we observe that for $m_s = 0.001$ and $m_s = 0.01$, the time between peaks is equal to the period of oscillation, $T = 2\pi/\sqrt{k/m}$. This is not true for the largest mass, $m_s = 0.1$, where $T = 19.7$. This is consistent with the observation that the mechanism of dissociation is different for the heaviest bath particle mass.

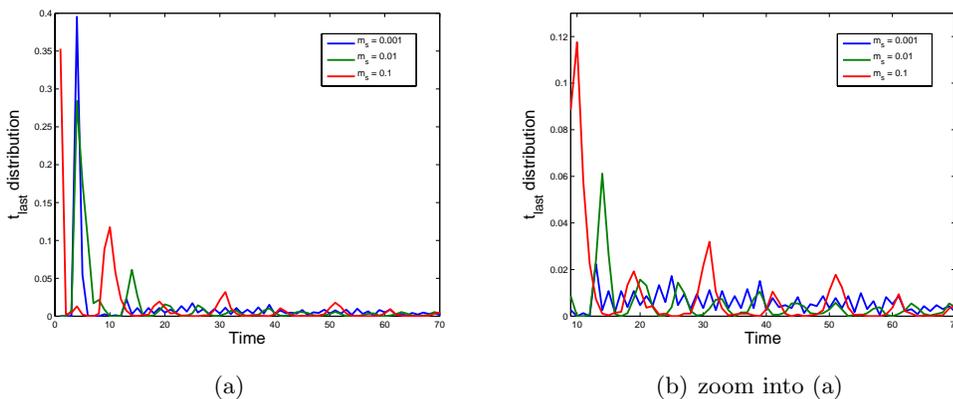


Figure 3.21: Distribution of times of last bounce, t_{last} , of the system before dissociating for different masses.

Table 3.2 also shows the percentage of simulations for which dissociation occurred and the average dissociation time for these reactive trajectories. The system with $m_s = 0.01$ is more reactive than systems with other bath particle masses. The effective potential for this mass has a shallower well.

Bath mass	% reactive trajectories	Average t_{last}	Time between t_{last} peaks	$T = 2\pi/\sqrt{k/m}$
0.001	69.2	34.4	1.98 ± 0.05	1.99
0.01	78.6	20.5	6.15 ± 0.48	6.28
0.1	72.5	21.4	10.16 ± 0.87	19.87

Table 3.2: Percentage of simulations leading to dissociation, the average lifetime before dissociation, the average time between peaks in dissociation times, and the period of oscillation of the secondary particles, for different masses.

We conclude that the influence of the internal centrifugal term in the effective potential becomes greater than that of the potential term once the bath particle mass increases from

$m_s = 0.01$ to $m_s = 0.1$. Evidence for this is the shape of the potential for the heaviest mass in Figure 3.18(a) and the form of the effective potential along the gyration radius in Figure 3.19. A consequence of this is that the time between peaks in the distribution of dissociation times is no longer given by the period of oscillation of the secondary particles.

The internal centrifugal term is responsible for the dynamical coupling of the bath modes to the reactive mode. As a result of this, we expect that the influence of the internal centrifugal term on the effective potential is also dependent on the spring constant of the spring potentials that couple the bath particles to one of the primary particles. This is an avenue for further investigation.

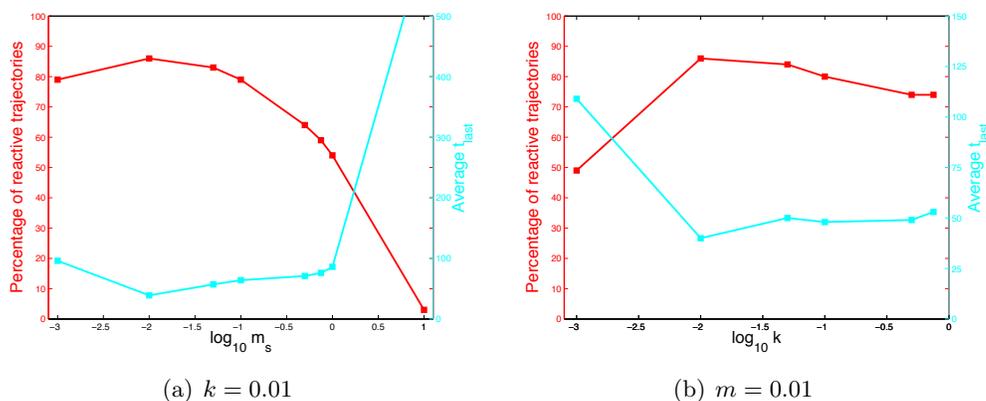


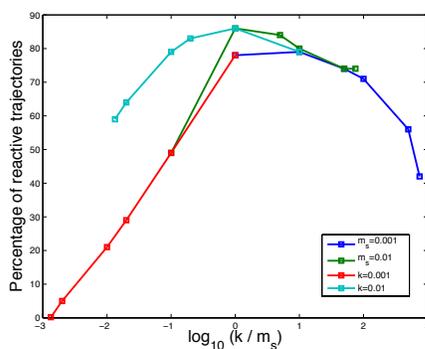
Figure 3.22: (a) Variation in reactivity with mass. Percentage of simulations that resulted in dissociation, for different masses, are shown on the left vertical axis in red. The average time to dissociation is shown on the right vertical axis in light blue. The horizontal axis is the logarithm to base 10 of the bath particle masses. (b) Variation in reactivity with spring constant, k . Vertical axes are the same as in (a). The horizontal axis is the logarithm to base 10 of k , the spring constant, that couples the bath particles to the primary particle. Each set of (m_s, k) parameters used 10,000 simulations with total energy $E = 0.0036$.

Figure 3.22 shows the variation in reactivity of the system as the bath particle parameters are changed. Figure 3.22(a) considers different secondary particles masses from $m_s = 0.001$ to $m_s = 10$ for a spring constant of $k = 0.01$. The logarithm to base 10 of these masses is shown on the horizontal axis. The vertical axes show the percentage of simulations that lead to dissociation on the left in red, and the average time to dissociation, as given by t_{last} , on the right in blue. We see that $m_s = 0.01$ leads to dissociation most often and is the fastest to dissociate. In contrast to the results presented earlier in this section, the results in Figure 3.22 are shown for 10,000 simulations for each set of (m_s, k) parameters

with an energy of $E = 0.0036$, and not the 4,000 simulations at an energy of $E = 0.003$ used previously.

Figure 3.22(b) considers different strengths of the coupling between the secondary particles and the primary particle that they are attached to. This is given by the spring constant k of the spring potential between them. Values of $k = 0.001$ to $k = 0.75$ were considered for a secondary particle mass of $m_s = 0.01$. The logarithm to base 10 of these spring constants is shown on the horizontal axis. The vertical axes are the same as in (a). For this mass, the most reactive system has $k = 0.01$ —it dissociates most often and the fastest.

Figure 3.23 shows the variation in reactivity with respect to k/m_s as either the mass or the spring constant is kept constant. The system is most reactive when $k = m_s = 0.01$. These simulations had a total energy of $E = 0.0036$. The examination of systems with different total energies and different bath masses in the last two sections has revealed that there is a competition between the internal centrifugal force and the potential force. The nature of the dissociation dynamics is determined by which of these is stronger for the given system parameters. In Section 3.3.3, as the energy was increased, the internal centrifugal force became dominant. This was also the case in the present section with a larger bath particle mass. Future work aims to define the region of (m_s, k, E) -space in which the internal centrifugal force is the dominant contributor to the effective potential governing the dissociation dynamics.



(a)

Figure 3.23: Variation in reactivity with k/m_s . The horizontal axis shows $\log_{10}(k/m_s)$. The different colors represent quantities that are kept constant while the other parameter is varied from 0.001 to 0.75. For example, the red line corresponds to $k = 0.001$ with masses varying from $m_s = 0.001$ to $m_s = 0.75$. The system energy is $E = 0.0036$.

3.3.5 The Internal Centrifugal Term in a Simpler System

The internal centrifugal term in the equation of motion along the gyration radius captures the kinematic effects that are not included in the potential force term. For some systems, this internal centrifugal term plays a larger role than the potential term in determining the dynamics. In this section we investigate a simpler system that exemplifies this.

The system that we consider is a modification of the model introduced in Section 3.1, with only two secondary particles instead of four. These secondary particles have a mass of $m_{3,4} = 0.1$ and the primary particles have mass $m_{1,2} = 1.0$. The total system energies that we consider are $E = 0.0012, 0.0030, 0.0036, 0.0042,$ and 0.0072 .

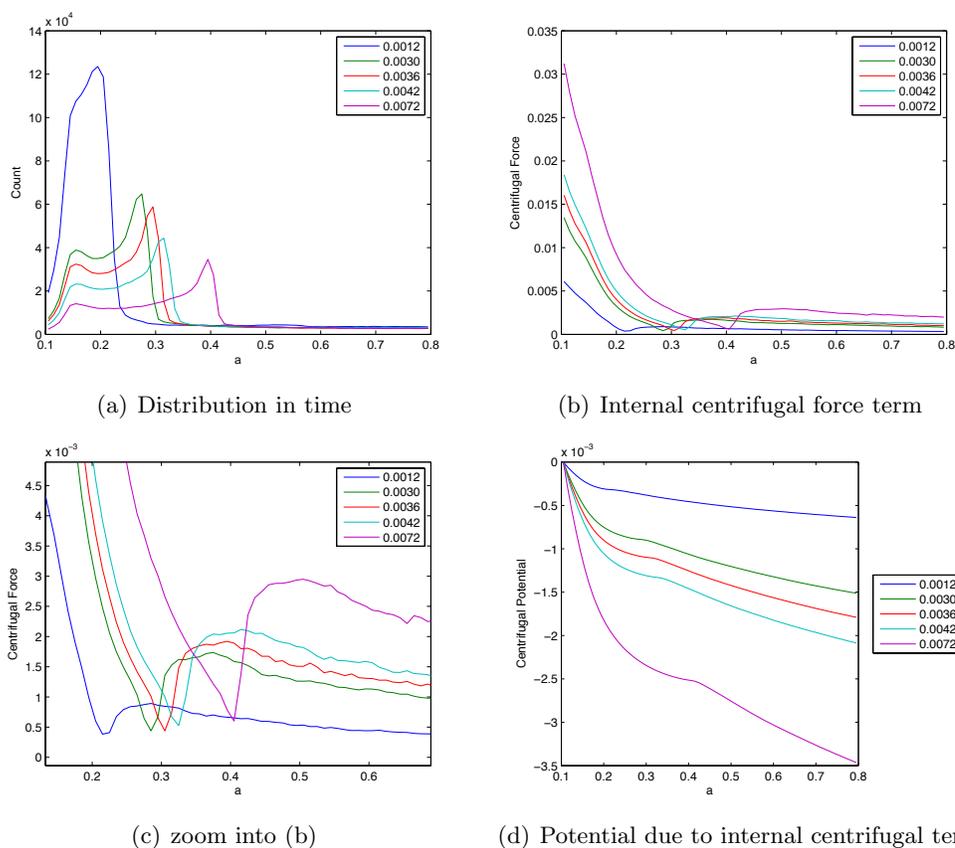


Figure 3.24: Distributions, internal centrifugal force, and internal centrifugal potential for different energies in a system with two secondary particles.

Figure 3.24(a) shows the distribution in time along the gyration radius, using both reactive and non-reactive trajectories. The distributions along a change with increasing energy, forming a bi-modal distribution above a certain energy threshold. The peak of the

distribution at lower values of a does not shift for larger energies, however the other peak moves to increasing a values for larger energies. This is consistent with having a well in the effective potential that has a hard repulsive potential, like the Lennard-Jones potential, at low a values. The well in the effective potential along the gyration radius gets wider with increasing energy. In this distribution in time, the peaks correspond to turning points as the system bounces back and forth within the effective potential well. It is for this reason that we look at the distribution in time here and not the distribution in space, the former giving a clear indication of the edge of the effective potential well.

Figure 3.24(b) and (c) show the internal centrifugal force term in the equation of motion along the gyration radius. Figure 3.24(d) shows the potential due to this internal centrifugal term. For all energies apart from the lowest one, there is a pronounced flat region in the internal centrifugal potential. For the lowest energy we see this flat region beginning to develop. The internal centrifugal force tends to increase the gyration radius by pushing the primary particles further apart. This effect is more pronounced for higher energies since it is a kinematic effect and there is more kinetic energy available at higher total energies. A comparison of Figure 3.24(a) and (c) reveals that the minimum in the internal centrifugal force term coincides with the location of the second peak in the distribution, for all energies. This strong correlation is evidence for the internal centrifugal force being the dominant force in this system, determining the width of the well in the effective potential. Further, the existence of the flat region in the internal centrifugal potential is consistent with our hypothesis that once this flat region develops, the internal centrifugal force will be the major contributor to the dissociation dynamics of the system.

3.4 Conclusion

In this chapter we have seen that for a correct description of the dynamics of transition in our dissociating system, the kinematic effects of the secondary particles must be taken into account. These effects are generally applicable in molecular transitions and frequently lead to observable effects that explain the transition dynamics.

In the Kinetic Energy Bias method, these kinematic effects are captured by an effective potential, which is calculated by quantifying the break-down of the law of equipartition of energy and accounting for the entropic force that this break-down leads to. The Gyration

Radius method shows that the kinematic influence of the bath particles is contained in the internal centrifugal force that is derived from the non-Euclidean nature of the internal shape space. The metric force due to the non-Euclidean internal space begins to dominate the force due to the potential of the system for higher energies and larger masses, as shown in Sections 3.3.3 and 3.3.4. The nature of the dissociation dynamics is determined by the competition between the internal centrifugal force and the potential force. The latter generally works to maintain the inter-atomic bond, whereas the internal centrifugal force works to break the bond. For some systems, such as the simpler one considered in Section 3.3.5, the internal centrifugal force dominates at all energies that we considered.

We have shown that the system parameter values at which the internal centrifugal force is dominant are those at which the internal centrifugal potential has a flat region. The location of this flat region corresponds to the edge of the well in the effective potential along the gyration radius.

Both the Kinetic Energy Bias method and the Gyration Radius method give the following physical insight: It is the redistribution of kinetic energy from the bath modes to the reactive mode that enables and governs the dissociation. The point at which the primary particles can be said to have dissociated is best given by the Gyration Radius method as the gyration radius value at the edge of the well in the effective potential. Dissociation can not be said to have definitely occurred if the distance between the primary particles is greater than the value at the right-hand edge of the Lennard-Jones potential. The Kinetic Energy Bias method indicates that beyond this primary well, there are secondary wells along r in which the system can be caught. These wells give the distances at which the primary particles can be temporarily caught; however, the most straightforward cutoff value at which dissociation occurs is given by the Gyration Radius method.

The most common times at which dissociation occurs (peaks in the distribution of dissociation times) was shown to depend on the mass of the secondary particles. For systems in which the potential force is dominant, peaks in the dissociation time distribution occurred every T time units, where T is the period of oscillation of the secondary particles. When the internal centrifugal force is more dominant, this no longer holds—that is, peaks in the distribution are not correlated with the period of oscillation.

The dissociation reactions that we have considered show that the Gyration Radius formalism provides a good coarse variable to use as the reaction co-ordinate and that the

effective potential calculated along this reaction co-ordinate explains the observed dissociation behaviour. The effective potential found in this way has a relatively simple form (one potential well) and is much smoother than the effective potential found in Section 3.2 using the Kinetic Energy Bias method. It is also much cheaper, computationally, to find the effective potential using the Gyration Radius method since all data points can be used, whereas in the Kinetic Energy Bias method, many data points were discarded as they did not lie along the reaction co-ordinate, which required the kinetic energy in the bath mode to be almost zero. For the results presented here, the Kinetic Energy Bias method used 25 times more simulations.

A further advantage of the Gyration Radius framework is its generalizability to higher-dimensional systems than the collinear one considered here. This is applicable in any system in which a change in the mass distribution can account for the transition of interest. The coarse variables would be the three gyration radii which are the singular values of the matrix of Jacobi vectors (Equation 3.23).

The Kinetic Energy Bias method and the Gyration Radius method together provide a detailed understanding of the dissociation dynamics in this system. Both methods highlight the importance of considering kinematic effects in this reaction.

A promising method for overcoming the problem of dimensionality in larger chemical systems is to combine the methods presented in this chapter with those of the previous, namely, to use the gyration radii of the system as a coarse variable and the effective potential along this variable as the model for the set-oriented method to use to compute transition rates. We believe this will be a fruitful combination of approaches.

Chapter 4

Lattice Quality Assessment Tools

In this chapter and the next, we move away from chemical reactions with a small number of constituents and instead focus on a large number of particles in the plane and their transition from a disordered to an ordered state. The transition can be captured by metrics that assess the quality of the patterns formed.

The structure of crystalline solids has been of interest to the condensed matter and materials science fields for a long time. The defects present in these solids determine many of the physical properties of the material. Investigations have been both theoretical and experimental, concentrating on how different types of defects, and the number of them, affect mechanical, electrical, and optical properties [50]. A large number of experiments have been performed that attempt to make materials that are free of defects, or examine the formation energies and movement of defects. The number of defects and their type is commonly deduced from bulk properties of the material, such as diffraction patterns or shear stress [3]. The focus is on the material as a whole rather than on the region surrounding each constituent particle.

This chapter is concerned with a quantitative assessment of the quality of two-dimensional lattices. We are interested not in measuring a bulk property of the whole lattice as it responds to some external stimulus, but rather assessing the positions of the constituent particles in a manner similar to the human eye. The aim is to quantify what the eye sees when comparing two lattices and deciding that one is better than the other. The notion of “better” may depend on which property of the lattice is more important to the assessor or the goal in assessing the lattice. The measures of the defects of a lattice that we have

developed are thus concerned with the local neighborhood of each particle, reflecting the eye's propensity to judge sub-regions of the lattice and how these regions combine.

We describe and compare five defect metrics that have this local nature. The most versatile of these, the Defect Measure, is used as a tool in applications such as those that arise in designing an isotropic inter-particle potential that leads to the self-assembly of particles into a specified lattice. All particles are identical and move in a finite two-dimensional area. This and other applications, including those of the other metrics, are presented in the next chapter.

The motivation for the development of the quality assessment tools was the self-assembly problem. The self-assembly of particles is of importance in the diverse fields of understanding how biological or chemical components form a coherent whole and multi-vehicular surveillance. In the latter, large numbers of small vehicles moving in a lattice formation may be an efficient way of surveying the landscape. The local deviations from a perfect lattice formation must be understood in this context in order to deal with issues such as loss of communication with neighboring vehicles and coverage of the area to be surveyed. Vehicles that communicate only with their neighbors should be able to assemble into and maintain a lattice formation, eliminating the need for each vehicle's specific trajectory to be programmed. The design of the ideal isotropic potential for doing so requires an effective measure of the defects in the lattice, as does the evaluation of the robustness of that potential.

4.1 Types of Lattice Defects

There are a number of different types of defects that can occur in a monatomic two-dimensional lattice [64]. The defects most commonly referred to are listed below.

- *Vacancy*: a lattice site that should have a particle is unoccupied.
- *Interstitial*: a particle occupies a lattice site that should not have a particle occupying it.
- *Frenkel pair*: a vacancy and an interstitial are nearby. A particle is at a lattice site which should be unoccupied, leaving a nearby lattice site that should be occupied empty.

- *Topological defect*: a region in a lattice where the ordered structure is different from the rest of the lattice. For example, in a honeycomb lattice, a region that has five particles in a unit cell rather than the required six.
- *Split interstitial*: two particles share a lattice site, typically by having their centre of mass at the lattice site where there should be one particle.
- *Edge dislocation*: an extra line of particles inserted part of the way into the lattice. The adjacent lines of correctly ordered particles bend around the line that terminates. Dislocations are breaks in the translational symmetry of the lattice.
- *Disclinations*: a line defect that results in a rotation if the orientation of the lattice around the defect is tracked.
- *Grain boundaries*: regions, typically lines, where the orientation of the lattice changes abruptly. Frequently caused by two sub-lattices growing separately and then meeting.

An effective quality assessment metric must capture these different kinds of defects.

4.2 Metrics for Assessing the Quality of Lattices

We describe five metrics for assessing the quality of lattices that focus on the local configuration of particles. These metrics are compared in the following section.

4.2.1 Defect Measure

The Defect Measure is a tool that was developed in order to compare the quality of lattices. The human eye is frequently a good judge of the quality of a lattice, however, a more quantitative assessment was sought in order to make this assessment more efficient and objective. The motivation was the need to provide a numerical assessment of the quality of lattices that are formed during the optimization procedure used for finding a potential that leads to the self-assembly of particles in a plane. This application of the Defect Measure is discussed in Sections 5.1.3 and 5.1.4.

Given the positions of particles in a plane, the quality of the lattice that is formed is determined by the desired or target lattice. If the target lattice is known, calculating the Defect Measure requires only Step 3 in the procedure described below. If the target lattice

is not known, or if it is necessary to find the type of lattice that the system of particles is attempting to form, then Steps 1 and 2 can be applied to determine what the target lattice is. Identification of a target lattice means identifying both the type of lattice (honeycomb, triangular, etc.) and the minimum lattice spacing, called the *lattice constant*, a . That is, identifying the target lattice involves identifying the shape and the scaling.

The algorithm for identifying the target lattice and computing the Defect Measure does not need the positions of all particles in the plane. It only requires a list of distances to the nearest m neighbors of each particle. To distinguish between the four types of lattices considered—triangular, honeycomb, square and kagome¹—it is only necessary to consider the distances to the nearest 15 neighbors of each particle,² however the results are more reliable (as the identification of the target lattice is improved) if more nearest neighbors are considered. We use $m = 20$.

This is an important feature for our surveillance example. Vehicles would only need to detect other vehicles that are in a certain range that covers an area in which there would be approximately m other vehicles. The direction of each detected vehicle is unimportant, only the distance to that vehicle. In this way, the position of each vehicle does not need to be tracked.

In the identification steps (Steps 1 and 2), it is assumed that the lattice is reasonably well formed. This assumption effectively means that the human eye would be able to distinguish the type of target lattice.

The number of particles in the lattice to be assessed must be large enough such that there are more interior particles than boundary particles. Note that “boundary” here refers to the boundary of the lattice and not the boundary of the domain containing the particles. The larger the number of inner particles with respect to boundary particles, the better the algorithm is able to identify the target lattice. The issue of identifying the boundary particles is discussed in Section 4.2.1.3.

¹A kagome lattice has is composed of hexagrams in which each particle has four nearest neighbors

²See Table 4.1.

4.2.1.1 Procedure for Identifying the Target Lattice and Computing the Defect Measure

Step 1: *Identify the lattice constant, a*

1. List the inter-particle distances d_{pj} for particle p in ascending order.
2. Find clusters in the distances for each particle.
3. Consider the cluster of shortest distances for each particle p . (a) Average over all of the values in all of these clusters to find a , or (b) use all of these values to construct a histogram of shortest distances and define a to be the maximum in the distribution of shortest distances (that is, the most common shortest distance).³

Step 2: *Identify the type of lattice*

1. Count the number of values in the first three clusters identified in Step 1.2, for each particle.
2. Find the mode of the number of particles in the first clusters (clusters of shortest distances). Repeat for the second and third clusters.⁴ The mode for the first clusters will be referred to as the number of particles at a distance given by the lattice constant, a , found in Step 1.
3. Compare the number of particles at each distance (modes of number of particles in each successive cluster) with the known values for possible types of lattices.

Mode of:	cluster 1	cluster 2	cluster 3
Triangular	6	6	6
Square	4	4	4
Honeycomb	3	6	3
Kagome	4	4	6

Table 4.1: Number of particles in perfect lattice.

³Method (b) for finding the lattice constant is the most accurate and will be used here.

⁴This step works because for a majority of particles, the number of values in each successive cluster should be the same.

Step 3: *Compute the Defect Measure*

Compare the given lattice to a perfect lattice of the same type with the same lattice constant, to find a measure of the defects, that is, the quality of the lattice.

For a particle in the target (perfect) lattice, find the distance r that is halfway between the distance to the closest neighbors and the next closest neighbors. This distance is shown as the red circle in Figure 4.1. For a triangular lattice, there are six closest neighbors at a distance of a and six next closest neighbors at a distance of $\sqrt{3}a$. The red circle has a radius of $r = (1 + \sqrt{3})a/2$.

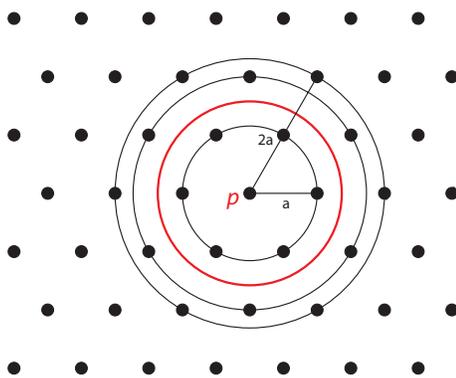


Figure 4.1: The nearest neighbors circle (red circle) of particle p is halfway between the closest particles and the next closest particles.

Define the *nearest neighbors* of a particle p to be those particles that are within a radius r of particle p .

1. Choose weights, $\omega_{\text{defect type}}$ for each type of defect (see discussion below).
2. For each particle, p , construct the nearest neighbors circle and compute the Defect Measure of that particle according to which of the following types of defects apply

(shown in Figure 4.2):

$$\begin{aligned}
 (\text{Defect Measure})_p &= \omega_{\text{displaced}} \times \left(\sum_{j \in \text{nearest neighbors}} ff_{jp} \times (d_{pj} - a)^2 \right) \\
 &+ \omega_{\text{missing}} \times n_{\text{missing}} \times a^2 + \omega_{\text{extra}} \times n_{\text{extra}} \times a^2 \\
 &+ \eta_{\text{lone}} \times \omega_{\text{lone}} \times a^2 + \eta_{\text{boundary}} \times \omega_{\text{boundary}} \times a^2 \quad (4.1)
 \end{aligned}$$

where the η s are indicator functions for lone and boundary particles. In the first term, ff_{jp} is the *fade factor* for particle j with respect to particle p . The fade factor allows particles to fade out of view of particle p rather than disappear as they cross the nearest neighbors circle. For triangular and honeycomb lattices, the fade factor is equal to 1 for $d_{pj} < (5 + 3\sqrt{3})a/8 \approx 1.27a$ and is equal to 0 for $d_{pj} > (1 + \sqrt{3})a/2 \approx 1.37a$ (the nearest neighbors circle). The fade factor decreases from 1 to 0 over a distance that is equal to a quarter of the distance from a to the nearest neighbors circle. Within this region, the fade factor decreases in a cubic polynomial fashion with horizontal tangent at the end points of the region.

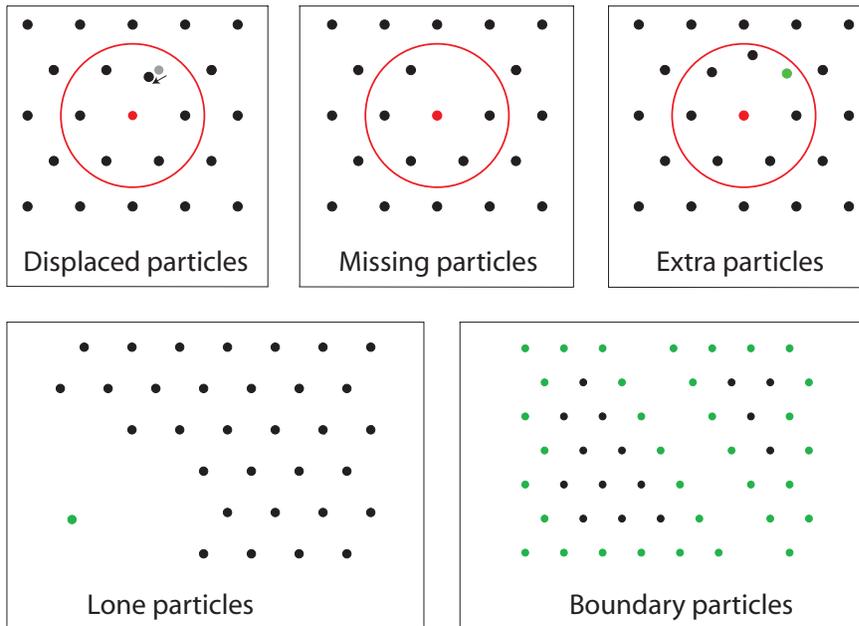


Figure 4.2: Types of defects used in computing the Defect Measure.

3. The Defect Measure for the lattice is given by summing over all particles p :

$$(\text{Defect Measure for Lattice}) = \sum_p (\text{Defect Measure})_p.$$

The Defect Measure is lower for higher quality lattices. A perfect lattice will only have a Defect Measure equal to zero if the weight for boundary particles is set to zero.

The weights, $\omega_{\text{defect type}}$, for each type of defect are chosen according to the severity of the defect. This depends on the goal. For example, in the surveillance example, if the collision avoidance of the vehicles is an issue, then a larger weight would be given for extra particles, ω_{extra} , in order to deter more strongly lattices with extra particles. When optimizing the potential for the self-assembly problem, a larger weight for boundary particles, ω_{boundary} , may be necessary to penalize the formation of distinct sub-lattices. Note that boundary particles are not penalized for missing particles in their nearest neighbors circle (see Section 4.2.1.3). The effect of changing the weights for the various types of defects will be discussed further in Section 4.3.

One of the advantages of the Defect Measure is that it allows for flexibility to penalize different types of defects more heavily. In this way, it is a tool that can be shaped for the specific task at hand.

Note that all of the varieties of defects discussed in Section 4.1 are taken into account by the types of defects listed in Step 3.2. For example, grain boundaries are penalized by contributions to the Defect Measure from displaced, missing, and extra particles.

The Defect Measure is composed of two factors: the local density around each particle (which is compared to that of the target lattice), and the distance between particles. The geometry of the particles is not explicitly considered. This is sufficient because, for any potential that could lead to self-assembly, if the correct number of particles are put into a region the size of the nearest neighbors circle, they will arrange themselves into approximately the correct configuration due to the potential between them. For this reason, only the inter-particle distances are required. In Section 4.3 we discuss the ability of the Defect Measure to be tailored to reflect the geometry of the particles.

4.2.1.2 Advantages of the Defect Measure

The Defect Measure has a number of advantages over other methods for quantifying the quality of lattices. Although the quality metrics described in the remainder of this section may share some of these advantages, none of the metrics share all of them.

- The Defect Measure gives a *local* assessment of the quality of a lattice. Apart from leading to the versatility of the Defect Measure in applications, a local assessment of a lattice is closer to the qualitative assessment that a human eye would make of a lattice.
- Each particle's contribution to the quality of the lattice can be quantified. In this way, regions of the lattice that are not well formed can be identified. This is useful in applications such as those discussed in Sections 5.3 and 5.4.
- The primary types of defects, and the number of such defects, that occur in a lattice can be easily identified.
- The flexibility of the Defect Measure due to the assignment of weights to defects, leads to a versatility that may be exploited in applications. This flexibility will be discussed further in Section 4.3.
- The target lattice does not need to be specified.
- The Defect Measure is invariant under rotations, reflections, and translations of the lattice.

4.2.1.3 Identifying the Boundary of a Lattice

To implement the procedure for calculating the Defect Measure (with non-zero weights for the boundary particles), the particles that form the boundary of the lattice must be identified using only the distances to the nearest m neighbors ($m = 20$ for us).

The boundary particles of a perfect lattice can be identified by counting the number of particles at a distance of a and the number of particles within a distance of $2a$ from each particle. For a triangular lattice, an inner particle has 6 neighbors at a distance of a , 6 neighbors at a distance of $\sqrt{3}a$, and 6 neighbors at a distance of $2a$. Boundary particles can have a maximum of 5 particles at a distance of a and 15 particles within a distance of

$2a$. This is shown in Figure 4.3. Particles that satisfy these conditions are designated as boundary particles. Note that these requirements specify the maximum concavity of the boundary that can be detected. A similar construction applies to different types of lattices.

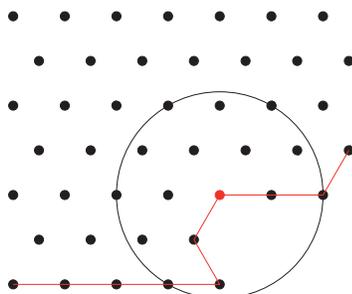


Figure 4.3: A boundary particle in a perfect triangular lattice can have a maximum of 5 particles at a distance of a and 15 particles within a distance of $2a$. The red particle is such a particle. The red line indicates the boundary of the lattice. The circle encloses particles that are within a distance $2a$ of the red particle.

For an imperfect triangular lattice, the condition of having a maximum of 5 particles at a distance of a is loosened to having a maximum of 5 particles within the nearest neighbors circle. The second condition is relaxed to having a maximum of 15 particles within a distance $(1 + \sqrt{7}/2)a$ of the candidate boundary particle. This distance is halfway between $2a$ and the distance to the next nearest neighbors (at $\sqrt{7}a$). Similarly relaxed conditions apply to other types of lattices.

Identifying the boundary of a lattice given the positions of particles. If a system of particles is not a sufficiently well-formed lattice and the goal is to compute the Defect Measure of the particles with respect to some specified target lattice, then it will be necessary to use a different method to identify the boundary particles. In this situation, the position of each particle is required, as the number of particles in the circles described above may not be comparable to that of the target lattice.

The convex hull of a set of points in the plane is the minimal convex set containing all the points. It may be visualised as the shape of an elastic band that has been stretched to encompass all the points and then allowed to collapse around them. The convex hull for a set of points and the points identified as boundary points by this convex hull are shown in

Figure 4.4.

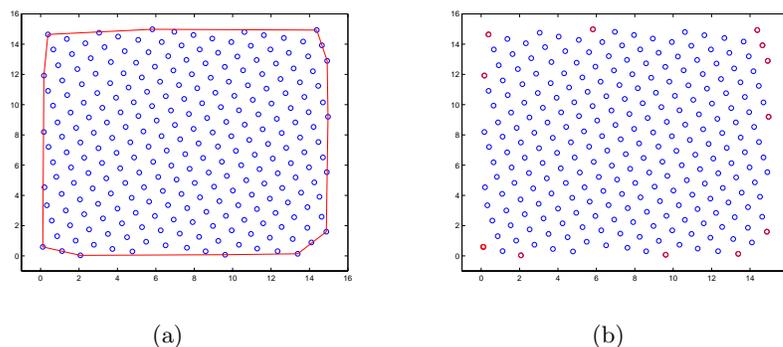


Figure 4.4: (a) The convex hull of a set of points in the plane. (b) The points identified by the convex hull as boundary points (shown in red).

Clearly, this is not what the human eye identifies as the boundary. Instead, the non-convex hull correctly identifies the boundary particles. First, a minimum concave curvature, ρ , must be chosen. The non-convex vertices of the boundary are those particles that are touched by a disk of radius ρ as it is rolled around the outside of the set of points. The following algorithm for finding the non-convex hull is due to Boje [8].

Algorithm for finding the non-convex hull (Boje).

1. Find the Delaunay triangulation⁵ of the set of points.
2. Find the outside triangles, that is, those triangles with an edge that does not touch another triangle's edge. Such edges together form the convex hull.
3. Recursively delete any outside triangle that has an outside edge longer than 2ρ .
4. Recursively delete outside triangles whose outside edge is the longest edge of the triangle and whose circumcircle⁶ has a radius greater than ρ .
5. Iterate steps 3 and 4 until no more triangles are deleted.

The points identified by this algorithm as the particles on the non-convex hull of the set of points in Figure 4.4 are shown in Figure 4.5 in red. These are the boundary particles.

⁵The Delaunay triangulation of a set of points is a triangulation such that no point is inside the circumcircle of any triangle in the triangulation [21]. It is the dual graph of the Voronoi tessellation of the points.

⁶The circumcircle of a polygon is a circle that passes through all of the vertices of the polygon.

Note that the boundary points identified by the convex hull corresponds to the points found by rolling a disk of infinite radius around the set of points.

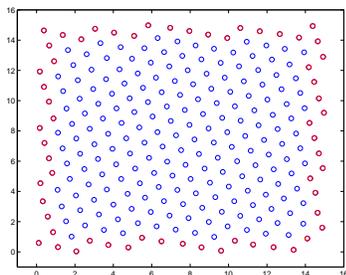


Figure 4.5: The points identified by the non-convex hull as boundary points (shown in red).

4.2.2 Geometric Defect Measure

The Defect Measure takes into account the local density of particles and the distance between particles; it does not consider local geometry. Despite this, as discussed above, the Defect Measure provides a good assessment of the quality of a lattice. The Geometric Defect Measure was developed as an alternative quality assessment tool that could be compared to the Defect Measure (with appropriate weights) to check that looking at local densities (with a correction for displacements) does indeed lead to a lattice with the correct geometry.

Computation of the Geometric Defect Measure requires the position of each particle in the system as well as the type of target lattice. It focuses on the shape of the lattice and not the scaling. Thus, two lattices that differ only by a scaling of the lattice constant will have identical values for the Geometric Defect Measure. It is computationally more expensive than the calculation of the Defect Measure.

The algorithm for computing the Geometric Defect Measure of a honeycomb lattice is outlined below and illustrated in Figure 4.6. The procedure is similar for other types of lattices.

Algorithm for computing the Geometric Defect Measure of a honeycomb lattice.

1. Find the nearest neighbor of particle p . Label it $nn1$. Let the distance between particles p and $nn1$ be d_{nn1} .

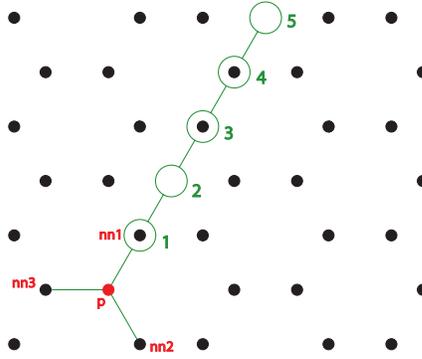


Figure 4.6: Construction used in computing the Geometric Defect Measure for particle p .

2. Extend the line from p to $nn1$ a distance of d_{nn1} . Determine whether there is a particle within a distance $d_{nn1}/8$ of this point, that is, whether there is a particle in region 2 in Figure 4.6.
3. Continue extending the line from p to $nn1$ in units of d_{nn1} and determining whether a particle is within a region around the end points. Do this for a total of s regions or steps from particle p . (Figure 4.6 shows $s = 5$.)
4. Compare whether or not a particle is found in each region to whether or not it is expected to be found if it were a perfect lattice. For the honeycomb lattice, there should be particles in regions 3 and 4, but not in regions 2 and 5. (Region 1 will have a particle by construction.)
5. If a region j does not have the correct number of particles then $m_{p1j} = 1$, otherwise $m_{p1j} = 0$.
6. Repeat steps 2–5 for the second and third nearest neighbors, $nn2$ and $nn3$.
7. Calculate the angle between pairs of nearest neighbors of p : θ_{p12} , θ_{p13} , and θ_{p23} .
8. Sum over all particles p in the lattice to obtain the Geometric Defect Measure of the lattice.

$$\text{Geometric Defect Measure} = \sum_p \sum_{i=1}^{nn} \left(\sum_{j=1}^s m_{pij} + \sum_{k=2, k>i}^{nn} |\cos \theta_{\text{perfect}} - \cos \theta_{pik}| \right) \quad (4.2)$$

where θ_{perfect} is the angle between pairs of nearest neighbors of p in the perfect lattice and nn is the number of nearest neighbors. For the honeycomb lattice $\theta_{\text{perfect}} = 120^\circ$ and $nn = 3$.

If the domain that contains the system of particles does not have periodic boundary conditions then the extension of the lines from particles to their nearest neighbors should be cut off when a boundary of the domain is reached.

A number of modifications to the algorithm outlined above are appropriate for most lattices. These modifications have been found based on experimentation. First, the number of steps, s , that the lines are extended may be increased or decreased, depending on the type of lattice. For the triangular lattice, looking at only three regions is sufficient to give a quality assessment that is comparable to what the human eye would judge. However, five regions is more appropriate for a honeycomb lattice because it has fewer neighbors than the triangular lattice. Changing the number of steps taken alters how local the quality assessment is. It is the local nature of the Geometric Defect Measure that makes it useful for detecting regions with many defects.

Secondly, the size of the regions used to determine whether a particle is in the correct position relative to the base particle p may be adjusted. It is appropriate to increase the size of the detection region the further the region is from the base particle. Doing so is compatible with judgements made by the human eye. It is also affected by the importance of having a correctly aligned lattice rather than a skewed lattice. For the honeycomb lattice, we have found that a good choice is allowing the radius of the detection region for region j to be $(j - 1) \times d_{nn1}/8$ for $j > 1$.

Last, as regions further away from the base particle are less important in quantifying the local geometry, the contribution of the more distant regions to the Geometric Defect Measure can be reduced. For our example, the first term in parentheses in Equation 4.2 then becomes $\sum_{j=1}^s m_{pij} \times (6 - j)/4$, for $j > 1$ where j denotes the number of the detection region.

4.2.3 Local Template Measure

The Local Template Measure compares the lattice particle positions to a small segment of the target lattice.⁷ For example, in the honeycomb lattice, a suitable segment of the

⁷Thanks to Jerry Marsden for the idea of the metric and to Philip du Toit for implementing it.

perfect lattice may be one honeycomb cell composed of six particles. For each particle in the given lattice, this template is pinned to the particle and rotated to find the best fit to the neighboring particles. The best fit minimizes the distance between points in the template and the nearest particle in the lattice. Once this minimum distance position of the template has been found, the Local Template Measure is given by⁸

$$\text{Local Template Measure} = \sum_p \left(\sqrt{\frac{1}{c} \sum_{i=1}^c (\mathbf{r}_i - \mathbf{r}_i^{\text{template}})^2} - \left(\frac{1}{c} \sum_{i=1}^c (\mathbf{r}_i - \mathbf{r}_i^{\text{template}}) \right) \right)^2 \quad (4.3)$$

where $\mathbf{r}_i^{\text{template}}$ is the position of the i^{th} point in the template, \mathbf{r}_i is the position of the lattice particle that is closest to this i^{th} template point, and c is the number of points in the template ($c = 6$ for the honeycomb cell template). The Local Template Measure for the lattice is found by summing these values for all particles, p , in the lattice.

Figure 4.7 shows a honeycomb lattice and the honeycomb cell template that is overlaid in calculating the Local Template Measure. To make the computation of the Local Template Measure faster, a feasible alternative is to use fewer cell templates than the number of particles in the lattice. In our experience, using a third as many templates as particles yields reliable results while cutting down the time taken to assess a lattice.

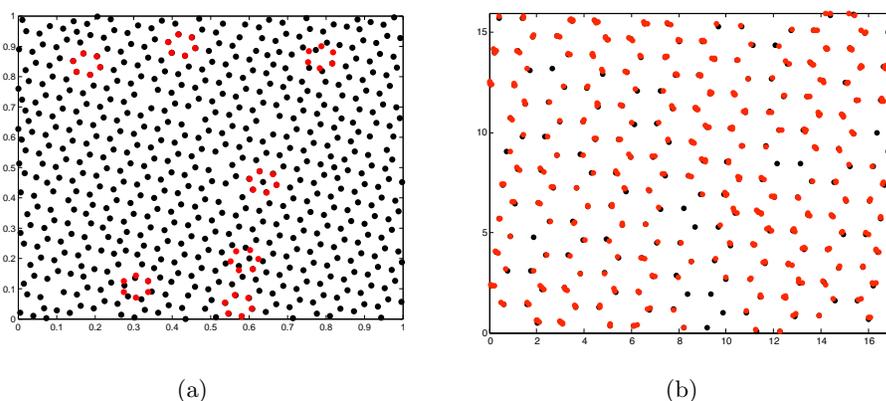


Figure 4.7: Illustration of the Local Template Measure. (a) A template of one cell of the target lattice (in red) is pinned to a particle and rotated to the position which minimizes the distance between points in the template and the nearest lattice point. (b) Using a third as many templates as the number of particles in the lattice provides good coverage of the lattice.

⁸This is the square root of the mean squared deviation.

The Local Template Measure suffers from the disadvantage that when minimizing the distance between template points and particles in the lattice, the same lattice particle may be attributed to more than one point in the template. For example, if a lattice particle is half way between two template points and there are no other lattice particles closer to the template points, then this one lattice particle will be used for both template points when calculating the distance between template points and lattice particles. This may lead to a region in a honeycomb lattice with an area equal to that of a unit cell that contains five lattice particles having a better value of the Local Template Measure than a similar region with six lattice particles.

Another disadvantage of the Local Template Measure is that it can not distinguish lattices where one lattice is a subset of the other. For example, the honeycomb lattice is a subset of the triangular lattice formed by adding a particle to the centre of each honeycomb cell. Since the templates used in the Local Template Measure do not consider particles in the centre of the template (unless close to one of the template points), this metric would judge a well-formed triangular lattice to be a well-formed honeycomb lattice. This is a serious drawback when using this metric since it requires independent analysis to determine whether the given lattice is the target honeycomb lattice or the competitor triangular lattice.

4.2.4 Voronoi Metric

The Voronoi Metric finds the Voronoi tessellation [86] of the particles in the lattice and compares the area of each Voronoi cell to the area of a Voronoi cell of the target lattice. The Voronoi tessellation of a set of particles in the plane is the partition of the plane into regions such that any point of the plane in the region corresponding to particle p is closer to p than to any other particle. The Voronoi tessellation of an imperfect lattice can be seen in Figure 4.8(a).

The Voronoi Metric is straightforward to apply to lattices with periodic boundary conditions that fill the available domain. For these lattices, the positions of the particles near the bounding box are mapped outside the bounding box on the opposite side in an appropriate way that is consistent with the periodic boundary conditions. This is illustrated in Figure 4.8(b). The particles of the lattice are shown in blue; the lattice has a bounding box specified by $0 < x < 1$ and $0 < y < 1$. The particles that have been mapped outside the bounding box are colored cyan. The particle on the left that is colored magenta lies inside

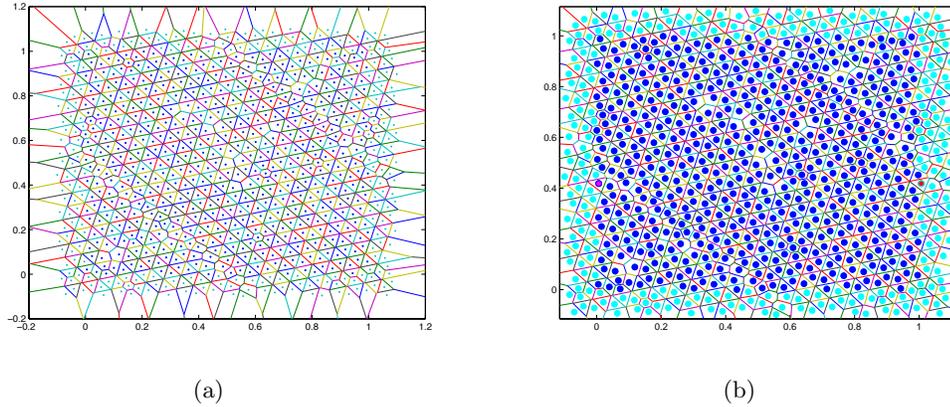


Figure 4.8: (a) Voronoi tessellation of a set of points in the plane. (b) The same Voronoi tessellation as in (a) with the particles of the original lattice colored blue and the particles that are mapped outside the periodic boundaries colored cyan. The magenta particle on the left is mapped outside the bounding box to the red particle position on the right.

the bounding box. It is mapped to the particle position on the right that is colored red. Such a construction allows the Voronoi cells of the particles close to the boundary to be calculated without edge effects, for periodic boundary conditions. Without this mapping of particles on the edges, these particles would have Voronoi cells that either extend to infinity or are partially determined by the bounding box rather than the other particles.

Algorithm for computing the Voronoi Metric.

1. Map particles that are close to the bounding box outside the bounding box, respecting the periodic boundary conditions.
2. Find a Voronoi tessellation of the particles.⁹
3. Compute the area of each Voronoi cell that contains a particle in the original lattice (not the particles that have been mapped outside the bounding box).
4. The Voronoi metric is given by

$$\text{Voronoi Metric} = \sum_p^N \left| \text{area}(V(p)) - \frac{\text{area of bounding box}}{N} \right| \quad (4.4)$$

⁹This can be done easily using the MATLAB[®] function *voronoi*. (MATLAB, 2007a, The MathWorks, Natick, MA.)

where N is the number of particles in the original lattice and $V(p)$ is the Voronoi cell containing particle p .

Equation 4.4 is motivated by the idea that for lattices that fill the domain completely (such as honeycomb lattices), each particle in a perfect lattice should take up the same amount of space.

An alternative expression for the area of a Voronoi cell in the perfect lattice may be obtained from the geometry of the target lattice and the best estimate for the lattice constant. The lattice constant may be found by Step 1 of the procedure for computing the Defect Measure described in Section 4.2.1.1. For a perfect honeycomb lattice (with periodic boundary conditions), the area of each Voronoi cell would be $3\sqrt{3}a^2/4$. The second term in Equation 4.4 may be replaced by this expression.

For honeycomb lattices formed in a bounding box that does not have periodic boundary conditions, or other types of lattices that do not completely fill the domain, an assessment of the quality of the lattice based on the Voronoi Metric can be made by ignoring the contribution from boundary particles. The boundary particles can be identified by using Boje's algorithm for finding the non-convex hull of a set of points (Section 4.2.1.3). In this case, the alternative expression for the area of a Voronoi cell in the target lattice should be used.

To include a contribution from the boundary particles of the lattice, there are a number of options. For lattices that completely fill the domain, modified Voronoi cells for the boundary particles can be formed by taking the area enclosed by the lines of the Voronoi tessellation and the bounding box. These areas for the boundary particles can then be compared to the areas of the similarly modified Voronoi cells that would be formed for the target lattice. Since the modified Voronoi cells of the boundary particles in the target lattice depend a great deal on the construction of the target lattice, it is best to sum the areas of the modified Voronoi boundary cells of the lattice and compare this to the sum of the areas of the modified boundary cells of the target lattice, rather than compare the areas of individual cells. However, this is not an effective way to assess the quality of a lattice because of the many possible ways of constructing the target lattice.

Another way to include the contribution of the boundary particles to a Voronoi-like metric is to define the Voronoi cells of the boundary particles to be the area enclosed by

the lines of the Voronoi tessellation and the non-convex hull of the particles. This method suffers from the same drawback as the previous one, but has the advantage that for a lattice that is not aligned with the edges of the bounding box, the contribution of the boundary particles to the value of the metric will be smaller. This effect is important for lattices that do not require the pressure from the domain walls to form, that is, lattices that do not completely fill the domain. For example, this method is the appropriate way to include the boundary particles in the Voronoi metric of a triangular lattice that is formed in a domain that is larger than the area occupied by the lattice.

None of these options for dealing with the boundary particles of the lattice are satisfactory. The Voronoi Metric, if one is unwilling to exclude boundary particles, is thus of limited use for lattices that do not completely fill the domain or are formed in domains without periodic boundary conditions.

4.2.5 Cumulative Distribution Function Metric

The cumulative distribution function of the inter-particle distances of a lattice can be used to assess the quality of the lattice. This metric was proposed by Mezić and Runolfsson in a different setting [66]. The cumulative distribution function (CDF) metric is most effective, for our purposes, when only inter-particle distances up to a distance of slightly above $2a$ are considered. In a perfect lattice, this includes the nearest neighbors, the next nearest neighbors and the third nearest neighbors of each particle.

Algorithm for computing the Cumulative Distribution Function Metric.

1. Let d_{max} be the distance to a point half-way between the third nearest neighbors circle and fourth nearest neighbors circle of particles in the target lattice.
2. For each distance d in $0 < d \leq d_{max}$, find all inter-particle distances of the lattice being assessed that are less than or equal to d .
3. Then

$$\text{CDF}_{\text{lattice}}(d) = \sum_{d_{ij} \leq d, i > j} d_{ij} \quad (4.5)$$

for $0 < d < d_{max}$, where the d_{ij} are the inter-particle distances of the lattice.

4. CDF_{target} is defined in a similar way using the inter-particle distances of the target lattice.
5. The Cumulative Distribution Function Metric is given by

$$\int_0^{d_{max}} |CDF_{\text{lattice}}(l) - CDF_{\text{target}}(l)| dl. \quad (4.6)$$

The CDF of a lattice and its target lattice is shown in Figure 4.9. CDF_{lat} is in blue and CDF_{target} is in green. The value of the Cumulative Distribution Function Metric is the area between the curves.

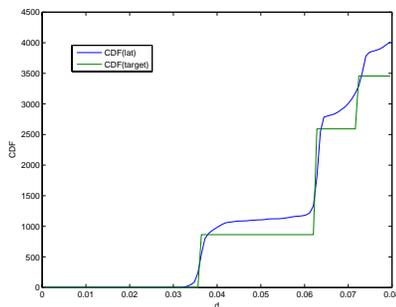


Figure 4.9: The Cumulative Distribution Function Metric computes the area between CDF_{lattice} shown in blue and CDF_{target} shown in green.

The question of how to treat the boundary particles of a lattice is also non-trivial for the CDF metric, especially for domains without periodic boundary conditions. For lattices formed in domains with periodic boundary conditions, the only concern is whether the domain can indeed be filled with a perfect lattice with the specified number of particles. Not all domains have a perfect lattice that completely fills the domain for an arbitrary number of particles; however, this introduces only a very minor error into the value of the CDF metric.

For domains without periodic boundary conditions, a target lattice must be constructed that can be used to find CDF_{target} . The inter-particle distances of the target lattice depend on where the boundary particles are placed, that is, the shape of the boundary.

A more serious problem with the CDF metric is that there is a cancellation between two different types of defects: missing particles and extra particles. If a particle is missing a neighboring particle at a distance d' from it and another particle has an extra particle

at a distance d' from it, then there will be some cancellation between these two defects, and the value of the CDF metric at d' will be lower than it should be. However, these two defects also affect the surrounding particles and this will add to the value of the CDF metric. How much is added depends on the arrangement of particles around the defect and not the defects themselves.

The CDF metric was designed to only look at inter-particle distances less than d_{max} partially for this reason. There is less opportunity for such cancellation of defects to occur. Another reason for limiting the CDF metric's horizon to d_{max} is that considering all inter-particle distances would put too much emphasis on the long range order of the lattice. When assessing the quality of a lattice, the human eye tends to focus more on the order within regions that have a radius of a few lattice constants, rather than the long range order of the lattice as a whole.

A feature of the CDF metric is that it tends to judge lattices with grain boundaries relatively harshly. This may or may not be a concern depending on the goal and how severe such a defect is considered to be.

4.3 Comparison of Quality Metrics

It is natural to ask which of the metrics for assessing the quality of a lattice discussed in Section 4.2 is the best.

Which quality metric performs best depends on which properties of a lattice are more important. Different metrics focus on different aspects, such as having the correct number of particles in approximately the right positions, or having the correct alignment of particles. This will be discussed further below.

There are two straightforward ways to compare metrics that assess the quality of a lattice: whether it can identify the best lattice from a set of lattices, and the computational time taken to compute the value of the metric for a lattice.

In order to compare the metrics from Section 4.2, a set of 20 lattices were generated, with 576 particles in a domain with periodic boundary conditions. The target lattice was the honeycomb lattice. Each metric was used to assess the lattices and rank them from best to worst. The time taken to calculate the value of the metric for each lattice was averaged over the 20 lattices. The results are shown in Figure 4.10 and Table 4.3. We chose

to examine 20 lattices because this allowed us to have a variety of different lattices that exhibited different features. The number of particles was chosen to be large enough such that boundary particles would not significantly contribute to the value of the metric, and because this number of particles can fill the domain with a honeycomb lattice.

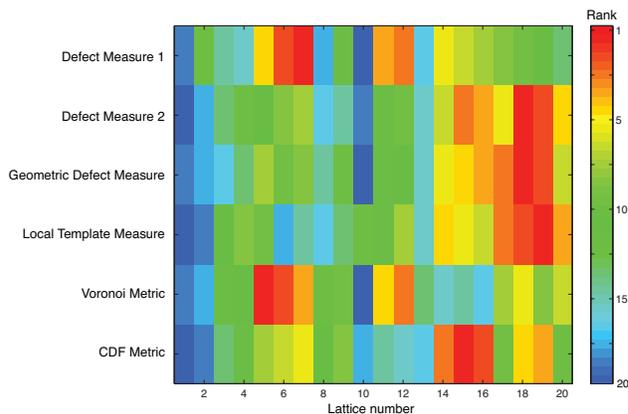


Figure 4.10: Comparison of quality assessment metrics. Each metric on the left hand side ranked the 20 test lattices (along horizontal axis) from best to worst. The colorbar indicates which color corresponds to which ranking. Red signifies the best lattice (rank 1) and dark blue the worst lattice (rank 20). Metrics that focus on similar features have a similar pattern of colors across their rows.

Figure 4.10 has the metrics discussed in Section 4.2 along the vertical axis and the index of each of the 20 lattices used for the comparison along the horizontal axis. Defect Measure 1 and Defect Measure 2 differ only in the weights assigned to the different types of defects.¹⁰ The colors in the figure represent the ranking of the 20 lattices, with a rank of 1 in red being the best lattice and a rank of 20 in dark blue being the worst lattice. For example, along the top row it can be seen that Defect Measure 1 ranked lattice number 7 as the best lattice and lattice number 10 as the worst lattice. Metrics that focus on similar features have a similar pattern of colors across their rows.

The index of each lattice is not important; however, groups of lattices with similar lattice numbers were generated by the same inter-particle potential at the same density. Two different types of inter-particle potentials were used: lattices 1–13 were generated using a Rechtsman-style potential and lattices 14–20 were generated using a polynomial potential.

¹⁰Weights for Defect Measure 1: $\omega_{\text{displaced}} = 1.0$, $\omega_{\text{missing}} = 1.0$, $\omega_{\text{extra}} = 0.8$.
Weights for Defect Measure 2: $\omega_{\text{displaced}} = 1.0$, $\omega_{\text{missing}} = 0.02$, $\omega_{\text{extra}} = 0.015$.
There were no boundary or lone particles in these lattices. These weights were chosen heuristically.

These potentials will be explained in Section 5.1. Beyond this, the lattices may be grouped according to $\{1, 2\}$, $\{3\}$, $\{4, 5\}$, $\{6-9\}$, $\{10\}$, $\{11\}$, $\{12\}$, $\{13\}$, $\{14-16\}$, $\{17-19\}$, $\{20\}$. These groupings are due to the parameters used in the potentials and the initial density of particles (see Appendix A).

	Time to compute (s)
Defect Measure	0.117
Geometric Defect Measure	6.69
Local Template Measure	44.4
Voronoi Metric	0.338
CDF Metric	0.163

Table 4.2: Time taken to compute quality metrics.

Figure 4.10 shows that Defect Measure 1 ranks the test lattices in an order that is similar to the ranking of the Voronoi Metric. Rows 2 and 3 of Figure 4.10 show that Defect Measure 2 assigns similar rankings to the test lattices as the Geometric Defect Measure. This highlights the flexibility of the Defect Measure, which results from the freedom to choose the weights for the different types of defects. It will be shown below that the Geometric Defect Measure and the Voronoi Metric consider different aspects of lattices to be important and thus apply to different situations. The Local Template Measure ranks lattices in a similar way to the Geometric Defect Measure and Defect Measure 2. One notable exception is that the Local Template Measure ranks lattice number 10 as being average whereas this lattice is ranked as one of the worst by all of the other metrics. The CDF Metric identifies the same few lattices as being the worst lattices that the other metrics identify. However, the lattices that are judged to be the best lattices by the CDF Metric are not judged to be that way by the other metrics. The Geometric Defect Measure ranks the CDF Metric's best lattices as being only moderately good.

Lattices number 18 and 5 are shown in Figure 4.11. It can be seen that in Lattice 18, although there are defects like missing and extra particles, the particles tend to be aligned with each other. For a majority of particles, the angles between nearest neighbor particles are close to that of a perfect honeycomb lattice. There is also more medium range structure than in Lattice 5. This is precisely what the Geometric Defect Measure focuses on. However, in Lattice 18, there are obvious defects. In Lattice 5 on the other hand,

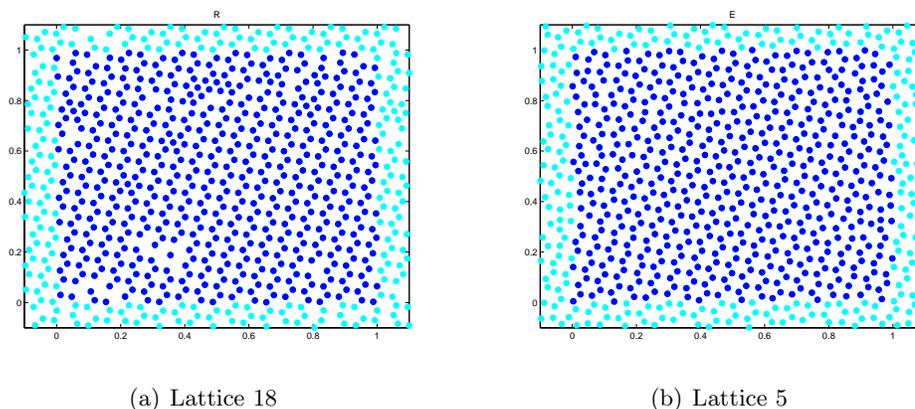


Figure 4.11: Best lattices. Test honeycomb lattices with 576 particles and periodic boundary conditions. The dark blue particles are inside the bounding box. The cyan particle positions show the structure at the edge of the bounding box. The lattice numbers correspond to those in Figure 4.10.

the density of particles is much more uniform across the domain. The particles are not aligned well into a honeycomb lattice but a majority of particles have the correct number of nearest neighbors and form rough rings of six particles. It is this focus on the local area of each particle that is characteristic of the Voronoi Metric. The shape of the Voronoi cells is not considered, only their area compared to a Voronoi cell of a perfect lattice. Thus local density is the most important feature.

Which of Lattices 18 and 5 is judged to be the better lattice depends on the goal. If the purpose of having a honeycomb lattice is to cover an area evenly, with each particle having three nearest neighbors and forming rings of six particles, then the Voronoi Metric is the one to use. If the goal is to form as much of a close-to-perfect honeycomb lattice as possible then the Geometric Defect Measure should be used.¹¹ It is interesting to note that the Defect Measure can be used to achieve both of these goals simply by adjusting the weights for the different types of defects.

The Voronoi Metric whose results are shown in Figure 4.10, used Equation 4.4. This metric can also be implemented using the alternative expression for the area of a Voronoi cell in the perfect lattice explained in Section 4.2.4. The alternative expression computes the average lattice constant and then sets the perfect Voronoi cell area to be $3\sqrt{3}a^2/4$,

¹¹Lattice 18 was generated using a polynomial potential while Lattice 5 was generated using a Rechtsman-style potential [75].

based on geometry. The results for this metric are shown in Figure 4.12 under the label Voronoi Metric 2. The assessment of this quality metric depends greatly on how accurate the estimate of the lattice constant is. The average lattice constant is calculated using the histogram method in Step 1.3(b) of the algorithm for computing the Defect Measure.

Voronoi Metric 2 ranked Lattice 17 so highly because of the peculiar coincidence of the error in the area of the Voronoi cells around the gross defects, summing to a similar error in area that is spread out across all Voronoi cells. Lattice 17 looked similar to Lattice 18, which is shown in Figure 4.11(a). This feature of the error can be seen by comparing the lattices in Figure 4.11. It is because of a few large errors in area that sum to a similar total as the sum of many small errors in area, and not because the Voronoi Metric particularly looks at geometry, that Lattice 17 has the best ranking. It can be seen from Figure 4.12 that Voronoi Metric 1 and Defect Measure 1 avoid this anomaly.

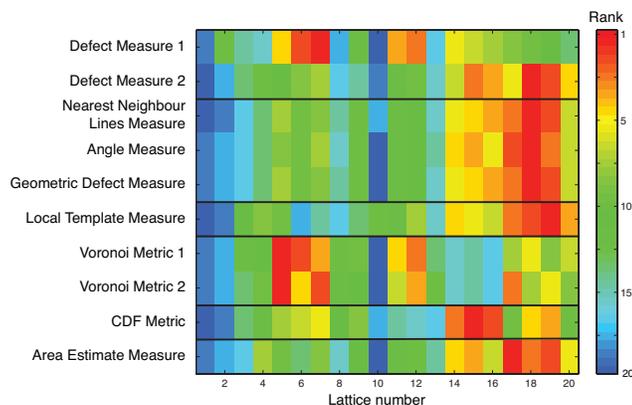


Figure 4.12: Comparison of quality assessment metrics. Each metric on the left hand side ranked the 20 test lattices (along horizontal axis). The colorbar indicates which color corresponds to which ranking. Red signifies the best lattice (rank 1) and dark blue the worst lattice (rank 20).

Figure 4.12 also shows the two components of the Geometric Defect Measure: one that determines whether particles are where they should be along the nearest neighbor lines and one that evaluates how close the angles between nearest neighbors are to what they should be in a perfect lattice. The rankings for these two components are very similar. This implies that either component can be used alone as a metric instead of the more complete Geometric Defect Measure.

The last row of Figure 4.12, labeled Area Estimate Measure, is a very rough quality

assessment metric. It compares the area that a perfect lattice would cover if it had the average lattice constant, to the area of the bounding box. That is,

$$\text{Area Estimate Measure} = \left| \text{Area of bounding box} - (3\sqrt{3} a^2/4) \times (\text{number of particles}) \right|.$$

This extremely fast calculation can roughly rank the lattices in a way similar to the Geometric Defect Measure, if the value of a that is used is accurate. Its advantage is that it requires nothing more than the calculation of the average lattice constant. All of the other metrics require this computation and additional computations.

Table 4.3 shows the time taken to compute the value of each metric.¹² These times are averages over the time taken to compute the metric for the 20 test lattices. Each metric was given the minimum information it needed in order to compute the value of the metric. The Defect Measure calculations were given the distances to the nearest twenty neighbors of each particle; the CDF Metric was given a list of all nearest neighbor distances up to a distance of $2.2a$; the Geometric Defect Measure, the Voronoi Metric, and the Local Template Measure were given the particle positions.

The rationale for giving each quality metric only the minimum information that it needs stems from the applications of the metrics. If the Defect Measure is used to assess the lattice formed by vehicles flying in formation then these vehicles need only detect their nearest twenty neighbors—this limits the range necessary for their relative distance sensors and the amount of information that must be transmitted. Further, when using the Defect Measure to assess the quality of lattices formed in a LAMMPS¹³ simulation, as was done in the optimization of self-assembly potentials discussed in Section 5.1.3, the distance to the nearest twenty neighbors is easily accessible due to the structure of the LAMMPS simulation code. This code calculates particle positions in parallel by dividing up the domain into smaller regions, thus keeping more detailed information about a particle's nearest neighbors. Similar reasoning holds for the CDF Metric.

We see that the Defect Measure is the fastest to compute. The Local Template Measure requires the most time to compute by far. Since the Local Template Measure, Geometric Defect Measure and Defect Measure 2 rank lattices in a similar manner, it is most efficient

¹²Metrics were computed using MATLAB[®] (2007a, The MathWorks, Natick, MA).

¹³LAMMPS is a software package that was developed by Sandia Laboratories for the simulation of molecular systems. See <http://lammps.sandia.gov> [71].

to use Defect Measure 2 whenever a metric focusing on particle alignment is needed.

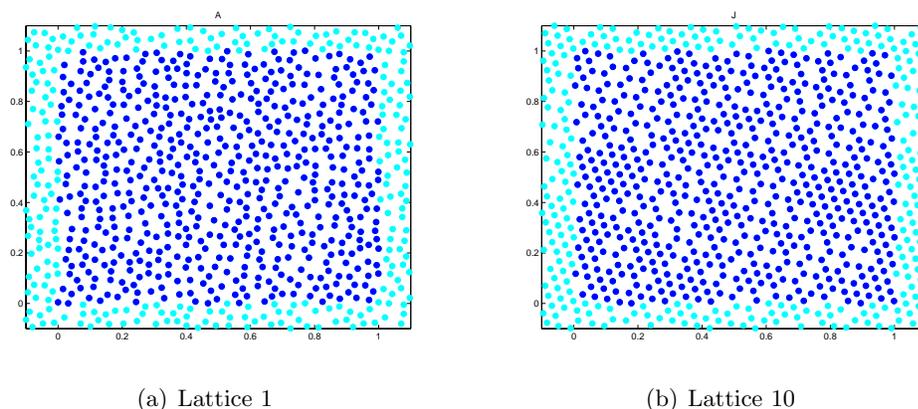


Figure 4.13: The two worst test honeycomb lattices with 576 particles and periodic boundary conditions. The dark blue particles are inside the bounding box. The cyan particle positions show the structure at the edge of the bounding box. The lattice numbers correspond to those in Figure 4.10.

Figure 4.10 shows that all of the metrics, apart from the Local Template Measure and the CDF Metric, found Lattices 1 and 10 to be the two worst lattices. These lattices are shown in Figure 4.13. Lattice 1 was judged to be bad by all of the metrics. Lattice 10 was ranked third to last by the CDF metric; however, the Local Template Measure judged it to be an average-quality lattice. It can be seen that these two generally worst lattices also exhibit the two different types of lattice that the best lattices in Figure 4.11 did. Namely, one has particles that are well aligned (albeit in the wrong locations) and the other has a more uniform number of particles per area (though not aligned into the honeycomb pattern at all). Both are considered to be bad lattices because of the large number of defects relative to the target honeycomb lattice.

The Local Template Measure assigned Lattice 10 an average ranking because it forms a reasonably good triangular lattice. The honeycomb lattice can be considered to be a subset of the triangular lattice: a triangular lattice can be formed from a honeycomb lattice by putting a particle in the centre of each honeycomb cell. Since the templates used in the Local Template Measure do not care whether or not there is a particle in the centre of the template, a well-formed triangular lattice would rank highly using a honeycomb Local Template Measure.

The CDF Metric assigns rankings that are quite different from all of the other metrics. It

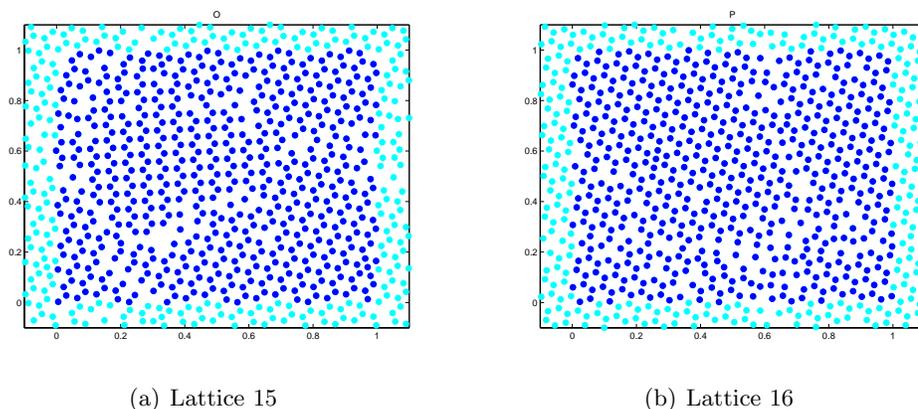


Figure 4.14: The CDF Metric’s two best test honeycomb lattices with 576 particles and periodic boundary conditions. The dark blue particles are inside the bounding box. The cyan particle positions show the structure at the edge of the bounding box. The lattice numbers correspond to those in Figure 4.10.

judges the worst lattices, similarly to the other metrics but chooses different lattices as the best lattices. These best lattices, shown in Figure 4.14, received a moderately good ranking from the Geometric Defect Measure but were rated as quite bad by the Voronoi Metric. One feature that they exhibit is having regions that are well-formed that are separated from other well-formed regions by areas with many defects. Lattices 15 and 16 have large regions with missing particles and also curves with a small distance between the particles (more evident in Lattice 16). This combination leads to some cancellation, causing the lattice to be ranked highly, as discussed in Section 4.2.5.

Thus, it seems that the CDF Metric does not assign rankings that are similar to what a human observer would assign, whether particle alignment or density is the focus. The CDF Metric is an indicator of how many inter-particle distances (within a limited range set by d_{max}) are correct. Since the distances are the focus rather than the local density or alignment (properties that are important for forming a lattice structure), the structure is less important with this metric. This renders the CDF Metric less useful as a quality metric for the particular problem of the self-assembly of particles into a target lattice.

4.4 Conclusion

We have introduced five metrics for assessing the quality of lattices in a two-dimensional domain. These metrics are based, to differing degrees, on a local perspective that quantifies lattice quality by analyzing the region immediately surrounding each particle in the lattice. This emphasis on local order leads to quality judgements that are much the same as those of the human eye. It differs from the global perspective that is prevalent in the condensed matter and materials science literature, since it does not investigate the number and type of defects by looking at bulk properties of the material.

Which of the lattice quality assessment tools introduced is better depends on which aspect of a lattice is most important to the assessor. A comparison of the metrics reveals that they fall into three main categories; those that focus on uniform local density, those that focus on local alignment, and those that focus on the correct distance between particles. The advantage of the Defect Measure, apart from being the fastest to compute, is that the weights assigned to each type of defect can be adjusted to reflect any one of these categories.

The Defect Measure and other metrics provide a succinct description of the order within a planar system of particles. They are useful in a variety of applications, particularly in tracking transitions in molecular systems from disordered to ordered states. This feature is exploited in the next chapter for the generation and assessment of potentials leading to the self-assembly of particles, the construction of a phase diagram, and identifying clusters of well-formed particles.

Chapter 5

Applications of the Defect Measure and Other Lattice Quality Assessment Tools

The Defect Measure is a tool that can be used to quantitatively assess the quality of lattices. As such, it can be regarded as a low-dimensional representation of a system that transitions between disordered and ordered states. It is useful in a variety of situations, such as finding potentials that lead to the self-assembly of particles, assessing the robustness of such potentials, and detecting phase transitions from a disordered to an ordered state.

Having a simple way to characterize the quality of a lattice opens up many avenues in the study of self-assembling particles. The quality assessment tools of the previous chapter enable potentials for self-assembly to be optimized directly for the quality of the resultant lattice. The quality metrics can also quantify the robustness of these potentials. We present a robust method for generating optimal self-assembly potentials that are themselves robust to changes in parameters of the potential and initial conditions. This method relies on (a) the quality metrics, and (b) a computationally efficient trend optimization scheme that uses the quality metrics as the objective function.

The Defect Measure is a good coarse description of the phase transition from liquid to solid of krypton atoms on a graphite substrate. We use the Defect Measure to construct the phase diagram for this system. Our results agree very well with those of experiments.

Clusters of well-formed particles in a lattice can be identified by the Defect Measure, since it assigns a quality metric value to each particle. The Voronoi metric is particularly useful in identifying global defects by considering local properties of lattices. It is a general feature of the quality metrics developed in the previous chapter that by examining the local

order around each particle, much can be said about the global pattern.

5.1 Generation of Potentials for the Self-Assembly of Particles

The self-assembly of constituent particles into an ordered final configuration is an active area of research from materials science and chemistry to nanotechnology. Through local interactions, the self-assembling entities organise themselves into a more regular configuration than the state in which they started.

Self-assembly has been studied experimentally in a variety of biological and nanotechnology systems. Examples include the formation of lipid bilayers, block copolymers organizing into ordered arrays, and the self-assembly of microspheres (see [75] and the references therein). Theoretical studies, such as that of Jagla [45], have typically started with a given interaction potential between particles and then examined the resulting structure and properties of the system. An exception to this is the reverse Monte Carlo method of Lyubartsev and Laaksonen [56] (and extensions by Toth [83]), which uses radial distribution functions to iteratively find the interaction potential. However, applications have only considered liquids and solutions (see also [55] for applications using experimental data). Early work in this area was carried out by Salsburg et al. [77] who derived an analytic expression for the molecular pair-distribution function for a system of particles interacting via an arbitrary nearest neighbor potential.

The focus of this section is on the generation of isotropic pairwise potentials that lead to the self-assembly of particles into a lattice configuration in the plane. We first describe the work done by Rechtsman and co-workers [75]. This was the motivation for the development of the quality metrics in Chapter 4 and the trend optimization scheme presented later in this section. Next we explain how the quality metrics can be used to design self-assembly potentials. A more efficient method for finding these potentials is to run an optimization procedure over the parameters in the potential, rather than do a brute force search. Section 5.1.3 presents two ways in which the Defect Measure can reduce the time taken to evaluate the objective function in an optimization. In Section 5.1.4 we present an improved optimization procedure that exploits the simple trend of the very noisy objective function. This scheme finds parameters of the self-assembly potential that not only lead to higher

quality lattices, but do so two orders of magnitude faster than a simulated annealing scheme. The robustness of such potentials is also an important issue. Ways of quantifying robustness are discussed in Section 5.1.5.

5.1.1 Background

Rechtsman and co-workers [75, 74] developed two computational algorithms to find potentials that lead to the self-assembly of particles into given target configurations. In contrast to previous work (both experimental and theoretical), their method was an inverse method in that the two-dimensional target lattice configuration was specified and an appropriate radially symmetric pair interaction potential was found. Previous work concentrated on determining the structure that would result from particles with a specified non-covalent interaction between them. Rechtsman et al. considered both energetic and mechanical stability by investigating the lattice sums and phonon spectra [50] of their potentials. Their [80] more theoretical work analyzed the asymptotic behaviour of the ground state energy of particle systems governed by Lennard-Jones-like pairwise isotropic interaction potentials.

The first scheme developed by Rechtsman et al. [75], a simulated annealing optimization, focuses on energetic stability by maximising the energy gap between the target lattice and its competitor lattices. This is done while maintaining mechanical stability via the requirement that phonon frequencies be real. Competitor lattices are lattice configurations that the self-assembling particles may tend towards or end up in that are not the target lattice. For example, when assembling a honeycomb lattice, the triangular lattice, which has twice as many nearest neighbors for a given particle as the honeycomb lattice, is the competitor lattice.

The second optimization scheme considered by Rechtsman et al. concentrates on maximising the stability of the lattice near its melting point, while requiring stable lattice sums and real phonon frequencies. Molecular dynamics simulations are run at temperatures close to, but below, the melting point of the target lattice and a measure of the extent of phase transitions, called the Lindemann parameter, is calculated. The Lindemann parameter is defined as

$$\sqrt{\frac{1}{N} \sum_i (\mathbf{r}_i - \mathbf{r}_i^{(0)})^2 - \left(\frac{1}{N} \sum_i (\mathbf{r}_i - \mathbf{r}_i^{(0)}) \right)^2}, \quad (5.1)$$

where $\mathbf{r}_i^{(0)}$ is the initial position of particle i , \mathbf{r}_i is the position some time later, and N is the number of particles.¹ It is the objective function for a simulated annealing optimization. Note that the Lindemann parameter compares each particle's position to the initial position of that particle, that is, the position of that particle when it is in the target configuration initially (particles start the simulation in the target configuration, the lattice is heated close to the melting temperature for the lattice, and then cooled). In this way, it requires information about the history of each particle and for each particle to be assigned to a specific position in the target lattice. Rather than measuring the general quality of the final lattice, it is a measure of how much a target lattice configuration has melted.

To test the potentials that they find from their optimization schemes, Rechtsman et al. run an NVT² Monte Carlo simulation, starting from a random initial configuration, and see whether the particles self-assemble into a lattice. Whether or not a lattice is formed is determined by looking at the lattice itself, deciding if the given configuration is ordered enough to be considered a lattice, and calculating the structure factor of the lattice, which gives an indication of the long-range order.

We improve on this eyeball test of the lattice by assigning a value to the quality of the lattice using the lattice quality assessment tools we have developed in the previous chapter. In this way, lattices can be compared to each other in terms of quality, and the extent to which the lattice differs from the target lattice configuration can be quantified.

In comparing potentials that lead to self-assembly, there are three criteria according to which potentials may be contrasted: the quality of the resulting lattices, the time taken to discover the potential, and how robustly good quality lattices are formed. The following sections will address these issues.

5.1.1.1 Baseline Geometric Method

With regard to the issue of the time taken to discover a potential that leads to self-assembly, there is a baseline to which the potentials found by the method of Rechtsman et al. or any other method, including the one in Section 5.1.4, may be compared. This is the method of

¹Compare with Equation 4.3. The Local Template Measure of Section 4.2.3 was developed in order to more directly compare our results with those of Rechtsman et al. [75] since Equation 5.1 is not a direct measure of the quality of a lattice.

²N = number of particles, V = volume of the system, T = absolute temperature. These quantities are held constant.

designing a potential by hand.

This baseline geometric approach is as follows: for a particle in the target lattice (not near the boundary), find the distances at which it has neighboring particles and count the number of particles at each of these distances from the initial particle. For example, in the honeycomb lattice, a particle has 3 nearest neighbors at a distance of a , where a is the minimum lattice spacing, called the *lattice constant*. The same particle has 6 neighbors at a distance of $\sqrt{3}a$ and 3 at a distance of $2a$. Particles must have a strongly repulsive core and interactions should be local so that the potential goes to zero beyond a distance of approximately $2a$ (this distance may be chosen to be any value; however, this is the distance we consider here to define local interactions³). Given this information, the shape of the potential can be deduced—there must be an inner repulsive core, a minimum at an inter-particle distance of a , another minimum at a distance of $\sqrt{3}a$ that is deeper (because there are more neighbors at this distance), and beyond this the potential goes to zero. The relative heights of the minima and the local maxima may be chosen to favor the honeycomb lattice and disfavor the competitor triangular lattice. An example of such a potential is shown in Figure 5.1. The potential is defined piecewise using cubic polynomials. The cubic splines are constructed so as to have zero derivative at the nodes thus ensuring that the potential is continuously differentiable on the whole positive real line. In order to ensure a repulsive core, a 6–12 Lennard-Jones potential is used on the first segment.⁴

Any potential that is found by any other method must perform better than this polynomial potential in tests of lattice quality and robustness to warrant the time taken to implement the optimization procedure used to find the potential.

5.1.2 Design of Potentials for Self-Assembly

The Defect Measure can be used as a tool not only to assess the quality of lattices that result from simulations with a particular isotropic potential between particles, but also to refine such potentials to achieve the best potential. A functional form for these potentials can be designed with a number of parameters that may be varied to achieve the potential that yields the best lattices, as measured by the Defect Measure.

³Polynomial potentials that were equal to zero beyond $5a$ (and not before) were also considered, but lattices formed with a tail of the potential that starts at $2a$ were better.

⁴The form of a Lennard-Jones potential can be seen in Equation 3.1.

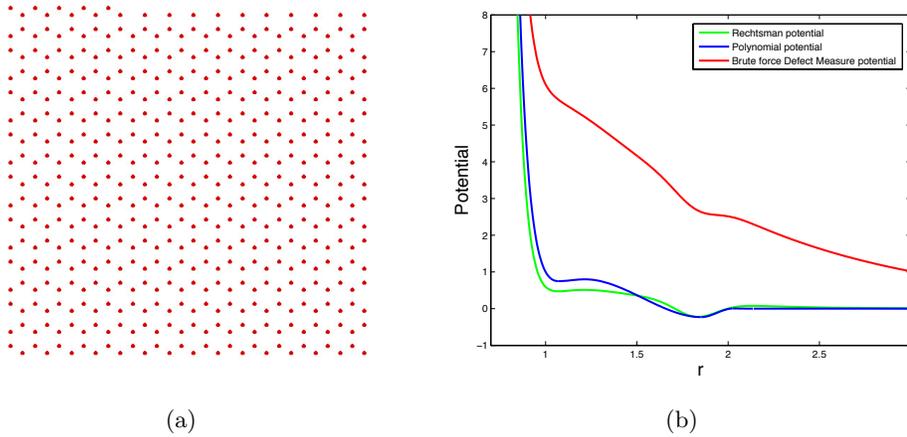


Figure 5.1: (a) A perfect honeycomb lattice. (b) Rechtsman et al.’s honeycomb potential (Equation 5.2), a polynomial potential designed using the baseline geometric method, and the potential found by a brute force search over parameters using the Defect Measure to assess lattice quality (Section 5.1.2).

Rechtsman et al. [75] found the following expression for the self-assembly of particles into a honeycomb lattice:

$$V_{HC} = \frac{5}{r^{12}} - \frac{a_0}{r^{10}} + a_1 e^{-a_2 r} - 0.4 e^{-40(r-a_3)^2} \quad (5.2)$$

with $a_0 = 5.89$, $a_1 = 17.9$, $a_2 = 2.49$, and $a_3 = 1.823$ being the best parameter values that they found for good lattices. This potential is shown in Figure 5.1. The functional form of this potential was chosen to be physically realizable in an experimental setting. We note that in [75], the quality of lattices was not quantitatively assessed, rather, particles were judged to have self-assembled into lattices or not by human observation.

Using the same basis as Rechtsman and co-workers, we can design a better potential by doing a brute force search over the different parameter values, and assessing the quality of lattices formed with these values. Any of the quality metrics introduced in Section 4.2 may be used for this. Although a brute force search over all four parameters has been conducted, and a global minimum found, we first illustrate the method with a search over only one parameter value, a_2 , for clarity. The other parameter values and conditions (such as the density and cooling schedule) were kept constant. One thousand simulations were run with 65 particles in a domain without periodic boundary conditions. The values of

the parameter a_2 used for the 1000 simulations formed a Gaussian distribution with mean $a_2 = 1.49$ and standard deviation of 0.6. Figure 5.2 shows the Defect Measure of the final lattice of each simulation versus the value of the a_2 parameter used to generate the lattice. The values of the weights for the Defect Measure (Section 4.2.1) are: $\omega_{\text{displaced}} = 1.0$, $\omega_{\text{missing}} = 1.0$, $\omega_{\text{extra}} = 0.8$, $\omega_{\text{boundary}} = 0.2$, and $\omega_{\text{lone}} = 2.0$. These weights are the same weights as for Defect Measure 1 in Section 4.3. The lattices corresponding to the red and green points are shown in Figure 5.3. The green point has the lowest Defect Measure of all the lattices. Note that each point in Figure 5.2 is the result of one simulation with random initial velocities for the particles.

Once the region in which the Defect Measure has a minimum is found, this procedure may be repeated. The brute force search over all parameter values reveals that the Defect Measure function over the space of all parameters has many local minima and that is it very noisy. Furthermore, different evaluations with the same parameters in the potential can lead to lattices with quite different quality, due to sensitivity to initial conditions.

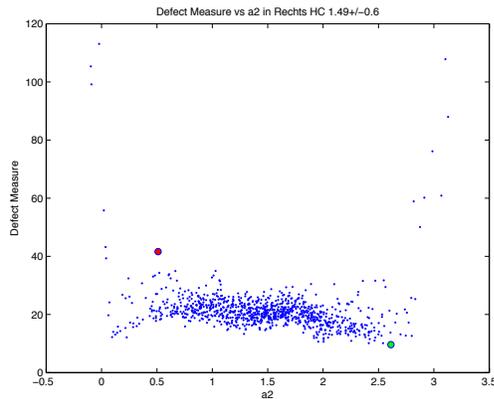


Figure 5.2: Defect Measure versus parameter a_2 in Equation 5.2. The 1000 simulations shown here used 65 particles in a domain without periodic boundary conditions. The a_2 values have a mean of 1.49 and a standard deviation of 0.6. The lattices corresponding to the red and green points are shown in Figure 5.3.

Figure 5.2 indicates that for the conditions under which the simulations were run, the best choice for parameter a_2 is a value of 2.6. Note that the honeycomb lattice is quite fragile and particularly sensitive to the density of particles and the boundary conditions. Thus, for different densities or boundary conditions, other values of a_2 may be more appropriate. However, this example illustrates the method for designing an isotropic potential by brute

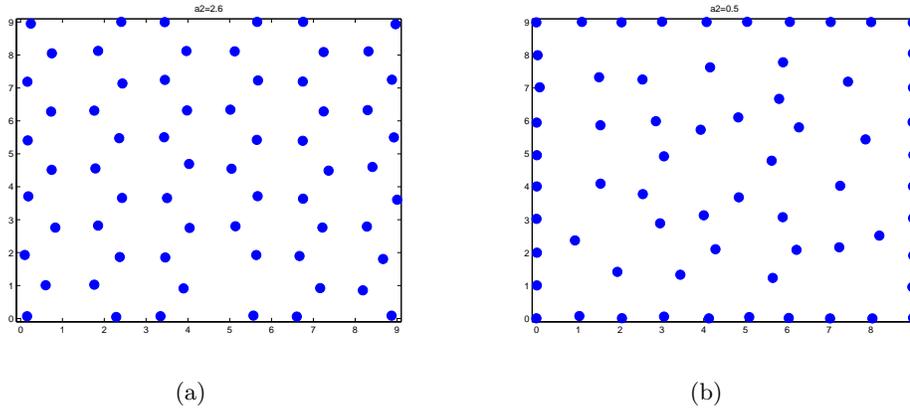


Figure 5.3: (a) Lattice corresponding to the green point in Figure 5.2, with $a_2 = 2.6$ and Defect Measure = 9.6. (b) Lattice corresponding to the red point in Figure 5.2, with $a_2 = 0.5$ and Defect Measure = 41.6.

force.

As mentioned in Section 4.3, the weights of the Defect Measure should be chosen according to the properties of the lattice that are most important for the particular application. Defect Measure 1, used here, focuses on the correct local density of particles. The Geometric Defect Measure or other metrics can also be used to design the potential for the formation of lattices.

Figure 5.4(a) shows the Geometric Defect Measure versus the parameter a_2 for the same simulations as above. The best lattice, that is, the one with the lowest Geometric Defect Measure value, shown in Figure 5.4(b), was formed with $a_2 = 2.23$. This value of a_2 is quite different from that found by Defect Measure 1. Thus, which value of a_2 is chosen depends on whether the focus is the correct alignment or local density of particles. Figure 5.5 depicts the two components of the Geometric Defect Measure: the Nearest Neighbor Lines Measure and the Angle Measure. The former looks at whether particles are aligned with their nearest neighbors and the latter measures how close the angle between nearest neighbors is to what it should be. Both components favour an a_2 value of 2.23.

Note that the a_2 values that are found to give the highest quality lattices (by whichever metric) are applicable only to self-assembly problems that have the same density and cooling schedule. The number of particles and the boundary conditions of the domain are less of an influence. At other densities and with other cooling schedules, different values of a_2 may yield better lattices. In general, the slower the rate of cooling, the better the resulting

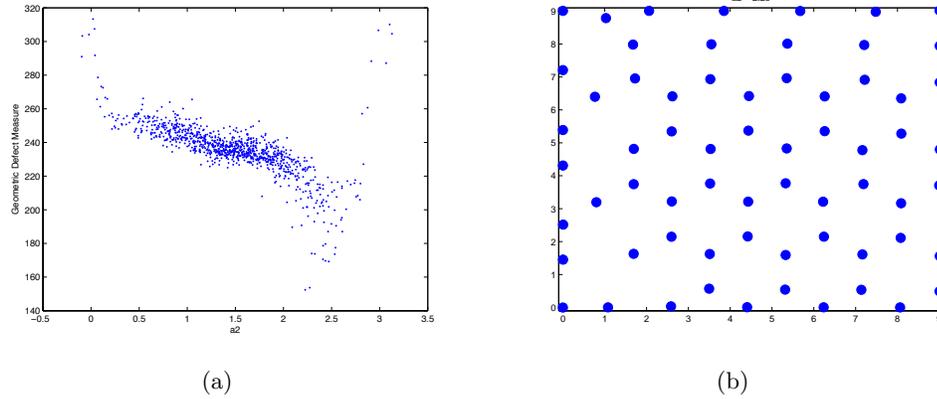


Figure 5.4: (a) Geometric Defect Measure versus parameter a_2 in Equation 5.2. (b) Lattice corresponding to the lowest Geometric Defect Measure in (a), with $a_2 = 2.23$.

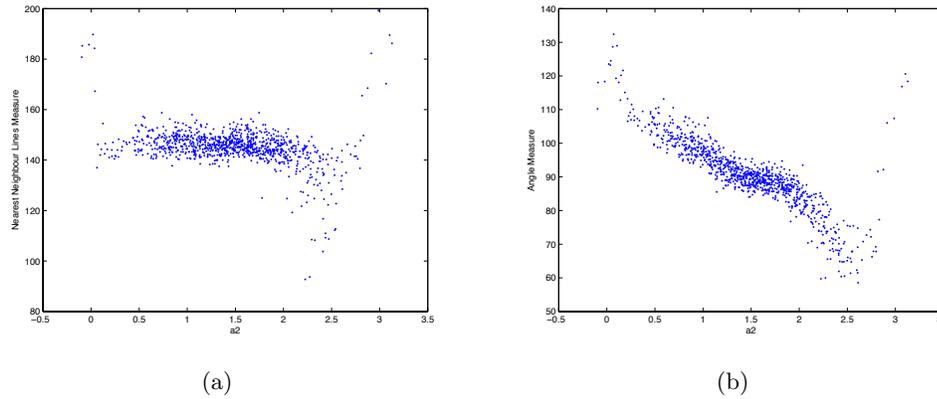


Figure 5.5: The two components of the Geometric Defect Measure. (a) Nearest Neighbor Lines Measure versus parameter a_2 . (b) Angle Measure versus parameter a_2 .

lattice is. This brute force method may be applied to the density of particles as well, and the best parameters for the potential found at each density of interest.

The brute force search over all of the parameters in the potential can be done to find the best combination of parameters, that is, the potential that leads to the highest quality lattice. Figure 5.6(a) shows the results of a search over parameters a_0, a_1, a_2 in Equation 5.2 using the Defect Measure as the quality metric⁵. Over 10^5 points were tested in the range

$$4.00 < a_0 < 7.50, \quad 15.0 < a_1 < 21.0, \quad 0.0001 < a_2 < 3.50. \quad (5.3)$$

⁵Thanks to Philip du Toit for running the brute force simulations and producing Figure 5.6.

Regions colored blue indicate higher quality lattices. We see that there is little sensitivity to a_1 since the slices at different a_1 values look very similar. However, in the $(a_0 - a_2)$ -plane there is a clearly defined region in which the Defect Measure is low.

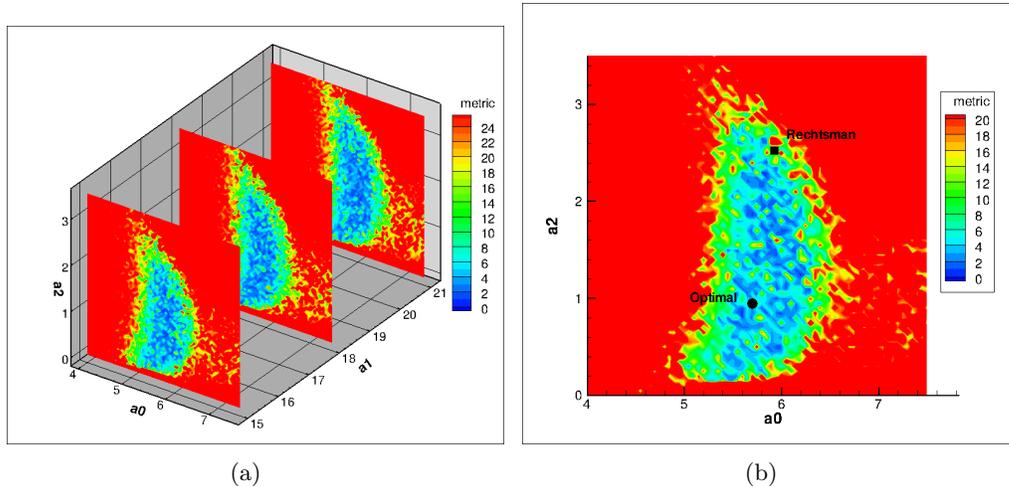


Figure 5.6: (a) Search over the parameters in the self-assembly potential, Equation 5.2, using the Defect Measure to assess lattice quality. (b) The slice from (a) in which $a_1 = 17.9$. The best parameters found by Rechtsman and co-workers is indicated with a square. The optimal parameter set as given by this search is indicated by a circle.

Figure 5.6(b) is the slice at $a_1 = 17.9$, with the location of the parameters found by Rechtsman et al. [75] marked by the square. The circle indicates the optimal parameter values as found by our brute force search. We see that Rechtsman et al.'s procedure found a point on the edge of the high lattice quality region (low Defect Measure). However, this is quite far removed from the optimal potential parameters. The self-assembly potentials found by Rechtsman et al. and this brute force search using the Defect Measure are shown in Figure 5.1. These potentials can be seen to have very different forms, the latter having a much larger repulsive section and having only a point of inflection at the lattice constant (where Rechtsman et al.'s potential has its first minimum).

Thus, the highest quality honeycomb lattices are formed from repulsive particles that spread out in the domain. The brute force potential shown in Figure 5.1 accounts for the observed sensitivity of the honeycomb lattice to the initial density of particles and for its fragility once formed.

Figure 5.7 illustrates that the lattice quality over the space of parameters for the potential is very noisy. Lattices are assessed here by the Defect Measure but similarly noisy results

were obtained for the other metrics considered, including the Local Template Measure.

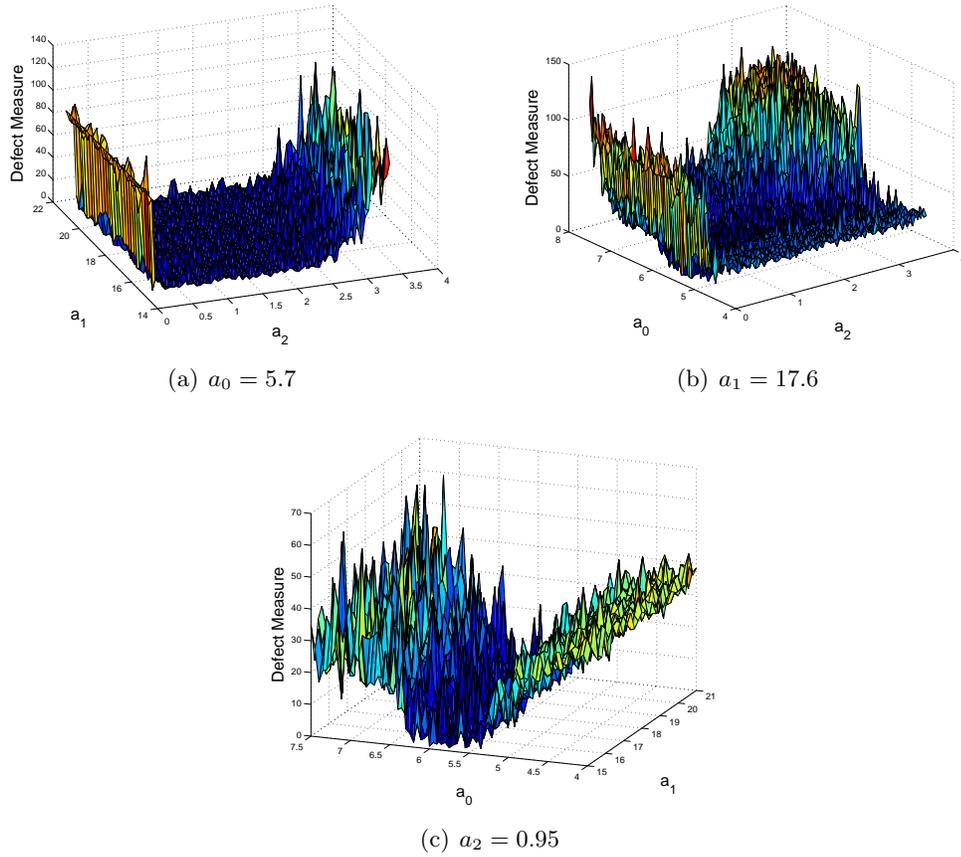


Figure 5.7: The Defect Measure over the parameters a_0, a_1, a_2 in Equation 5.2, with $a_3 = 1.823$. Each set of parameters yields a different isotropic inter-particle potential. Each potential was used in a simulation of particles cooling and the Defect Measure of the final configuration was computed to assess the quality of the lattice. The lattice quality is a noisy function but regions of higher quality can be seen as the dark blue areas.

5.1.3 Optimization of Potential for Self-Assembly

The brute force approach to designing the best potential for the self-assembly of particles in Section 5.1.2, is computationally expensive. Each set of parameter values in the potential requires at least one simulation to be run in order to assess the quality of the lattice formed. A more efficient way to find the best parameters is to use an optimization scheme that has the chosen quality metric as the objective function. Ideally, the optimization would be a more robust way of finding potentials and would also provide some information on the robustness of the potentials found.

Rechtsman and co-workers used a simulated annealing procedure to find the potential in Equation 5.2. In Section 5.1.4 we implement a trend optimization scheme that yields better lattices and the parameters for the potentials that form these better lattices are found up to one hundred times faster. The Defect Measure and the Local Template Measure were used as both the objective function in the optimization scheme and as metrics for assessing the quality of the lattices that were formed using the potentials found by the optimization procedure.

Any quality metric may be used as the objective function in optimization schemes and to assess lattice quality. However, the Defect Measure has two other features that can be used to shorten the time taken to evaluate the objective function. These objective function evaluations are frequently a large part of the cost of optimization schemes as they involve running simulations of particles being slowly cooled to see what configuration they self-assemble into.

The first application takes advantage of Step 2 of the Defect Measure algorithm, which can identify the type of lattice being formed (if it is reasonably well formed). If the lattice being formed in a simulation is not the same type of lattice as the target lattice then the simulation may be cut short. This is particularly useful when the target lattice is the honeycomb lattice because the competitor lattice is the very stable triangular lattice. For many values of the parameters in the potential, a triangular lattice (with lattice constant equal to the location of either of the first two minima) will form. Since triangular lattices are very stable, there is no need to continue the simulation; the particles will not form a honeycomb lattice with further cooling.

The second application of the Defect Measure in an optimization scheme is indicating when the simulations can be halted because the particles will not significantly change their positions with further cooling. For each set of parameters, a simulation will have to be run to determine whether those parameters lead to a good potential for self-assembly. Rather than having to run each large simulation for a long time, the Defect Measure of the particles can be computed during the simulation and when the Defect Measure levels off, the simulation can be cut short. Figure 5.8 shows the Defect Measure plotted against the time step for a single simulation. From time step 110 until the end of the simulation at time step 150, the Defect Measure stays relatively constant, indicating that the simulation could have been cut off at time step 110. Naturally, the Defect Measure would not need to be computed at

every time step, and not at the beginning of the simulation. This is especially useful if the simulations involve a large number of particles. As the number of particles increases, the relative expense of calculating the Defect Measure decreases, such that for large systems, the computational savings from ending simulations early outweighs the cost of computing the Defect Measure.

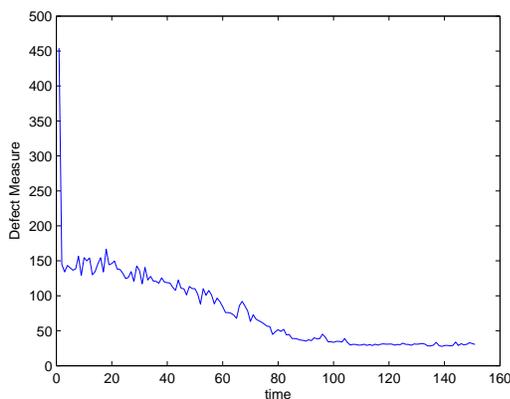


Figure 5.8: Defect Measure versus time for a single simulation.

5.1.4 Trend Optimization of Potential for Self-Assembly

Rechtsman et al. [75] used a simulated annealing optimization procedure to find potentials that lead to the self-assembly of particles into a lattice formation, as the temperature of the system is reduced. The final method described here leads to a one hundred times speed-up in the optimization of the parameters of the potential, and to higher quality lattices. The procedure for finding the best potential is more robust and the resulting potentials form the target lattice more robustly.

Choosing an optimization method requires the selection of both an optimization scheme and an objective function to optimize. We compare simulated annealing and trend optimization as our optimization schemes. The objective functions that we consider are the Defect Measure (Equation 4.1), the Local Template Measure (Equation 4.3), and the Lindemann Parameter (Equation 5.1).

Whereas the baseline geometric method utilizes only the geometry of the static target lattice, the optimization methods discussed in this section incorporate information about particle dynamics by optimizing parameters in the potential with respect to actual particle

simulations. All molecular dynamics simulations were performed on eight parallel processors using the LAMMPS software package from Sandia Laboratories.⁶

5.1.4.1 Objective Function Evaluation

Evaluation of the objective function $f(x)$, where $x = (a_0, a_1, a_2, a_3)$ is a point in the space of parameter values for the inter-particle potential given by Equation 5.2, requires a particle simulation with this inter-particle potential. These simulations were carried out with a relatively small number of particles (74 for the Lindemann parameter and 64 for the Defect Measure and Local Template Measure) in a periodic domain. As noted in Section 5.1.1, the Lindemann parameter is a measure of how much a target lattice configuration has melted and not the quality of the lattice, but using the Lindemann parameter as an objective function is advantageous since it is faster to evaluate than the other objective functions, which require much longer molecular dynamics simulations. Optimizing the Lindemann parameter is, however, an indirect method in that the quantity being optimized is not the quantity that will be used to determine the final quality of the potential. A more direct approach is to explicitly optimize the quality metrics. Although computationally more expensive to evaluate, optimization of these objective functions guarantees optimization of lattice quality and does not rely on the correlation of the objective function with lattice quality. Indeed, in some cases we have observed that a potential that produces a low value of the Lindemann parameter does not produce lattices of high quality.

To evaluate the Defect Measure and Local Template Measure objective functions, we start the molecular dynamics simulation with a temperature well above the lattice melting point and then reduce this temperature slowly until the particle motion is very small, that is, to a temperature very close to zero. The quality metrics are computed for the final lattice configuration. These simulations are expensive to carry out. Even though they may be run in parallel, objective function evaluations are still the major contributor to the time taken by the optimization procedure to optimize the potential parameters.

5.1.4.2 Objective Function Difficulties

As shown in Figure 5.7, the Defect Measure objective function is very noisy or spiky. Despite this, a region of generally lower Defect Measure can be clearly distinguished (the dark blue

⁶<http://lammmps.sandia.org> [71]

region). Repeating the simulation for the same parameter values frequently lead to very different Defect Measure values—up to an order of magnitude different. This is the main difficulty with our objective function. It makes the optimization non-trivial for available optimization techniques, and prompted the development of the trend optimization scheme. The other objective functions considered have the same characteristics.

Traditional gradient-based optimization methods can not deal with such noisy objective functions since gradients can not be computed using finite differences. If the Δx in $\Delta f/\Delta x$ is chosen larger, so as to try to avoid the spikes, the method becomes more like a patterned search method (discussed in Section 5.1.4.3). Such methods can be robust despite noisy objectives. Interpolation is another class of techniques used in the optimization community that would have difficulties with a noisy objective. The interpolation between sites of objective function evaluations will pass through spikes in the objective and the interpolation itself will be as noisy as the objective. Furthermore, these methods perform poorly with objectives that are highly variable in the sense that subsequent evaluations of the objective function at the same location may yield very different results.

A number of fitting or trend approaches that approximate the objective are in use and some of these are discussed briefly in Section 5.1.4.3. They can be very efficient even for noisy objective functions; however, highly variable objective functions frequently present a problem here too.

Thus, the primary features of our objective function are that it is

1. expensive to evaluate,
2. highly variable—repeated evaluations of the objective function at the same point in the domain can give values that are up to an order of magnitude different,
3. very noisy or spiky;
4. yet has a smooth trend.

After a brief history of gradient-free optimization we discuss the initial investigations that allowed us to construct the efficient procedure that we call *trend optimization*.

5.1.4.3 A Brief Survey of Similar Optimization Techniques

This section briefly surveys methods related to the trend optimization technique.

Trend optimization combines:

- trends, or surfaces that fit the objective function evaluations,
- global and local optimization of the trend,
- hierarchical optimization approaches using a hierarchy of coarser and finer trends, and
- handling of noisy or highly oscillatory objectives by envelopes.

Trend optimization is designed for the global optimization of highly noisy objective functions that are expensive to evaluate (possibly with crashes). Such problems are often encountered in engineering, finance, and life sciences.

There are many optimization approaches based on approximations of the objective function. Such approximations could be, for example, response surfaces, reduced-order models, surrogates, trends, coarse models, and many more. Some of these approximations are physics based (such as coarse models), some are based on data such as regressions (data fit) or interpolation, and some are hybrid, such as physics-based models with coefficients tuned to best fit the given data. A major difference should be noted between surfaces that interpolate the objective function, that is, exactly pass through the objective function values, and those that just fit the objective function values, for example, approximate the objective function in a least squares way.

For objective functions that are noisy, optimization techniques based on interpolation are prone to fail since they may try to follow the spikes of the objective function. Optimization approaches that use trends or data fitting approaches may be much more robust since the trends tend to reproduce the general shape of the objective function by smooth surfaces that ignore the noise and local details of the objective function. Hence, trend based approaches may be much more successful at finding the large local minima (as opposed to the deep spikes) of non-smooth objective functions.

There are a great number of interpolation and fitting approaches. Approximations that are used include polynomial interpolation, splines, Kriging methods, Taylor and other expansions, distance based interpolation, linear and nonlinear regressions, radial basis functions, neural networks, kernel based approaches, optimization based approaches, and more. For a survey see, for example, [38].

There is a vast body of literature on global optimization and multidisciplinary design optimization (MDO) that encompasses different hierarchical optimization approaches and surrogate approaches. Some reviews are [41, 69, 7, 10, 2]. Seminal papers in the area include [47, 46, 78, 37, 42].

Noisy or non-differentiable objective functions can be handled by optimization techniques such as pattern searches (discussed below), deterministic and stochastic searches (for example, Monte Carlo, genetic algorithms, simulated annealing, evolutionary algorithms), and techniques that smooth the objective (for example, multigrid optimization) or build smooth trends of the objective function.

Trend optimization has a large overlap with the pattern search algorithms developed by John Dennis, Charles Audet, and co-workers ([5] and references therein). They may be combined with surrogates and response surfaces, such as Kriging surfaces. Some key techniques are the Surrogate Management Framework (SMF) [9], the Generalized Pattern Search (GPS) [81, 82], Mesh Adaptive Direct Search (MADS) [6], and the Variable Neighborhood Search (VNS) [4]. The SMF incorporates the use of surrogate functions into a pattern search framework. The SMF and the other pattern search techniques provide a robust and efficient alternative to traditional gradient-based optimization methods and they can minimize functions that are noisy, have many local minima, and are relatively expensive to evaluate. Audet et al. [4] propose a way to combine the MADS, which extends the Generalized Pattern Search (GPS), with the Variable Neighborhood Search (VNS) metaheuristic for non-smooth constrained optimization. This paper also proposes a generic way to use surrogate functions in the VNS search. An application of SFM to a CFD problem is shown in Marsden et al. [58], which also has a summary of how SFM works. Pattern searches date back to the work of Hooke and Jeeves [40], where a robust algorithm that can handle noisy objective functions is presented; it is implemented in iSIGHT⁷.

Leading optimization software packages have algorithms that combine global optimization with response surfaces or surrogates. Some of these are Tomlab⁸, DAKOTA⁹, iSIGHT, IOSO¹⁰, and NOMAD¹¹.

The trend optimization scheme presented here is similar to other response surface meth-

⁷<http://www.engineous.com>

⁸<http://tomopt.com/tomlab>

⁹<http://www.cs.sandia.gov/DAKOTA/software.html>

¹⁰<http://www.iosotech.com/iosotech.htm>

¹¹<http://www.afit.edu/en/ENC/Faculty/MAbramson/NOMADm.html>

ods in global optimization that fit functions to the objective function rather than use interpolation. The added feature of our method is the upper and lower envelopes, which provide a measure of the variability in the objective function values. It is ideal for dealing with noisy (or discontinuous), widely variable, expensive-to-evaluate objective functions such as the one we have in the self-assembly problem.

5.1.4.4 General Trend Optimization Algorithm

1. Evaluate the objective function at some number of points in the domain.
2. Construct a trend surface to the objective function at these points and previously evaluated points.
3. Construct an upper envelope surface using information from all objective function evaluations thus far.
4. Construct a lower envelope surface in a similar manner.
5. Optimize the trend and determine points of the trend that are local minima.
6. Refine the domain by excluding regions in which the upper and lower envelopes are more than a specified maximum variability tolerance apart.
7. Repeat Steps 1–6 until a stopping criterion is met. Subsequent objective function evaluations in Step 1 should select points both near the local minima and in regions of lowest point density.
8. Perform a local search by any optimization technique in this region of flat envelopes.
9. The global optimal point is declared to be the point found by this local search.

This trend optimization algorithm¹² is very general in that

- many methods of choosing points at which to evaluate the objective function may be used;

¹²This algorithm was the result of discussions between the authors of [16], and Ronald Coifman from Yale University and his group.

- the trend and envelopes may be constructed by local averaging, Kriging, other surface fitting techniques, least squares fit constructed from a vast range of basis functions, or any other method;
- the optimization of the trend may be performed by any of the previously mentioned techniques, including gradient-based methods; and
- the local search can employ a wide range of methods.

The feature that makes the algorithm efficient and robust is the upper and lower envelopes. This allows us to handle highly variable objective functions and, importantly, gives a meaningful stopping criterion. Any other stopping criterion (given number of objective function evaluations, given computing time, an objective function evaluation that is lower than a given value, etc.) may be used, but the envelopes provide a natural stopping criterion that is related to the acceptable variability in the objective function. This stopping criterion is that the envelopes are flat and less than a maximum tolerance apart in a neighborhood of the global minimum of the trend. Flat here means that their variance is less than a given threshold.

The upper and lower envelopes specify which regions of the domain can be excluded in the refinement step, thereby focusing the procedure on regions with less variability in the objective function. The number of expensive objective function evaluations required is reduced by this refinement step. The procedure also gains efficiency by using a cheap method to optimize the trend.

Not only is the algorithm robust, but it provides a measure of the robustness of the final solution. Flat upper and lower envelopes indicate that all points in this region exhibit the same variability in repeated evaluations of the objective function. The minimum tolerance distance specifies the maximum variability that is acceptable in the solution. Any point in the region of flat envelopes near the global minimum of the trend is an equally good solution. The closer the global optimal point is to the center of the region in which the envelopes are flat, the more robust it is to variations in the parameters spanning the domain. Thus, the trend optimization scheme provides a measure of the robustness of the optimal point. A point that is at the edge of the region of flat envelopes may be moved to a point of high variability in the objective function by a small variation in the domain parameters. Such a point is not robust even though it is in the region of equally good solutions.

A note on convergence of the algorithm

Jones et al. [47] studied issues related to balancing the exploitation of the approximating surface for optimization with the need to improve the approximation in the response surface methodology. They prove that if the objective function is continuous on a compact, finite-dimensional domain and a response surface interpolates the objective, then the above algorithm, without the refinement of the search area in Step 6, will converge towards the global minimum of the objective.

When the objective function is not continuous, this convergence result no longer holds. However, it is intuitively plausible that if we sample points in the domain such that the maximum distance between sampled points is no greater than a defined characteristic width of the deep spikes, then the algorithm will find the “global minimum”, ignoring features smaller than this characteristic distance.

5.1.4.5 Hierarchical Nature of the Trend Optimization Scheme

Trend optimization can be performed in a hierarchical manner by building a hierarchy of trend surfaces. Each successive iteration of the procedure can build its own unique trend. These trends can differ in four key ways. The most straightforward of these is that successive trend surfaces may use previous objective function evaluations to build a surface which better fits the objective function, since more points are used than in the previous iteration. The second property of the family of trends is that after the initial global trend, subsequent trend surfaces may be more local due to the refinement of the search area (by excluding regions in which the envelopes are more than a maximum variability tolerance apart). Third, successive trends may use more complex basis functions or a larger basis when fitting surfaces so as to get a better trend. Finally, the method used to construct the trend may be more sophisticated in successive iterations. The last two properties yield a hierarchy of coarse to fine trends.

Note: In Step 7 of the algorithm, points are sampled not only in the region of the global minimum but also at areas of lowest point density to ensure that no areas of local minima have been missed by previous sampling steps.

5.1.4.6 Initial Investigations

We present here the initial investigations that were carried out before implementing the final trend optimization scheme for our self-assembly problem. It would not be fair to present the final method without a discussion of the previous work that lead to confidence in our results. This previous work lead to the development of the trend optimization method that is expected to be useful in many other applications with similarly difficult objective functions.

We seek the lowest point of the objective function in the regions of the relevant local minima of the trend. We first briefly discuss the method we used to construct the trend and envelopes.

Points from the domain are chosen using Monte Carlo sampling, and the trend, and upper and lower envelopes, are constructed from the objective function evaluations at these points, as shown in Figure 5.9. We first define a radius in parameter space, r_{nbhd} , that is the distance, or neighborhood, around each point that will be taken into account when constructing the envelope. These envelopes are interpolated based on local averaging. For each point, find the minimum and maximum value of the objective function in a neighborhood around that point,

$$\begin{aligned} f_{nbhd}^{min}(x_p) &= \min\{f(x) : |x - x_p| < r_{nbhd}\} \\ f_{nbhd}^{max}(x_p) &= \max\{f(x) : |x - x_p| < r_{nbhd}\}. \end{aligned} \quad (5.4)$$

Our upper envelope is a surface that goes through these $f_{nbhd}^{max}(x)$ values and the lower envelope is a surface that goes through these $f_{nbhd}^{min}(x)$ values. An infinite number of surfaces can qualify as an upper envelope (likewise for lower); we require only that it should touch $f(x)$ at least at one point and that it should be smoother than $f(x)$.

The trend is a smoothed version of the objective function; it shows the simple smooth shape underlying the objective function. A straightforward way to find a trend is by finding the average value of the objective function in a neighborhood around each point, $f_{nbhd}^{avg}(x)$, and constructing the surface through these values.

In our method we fit a linear regression such that the difference between $f(x)$ and the trend at x is minimized in some norm. We tested a variety of basis functions for the trend.

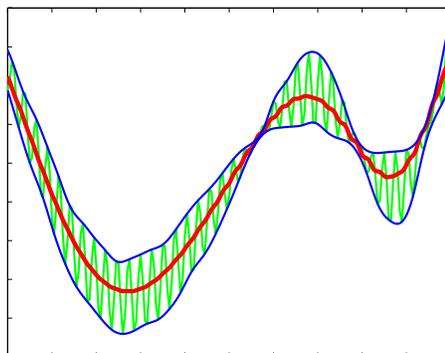


Figure 5.9: Schematic of the trend and upper and lower envelopes for a noisy objective function with a simple trend. The objective function is shown in green, the trend is red, and the upper and lower envelopes are blue.

Once a satisfactory trend has been found, any desired method for finding the global minimum of it can be used. In the regions of the minima, local searches using more objective function evaluations can be conducted to find the multiple relevant minima. The evaluation of the trend is very fast, unlike the evaluation of the objective function, so it is feasible to evaluate the trend at a large number of points, and thus perform a robust global optimization of the trend.

The process of choosing points by Monte Carlo sampling and evaluating the trend and envelopes at these points can be repeated any number of times to get better and better surfaces. Successive selection of points should be chosen not only around the region of the global minimum of the trend but also in other regions of lowest point density.

Initial sensitivity studies were conducted on a one-dimensional slice of the objective function with the goal of tuning the trend optimization method to our particular lattice quality objective function. These investigations varied the number of points sampled at each step, as well as the number of steps in the procedure. They revealed appropriate basis functions and methods to use to fit the trend, tolerances on acceptable variability, and methods for optimizing the trend. We present our final method below.

Trend Optimization Algorithm used in Self-Assembly Problem

1. Evaluate the objective function at 20 randomly sampled points drawn from a uniform

distribution.

2. Construct a trend by computing a least squares fit to these points.
3. Evaluate the quadratic at 10,000 Monte Carlo-selected points and determine the points at which the trend has local minima.
4. Refine the size of the domain by reducing the length along each dimension by a factor of 3 and re-centering the refined domain on the minimum of the quadratic.
5. Repeat the above steps, using all objective function evaluations performed thus far in Step 2.
6. Evaluate the objective function at another 20 points in this doubly refined area of the global minimum.
7. The optimal parameter values (minimum of the objective function) is declared to be the point in this refined area at which the objective function is lowest.

This procedure is fast because although it is expensive to evaluate the objective function, fitting the quadratic and evaluating it at 10,000 points requires very little time. Each set of 20 objective function evaluations may also be done in parallel thus reducing the wall-clock time for the optimization. We also note that all previous objective function evaluations are used to construct the trend and to find the global minimum in the final step. It is these features and the simple smooth nature of the trend that make this algorithm so efficient for our particular implementation.

We note that up to 14 successive sampling steps were considered. The factor by which the search area could be refined at each step was also explored, using factors from 0.1 to 0.5. These studies indicated that 2 sampling steps were sufficient, each consisting of 20 sampled points, and that the search area could be refined by a factor of 1/3. It is this realisation that lead to this method being 100 times faster than a simulated annealing procedure.

These values were demonstrated to be sufficient by examining the upper and lower envelopes. Flat upper and lower envelopes after two sampling steps indicate that all points in this region exhibit the same variability in repeated evaluations of the objective function, and thus, are equally good. Hence, potentials using parameters from this region are equally robust in the sense of producing the same quality lattices with the same potential. The

robustness of the potential with respect to changes in its parameters is proportional to the distance to the center of the flat region of the envelopes (Section 5.1.4.4).

Although designed with the potential for the self-assembly of a honeycomb lattice in mind, the above procedure has been successful in optimizing the parameters for other self-assembly potentials as well.

5.1.4.7 Comparison with Simulated Annealing

Simulated annealing is known to perform badly with objective functions such as the one we have here. The results presented in Section 5.1.4.8 used a contributed simulated annealing procedure available on the MATLAB[®] website.¹³ An enhanced simulated annealing procedure with many more parameters was also used that was tuned to better deal with our objective function. Five hundred runs of this simulated annealing optimization, with different combinations of parameters, gave greatly varying optimal potentials. This variety of optimal potentials naturally lead to great variability in the final lattice quality of simulations with each potential. In contrast, the trend optimization procedure gave consistently higher quality lattices (via better potential parameters) and did so 100 times faster than the best out of the 500 enhanced simulated annealing runs. That is, it found a better global minimum of the objective function and did so 100 times faster than simulated annealing.

Averaging multiple objective function evaluations at the same point in the domain would give a smoother function on which to run simulated annealing; however, this is extremely inefficient. It is faster and more robust to follow the trend optimization procedure explained above. We note that simulated annealing may be used in the trend optimization algorithm in Step 5 or 8, and we would expect good results in this case.

5.1.4.8 Results

After running an optimization method, we can assess the quality of the resultant potential by measuring the quality of the lattices produced by the potential under cooling. The Defect Measure and the Local Template Measure can be used to measure lattice quality but the Lindemann parameter is not a measure of lattice quality (Section 5.1.1). Since we know from the envelopes that repeated simulations with the same potential can produce lattices with quality metric values varying by a factor of ten, for the results presented here, we

¹³MATLAB, 2007a, The MathWorks, Natick, MA

average over 15 cooling simulations with different random initial conditions for each set of potential parameters. The cooling simulations started from a temperature approximately 1.5 times the melting temperature of the lattice and cooled, using a Nosé-Hoover thermostat [30], to less than 10% of the melting temperature.

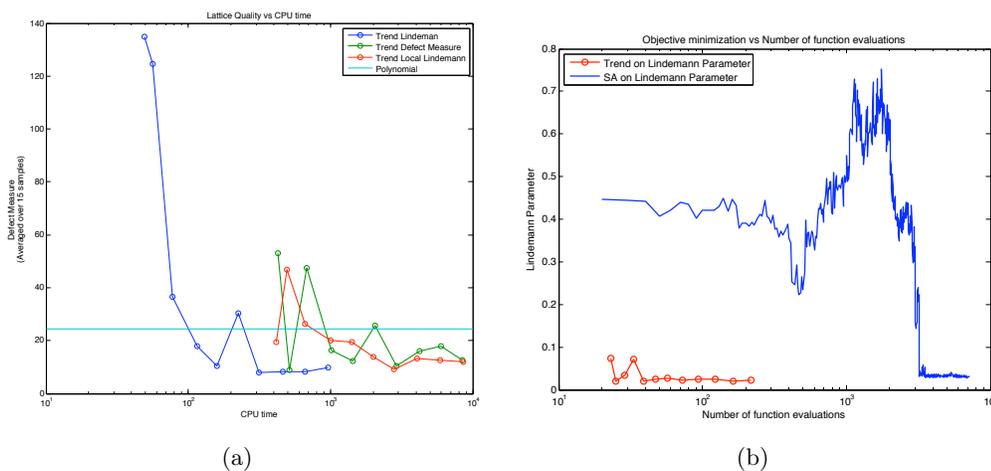


Figure 5.10: (a) Defect Measure versus CPU time for the different trend optimizations that were run. The objective functions of these optimizations were the Lindemann Parameter, the Defect Measure and the Local Template Measure (called Local Lindemann in the figure). Each data point represents a separate simulation run for the given CPU time. The polynomial potential from the baseline geometric method (Section 5.1.1.1) is indicated for comparison; it required no CPU time. (b) Lindemann Parameter versus the number of function evaluations both for a trend optimization and a simulated annealing run with the Lindemann Parameter as the objective function.

Figure 5.10(a) shows the average lattice quality (assessed by the Defect Measure) for each potential found versus the CPU time taken to find the potential. Results are shown for trend optimization using the Lindemann Parameter, the Defect Measure, and the Local Template Measure (called Local Lindemann in the figure). We see that evaluating the Lindemann Parameter is ten times faster than evaluating the other two metrics that measure quality. The baseline geometric method's polynomial potential is indicated for comparison; it required no CPU time. Similar results were obtained using the Local Template Measure as the final lattice quality metric.

Figure 5.10(b) is a direct comparison of the trend optimization scheme and simulated annealing, using the Lindemann Parameter as the objective function. The latter is the method employed by Rechtsman et al. to find their self-assembly potential parameters.

The number of function evaluations is plotted on the horizontal axis. Trend optimization can be seen to require 100 times fewer function evaluations than simulated annealing, for the same final Lindemann Parameter. Not all simulated annealing runs performed as well as the one shown here but the trend optimizations performed consistently well. The speed-up can be attributed to the fact that the objective function is noisy yet has a simple trend, two properties for which trend optimization is ideally suited. Furthermore, using trend optimization on the more expensive quality metric objective functions still provides a ten-fold speed up and provides the guarantee that the final potentials produce high-quality lattices.

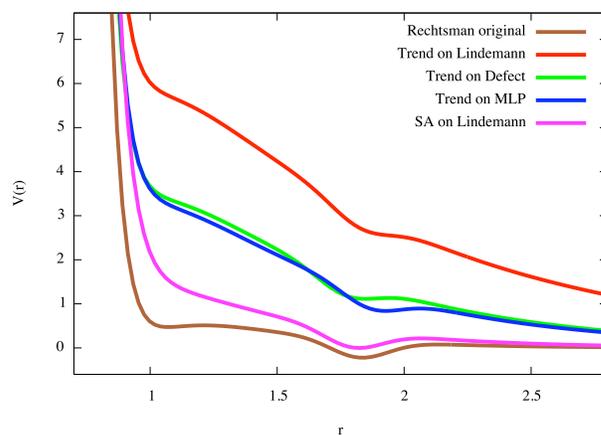


Figure 5.11: Potentials for the self-assembly of particles into a honeycomb lattice. The different colors denote different methods used to find the form of the potential. Brown: Rechtsman et al.’s potential (Equation 5.2); Red: potential resulting from trend optimization with the Lindemann parameter (Equation 5.1) as the objective function; Green: potential resulting from trend optimization with the Defect Measure (Equation 4.1) as the objective function; Blue: potential resulting from trend optimization with the Local Template Measure (MLP in the figure; Equation 4.3) as the objective function; Magenta: potential resulting from simulated annealing with the Lindemann parameter as the objective function.

Figure 5.11 shows the potentials generated by the various methods discussed. The potential found by Rechtsman et al. is also shown for reference. When comparing the potentials, the most striking feature that we observe is that trend optimization generates potentials that are more repulsive and do not have potential wells. This leads to higher quality lattices since particles do not get stuck in local minima associated with the potential wells.

5.1.5 Quantifying Robustness of Potentials

A good potential for the self-assembly of particles will be robust to uncertainty in the parameters of the potential, the density of particles, the particle initial conditions and the cooling schedule. The larger the range of values over which the final lattices formed are acceptable, the more robust the potential is.

Two ways for finding this range of values will be presented. The first looks at lattices formed from potentials that are equally good and the second looks at lattices that are acceptably good. The notion of equally good potentials comes from the discussion of the envelopes used to establish the trend optimization procedure.

5.1.5.1 Equally Good Potentials

Potentials that are equally good at producing the highest quality lattices possible, can be found from the envelopes discussed in Section 5.1.4. Once the search area over the domain of parameters of the potential has been refined so that the upper and lower envelopes over the refined area at the global minimum of the trend are flat and less than a maximum variability tolerance apart, this refined area defines the potentials that are equally good at producing high quality lattices. They are equally good because the variability in repeated objective function evaluations for these parameters is the same. Thus, these potentials are equally robust to changes in the initial conditions of particles.

The robustness of the potential with respect to changes in its parameters is proportional to the distance to the center of this region where the envelopes are flat.

5.1.5.2 Acceptably Good Lattices

There are two ways to define what an acceptably good lattice is. Lattices that have a value of the chosen quality metric that is below a certain threshold value can be accepted as good lattices. The value of this threshold will depend on what the assessor deems to be a good enough lattice and which metric is being used.

The second way to define the acceptably good lattices is by using Step 2 of the procedure to compute the Defect Measure. If the lattice can be identified by the algorithm to be of the same type of lattice as the target lattice, then it is an acceptably good lattice. Such lattices will have a majority of particles that have the correct number of nearest neighbors,

next nearest neighbors, and third nearest neighbors. This definition of a good lattice works best for lattices with many more inner particles than boundary particles.

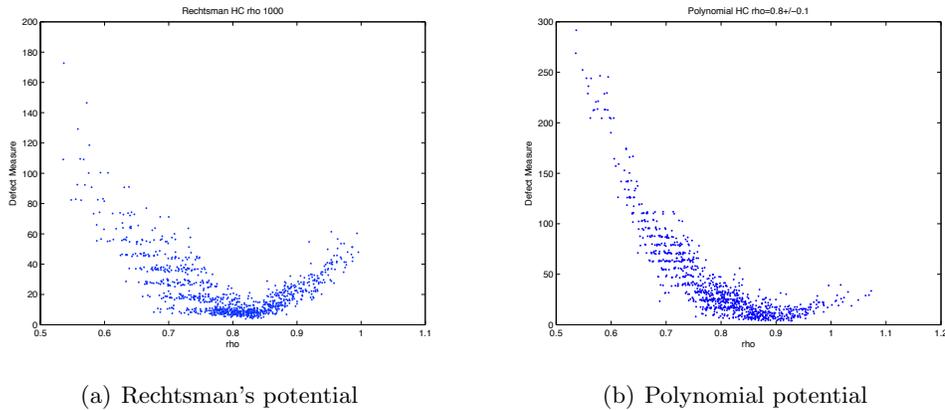


Figure 5.12: Defect Measure versus density of particles for (a) Rechtsman's potential and (b) a piecewise polynomial potential. Each point corresponds to one simulation of 65 particles in a domain without periodic boundary conditions. The weights corresponding to Defect Measure 1 were used.

Figure 5.12 plots the Defect Measure versus density of particles for two honeycomb potentials: Rechtsman's potential and a simple piecewise polynomial potential designed using the baseline geometric method explained in Section 5.1.1.1. The weights of Defect Measure 1 were used in the computation. Note that 15 points were omitted from Figure 5.12(a). These points all had a Defect Measure greater than 495 and a density greater than 1. Each point in the plot corresponds to a simulation of 65 particles in a domain without periodic boundary conditions. There were 1000 simulations with densities given by a Gaussian distribution with mean of 0.8 and standard deviation of 0.1.

Defining acceptably good lattices to be those having a Defect Measure less than 30, Rechtsman's potential yields good lattices from densities of 0.6334 to 0.9578. The polynomial potential yields good lattices over a range of densities from 0.6884 to 1.0690. This corresponds to a spread of 0.3244 for Rechtsman's potential and a spread of 0.3806 for the polynomial potential; the latter is 17% larger. Thus, the polynomial potential can be said to be more robust to uncertainty in the density of particles. It forms good lattices over a wider range of densities.

The sensitivity of potentials to the parameters in the potential and the cooling schedule can be examined by a similar procedure.

5.2 Construction of the Phase Diagram of Krypton Atoms on a Graphite Substrate

The lattice quality assessment tools developed in Section 4.2 can be used to detect phase transitions from a disordered to an ordered state. In this way, they provide a low-dimensional representation of the system. As the system is cooled, the quality metric can be monitored until it reaches a threshold value that corresponds to an acceptably good lattice being formed. The temperature at this time indicates the temperature at which the system has transitioned for the density of particles in the simulation. By repeating this procedure at different densities, a phase transition curve can be constructed. This method is straightforward to implement and we do not discuss it further here. Instead, we implement a temperature quench method that is not as computationally expensive.¹⁴ We will show that our results agree very well with experimental results.

Krypton atoms above a graphite substrate are an example of a system that forms a lattice in one phase. The graphite substrate provides the krypton atoms with adsorption sites in the form of a triangular lattice [11, 53]. The LAMMPS simulation package from Sandia Laboratories was used to model this system.¹⁵ A monolayer of krypton atoms was maintained by a reflective wall above the substrate and periodic boundary conditions in both co-ordinate directions were enforced. The Defect Measure was used as the quality metric because of its adaptability and short computation time.

The simulations follow a temperature quenching procedure, as explained by Gelb and Müller [32]. The system is equilibrated at a temperature above the phase transition curve, thus providing a system that is all in the fluid phase. The temperature is then quickly dropped to a temperature that is below the phase transition curve. The density of the system must be between the densities of the phase transitions to the solid and fluid phases at the final temperature. This procedure is illustrated in Figure 5.13 where the system is initialised at some point *A* that is at a high temperature. After the temperature drop, the system moves to point *B* that is at a lower temperature. It has a density between points *C* and *D* on the phase transition curve that have the same temperature. It has a temperature that is below the broad, flat part of the phase transition curve.

¹⁴Thanks to Marco Arienti for suggesting the use of this method.

¹⁵Thanks to Vladimir Fonoferov from Aimdyn for modifying the code to enable implementation of these simulations.

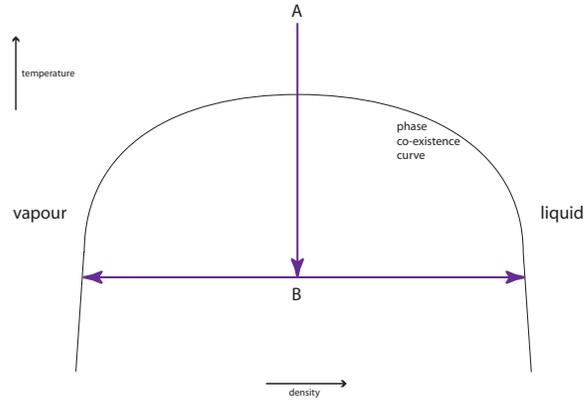


Figure 5.13: Schematic illustration of the temperature quench method. Based on Figure 1 in [32].

The temperature quenching (sharp drop in temperature) causes the system to become unstable and to spontaneously separate into fluid and solid regions. It is not necessary to let the system as a whole equilibrate; only local equilibrium is required. This leads to some savings in computational time. Two points on the phase transition curve can now be found by identifying the solid and fluid phases and calculating their densities.

Constructing the phase diagram involves two methods, one of which is used at low temperatures and the other at higher temperatures. The main difference between the two methods is the form of the Defect Measure that is used.

5.2.1 Low Temperature Region

In the low temperature region of the phase diagram, the densities at which there is a transition to a purely fluid and a purely solid phase are far apart. Thus, the phase that an individual particle belongs to (in a system with both fluid and solid phases) can be determined by counting the number of nearest neighbors of the particle. Since the graphite substrate provides adsorption sites in a triangular lattice formation, a particle having six nearest neighbors is in the solid phase. Particles with fewer than six nearest neighbors are either fluid particles or particles on the boundary of fluid and solid regions. Counting the number of nearest neighbors of particles corresponds to using the Defect Measure with weights $\omega_{\text{missing}} = \omega_{\text{extra}} = 1$ and $\omega_{\text{displaced}} = \omega_{\text{boundary}} = \omega_{\text{none}} = 0$. Particles in the solid phase will have a Defect Measure less than one, whereas fluid and boundary particles will have Defect Measure greater than one. Boundary particles are identified as particles that

have fewer than six nearest neighbors and that have at least one nearest neighbor that has the correct number of nearest neighbors.

Once each particle has been identified as either a solid, fluid, or boundary particle, the density of the two phases can be calculated by attributing an area to each particle. This is done by finding the Voronoi tessellation of the particles in the plane [86]. Figure 5.14 shows a typical snapshot of the system at a temperature of 70K after the temperature quench. Solid particles are colored red, fluid particles are colored blue, and boundary particles are green. Voronoi cells that correspond to fluid or boundary particles are colored yellow. The small points around the edge of the figure show where the particles on the other side of the domain are mapped to under the periodic boundary conditions. This is necessary to account for when finding the Voronoi cell of each of the particles because of the edge effects discussed in Section 4.2.4.

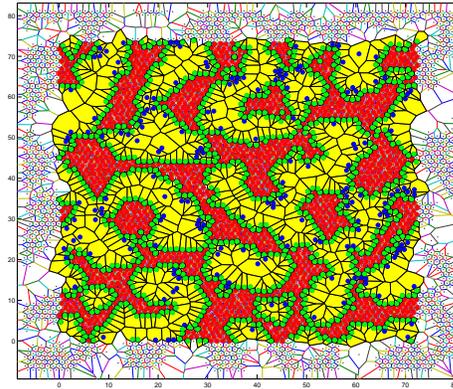


Figure 5.14: A typical snapshot of krypton atoms (that are on a graphite substrate) at 70K. Solid particles are colored red, fluid particles are colored blue and boundary particles are green. Voronoi cells that correspond to fluid or boundary particles are colored yellow. Points around the edge show where the particles on the other side of the domain are mapped to under the periodic boundary conditions.

Let n_{solid} be the number of particles identified as being in the solid phase, and A_{solid} be the sum of the areas of the Voronoi cells corresponding to solid particles. Similar notation applies to fluid and boundary particles. The density of the solid and fluid phases are then

calculated as

$$\rho_{\text{solid}} = \frac{n_{\text{solid}}}{A_{\text{solid}}}, \quad (5.5)$$

$$\rho_{\text{fluid}} = \frac{n_{\text{fluid}}}{(A_{\text{fluid}} + A_{\text{boundary}}) - \left(n_{\text{boundary}} \times \frac{A_{\text{solid}}}{n_{\text{solid}}}\right)}. \quad (5.6)$$

The area attributed to the fluid phase (the denominator in Equation 5.6) is the sum of the area of Voronoi cells of particles that are either fluid or boundary particles less an approximation of the area taken up by the boundary particles. The area of boundary particle Voronoi cells is necessary to take into account so as to include all the free space not taken up by the solid phase. However, the boundary particles are not to be counted as fluid particles so we assign an area to them corresponding to the average area of a solid particle Voronoi cell.

A lower estimate on the fluid density can be found by not subtracting the area taken up by the boundary particles. An over-estimate of the fluid density can be found by looking for particles that have Voronoi cells with an area greater than 1.5 times the average area of a solid particle. Such particles are colored black in Figure 5.15 and their Voronoi cells are colored yellow. This includes many of the boundary particles in the numerator. Thus

$$\text{Fluid density under estimate} = \frac{n_{\text{fluid}}}{A_{\text{fluid}} + A_{\text{boundary}}}, \quad (5.7)$$

$$(5.8)$$

$$\text{Fluid density over estimate} = \frac{n_{\text{large cell}}}{A_{\text{large cell}}}. \quad (5.9)$$

The results of this low temperature method are shown in Figure 5.18 as the black points for temperatures below 85K (vertical axis). The horizontal axis shows the commensurate monolayer density. This is found by comparing the densities calculated above to the density of krypton atoms if there were a complete monolayer. Thus, a commensurate density of 1 corresponds to all adsorption sites of the krypton and graphite system being occupied. The fluid density under- and over-estimates are indicated by the green lines. The red points are the results of a 10,000 atom molecular dynamics simulation using the LAMMPS software package.¹⁶ The blue lines indicate experimental results obtained by Butler and Lehrer

¹⁶Thanks to Vladimir Fonoberov from Aimdyn for these results.

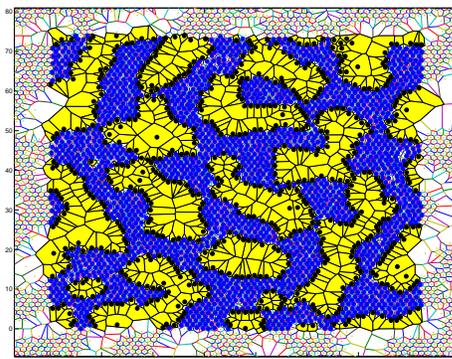


Figure 5.15: A snapshot of krypton atoms on graphite showing which particles and Voronoi cells are used for the fluid density over-estimate. Cells with an area greater than 1.5 times the average area of a solid particle are colored yellow and the particles belonging to these cells are colored black. Points around the edge show where the particles on the other side of the domain are mapped to under the periodic boundary conditions.

[11, 53].

There is good agreement between our method and the two other methods. The advantage over the molecular dynamics simulation is that the simulations did not need to be run for as long since we require only local equilibrium and not equilibrium of the whole system. There was also a reduction in computational cost as our results were obtained using approximately a quarter of the number of atoms. Other advantages of this method are that only one snapshot of the system is needed, any snapshot provides two points on the phase diagram (one for fluid and one for solid density), and the same initial equilibrated system can be used for each final temperature considered.

This method of constructing the phase transition curve only works well in regions of the phase diagram where the transition to pure fluid and pure solid phases are far apart. The method is not suitable in the neck region at higher temperatures or in the broad flat region around 85K. In these regions we must use a different form of the Defect Measure.

5.2.2 High Temperature Region

At temperatures where the solid and fluid phase transition densities are similar, the order of the particles is the feature that distinguished the two phases. The Defect Measure weights that are used are similar to that of Defect Measure 1 in Section 4.3: $\omega_{\text{displaced}} = \omega_{\text{missing}} =$

$\omega_{\text{extra}} = 1$, and $\omega_{\text{boundary}} = \omega_{\text{lone}} = 10$.

At these higher temperatures, the krypton atoms oscillate around their adsorption sites more than at lower temperatures. To account for this, we average the positions of each particle for a number of timesteps after the system has reached local equilibrium. In the simulations used to produce the results shown in Figure 5.18, we average for 500 timesteps, corresponding to 3.75×10^{-12} s. The oscillation of a particle in the solid phase around its adsorption site for this period is shown in Figure 5.16, with the green point indicating the average position. The oscillations are about 0.3 times the distance between adsorption sites. Note that it is feasible at most temperatures considered to average for fewer than 500 timesteps.

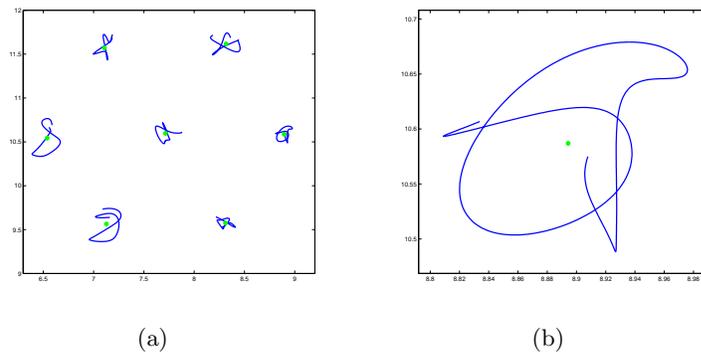


Figure 5.16: (a) Trajectory of krypton atoms around their adsorption sites on the graphite substrate for 500 timesteps. The green points indicate the average position of each particle. The width of the oscillations are about 0.3 times the distance between adsorption sites. (b) Close-up of one particle's oscillations.

After calculating the Defect Measure of each particle (using the average particle positions), the histogram of the Defect Measures will have a peak at low Defect Measures (at some value less than 1) and other peaks at higher Defect Measures. This can be seen in the example shown in Figure 5.17 (note that this figure omits higher Defect Measures). We choose a threshold value between the first and second peaks in the distribution. In the example, and in the simulations used to produce Figure 5.18, the threshold Defect Measure that defines the phases is a value of 0.5. Below this, particles are in the solid phase and above this they are in the fluid phase. For the weights specified above, the threshold value will be less than one. Other weights may give different threshold values but the appropriate threshold value is between the first and second peaks of the distribution.

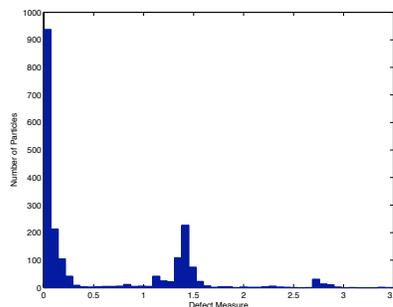


Figure 5.17: Distribution of Defect Measure values (horizontal axis is truncated so higher values of the Defect Measure are not shown). Below the threshold Defect Measure of 0.5, particles are in the solid state and above the threshold they are in the fluid state.

The density of the phases is found by using the area of the Voronoi cells of particles corresponding to each phase, as explained in Section 5.2.1. The commensurate density of the points on the phase transition curve is shown in Figure 5.18 for temperatures from 85K upwards (from the black diamonds upwards).

It can be seen that there is quite good agreement in the high temperature region with the 3D molecular dynamics and experimental results. The shape of the phase transition curves are similar in all three cases, however the Defect Measure method seems to be offset slightly.

The advantage of our method is that we have used only a quarter of the particles that the molecular dynamics simulation results used. It is possible to use fewer particles still. The other advantages mentioned above for the low temperature method also hold, apart from needing only one snapshot. It is necessary to find the average position of the atoms as they are oscillating about their adsorption site. However, good results can be obtained with averaging fewer than 500 timesteps.

The example of krypton atoms on a graphite substrate considered here illustrates that the Defect Measure is a good coarse description of the system as it transitions between fluid and solid phases. The temperature quench method reduces the computational time taken to find the phase diagram of the system. The phase transition curve found by using these tools is in agreement with experimentally found values.

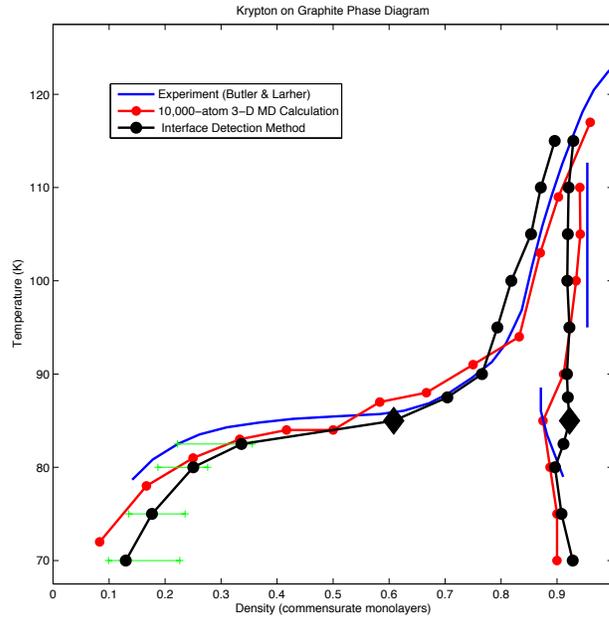


Figure 5.18: Phase diagram for krypton atoms on a graphite substrate. Results using the Defect Measure method are shown in black. The low temperature method is used for temperatures below 85K and the high temperature method for temperatures 85K (shown as a black diamond) and above. Fluid density under and over estimates are represented by the green lines. Approximately 2500 atoms were used for each simulation. Molecular dynamics results, using 10,000 krypton atoms are shown in red. Experimental results are shown in blue [11, 53].

5.3 Identification of Global Defects

The Voronoi metric can be used to identify global defects in a lattice by looking at the local area of each particle. Calculating the Voronoi metric of each particle and shading the Voronoi cell of particles that have a Voronoi metric value greater than some threshold, yields Figure 5.20(a). This lattice has 576 particles and periodic boundary conditions in both directions. The threshold Voronoi metric value used is 5×10^{-5} . This threshold value can be found by looking at the distribution of Voronoi metric values.

For example, Figure 5.19 shows the distribution of the area of Voronoi cells in blue. The green dashed line represents the value of (area of bounding box)/(number of particles), and the shorter, red dashed lines represent this value plus/minus the threshold value. This threshold value was chosen to include the part of the distribution centered around the (area

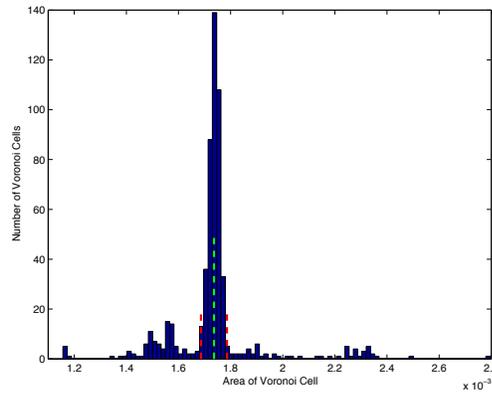


Figure 5.19: Distribution of the area of Voronoi cells. The dashed green line represents the value of $[(\text{area of bounding box})/(\text{number of particles})]$, and the shorter, red dashed lines represent this value plus/minus the threshold value. This threshold value was chosen to include the part of the distribution centered around the $[(\text{area of bounding box})/(\text{number of particles})]$ value.

of bounding box)/(number of particles) value.

Local as well as global defects of the lattice can be seen in Figure 5.20(a). The hexagonal shape (made up of seven Voronoi cells) indicates an extra particle in the lattice, whereas the triangular shape (made up of three Voronoi cells) indicates a missing particle. These local defects are much easier to locate at a glance by this method.

Global defects such as domain walls or grain boundaries can be seen by observing contiguous shaded Voronoi cells. The lattice in Figure 5.20 has a domain wall extending from the left-hand side down to the lower edge, which then continues from the top to the right-hand side due to the periodic boundary conditions. There is another domain wall that joins up with this domain wall, beginning in the top right and extending upwards until it comes back into the bounding box at the bottom left. These domain walls are more clearly defined on Figure 5.20(b) where Voronoi cells that share at least two vertices with cells that have a large value of the Voronoi metric (shaded magenta), are colored green.

Note that the Voronoi tessellation yields cells that have many sides, including short sides that can not easily be seen in Figure 5.20. The vertices that are used in determining which cells are colored green must adjoin a side that is not one of these short sides, that is, a side that is visible in the figure. Short sides are defined heuristically to have a length of less than 0.03 in Figure 5.20.

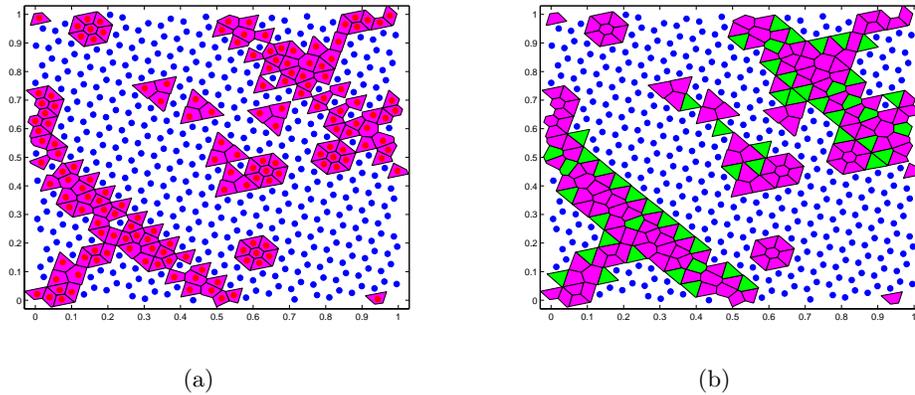


Figure 5.20: Identification of global and local defects using the Voronoi metric. (a) Voronoi cells of particles with a high Voronoi metric are colored magenta. These cells identify areas where there are defects. Local defects such as missing and extra particles can easily be located by their characteristic triangular and hexagonal, respectively, shaded regions. Global defects such as domain walls can be identified by contiguous shaded cells. (b) Global areas of defects, for example domain walls, are more clearly defined by shading cells green that share at least two vertices with magenta cells.

Apart from the domain walls mentioned above, the shading rule that yields Figure 5.20(b) also joins the defects in the middle of the domain that were isolated defects in Figure 5.20(a). Thus, this procedure based on the Voronoi metric highlights regions of the lattice that are weaker or not as well formed, as well as identifying domain walls.

5.4 Identification of Clusters to be treated as Rigid Bodies

The self-assembly of particles with pairwise isotropic inter-particle potentials into a lattice requires long and computationally expensive simulations. The temperature of the particles must be decreased slowly to allow a lattice to form. One way to speed up such simulations is to identify particles that have already formed a lattice structure and treat these particles as a rigid body, thus decreasing the number of evaluations of the potential function. This procedure works best for lattices that form nuclei that grow and join up to form a lattice. The triangular lattice is ideal because of its stability once formed. The polynomial potential in Section 5.1.5 for the honeycomb potential also forms in this way.

A cluster of particles that may be treated as a rigid body can be identified by observing how much the inter-particle distances change over time. If these distances have not changed

significantly for an appropriate length of time then the particles are most likely in the lattice configuration and may be treated as a rigid body.

However, there may be some defects in this cluster of particles. A lattice with a defect is typically less stable than a lattice free of defects. Over time, the particles will try to eliminate the defect or move it towards the edge of the otherwise well-formed cluster. This takes more time than the free movement of particles that are not a part of a cluster. So a cluster that has a defect may be identified as a cluster to treat as a rigid body because the inter-particle distances do not change much.

To allow defects within clusters of particles to be eliminated, the Defect Measure can be calculated for each candidate rigid body particle. Those particles with a high Defect Measure (those close to the defect) would then not be included in the cluster of particles to be treated as a rigid body. This would allow the defect to move to the boundary of the cluster. If the particles with a high Defect Measure are completely surrounded by particles with a low Defect Measure, the defect is likely to be stuck in the interior of the cluster.

A related application of the procedure for calculating the Defect Measure is in that of simulating nucleation. Clusters of particles that have formed a lattice either grow or combine with other clusters to form a larger lattice cluster. This process of nucleation can be modelled by allowing the outer particles of a cluster to move freely while keeping the inner particles in a rigid body core, since cluster growth is a process that takes place on the boundary of clusters. The method for identifying the boundary of a group of particles in a lattice configuration (see the Defect Measure algorithm in Section 4.2.1.1) can be used to identify these boundary particles, which are then taken out of consideration while the next layer of boundary particles is found. As many layers as necessary may be found in this way, although for the triangular lattice, without defects, two layers or “skins” is sufficient.

5.5 Conclusions and Future Directions

The Defect Measure, and other quality metrics, have been used in implementing an optimization procedure that finds parameters for potentials that lead to the self-assembly of particles into desired lattice configurations. Apart from being the objective function, it was used to measure the robustness of potentials. It also has other uses in optimization schemes, such as speeding up the simulations required and identifying what type of lattice

is forming.

The trend optimization scheme presented here is an efficient method for finding the best potential parameters in a robust way. We compared it to a simulated annealing procedure and found it to be at least ten times faster, depending on which objective function is used. Trend optimization is ideal for objective functions that are very noisy, highly variable, and expensive to evaluate, but have a simple trend. The development of a way to characterize the quality of particle configurations has enabled the optimization of potentials based on quality.

Phase diagrams that involve a phase transition from a disordered to an ordered state can be constructed using lattice quality metrics. Krypton atoms on a graphite substrate are an example of such a system, and the phase transition curve for this system was found by using the Defect Measure. A different formulation of the Defect Measure is needed for the low temperature and high temperature regimes because of the densities at which phase transitions occur. The results are in good agreement with both molecular dynamics simulations using many more particles and experimental results.

The Voronoi metric, by looking at local properties, can be used to find global defects. This notion of using local properties to find global features allows the weaker regions of a lattice to be determined. If the lattice represented a material, it is along these weaker regions that deformations due to the effects of a stress on the lattice as a whole would occur. Thus, it is useful in predicting the behaviour of materials.

Similar ideas find an application in a very different system in the area of vehicular surveillance. Small, unmanned vehicles, with a limited range of communication, may be given the task of searching an area. Depending on the size and shape of the area to be searched, the vehicles flying in formation may be the best way to conduct the search. If vehicles can only communicate with their nearest neighbors, it is important for a central controller to know where there are weaknesses in the communication links. These correspond to areas of global defects in the lattice formation. Identifying these areas of weakness is the same problem that was solved using the Voronoi metric to find global defects by examining local properties.

The local perspective advocated here yields a toolkit of lattice quality metrics that can be used to further understand the structure of planar ordered systems. These can be either computational or experimental in nature. Although the applications discussed here

have been chemical systems, the metrics developed are, in principle, applicable to biological systems as well.

The self-assembly of nano-particles promises to yield a new generation of materials with novel applications (see [43, 13] for theoretical and experimental examples, respectively). The quality metrics developed here have a role to play in studying the behaviour of such materials that have not formed perfectly. Finding global defects by examining local properties can help to understand the structural properties of these materials. Optimization of features of the nano-particles with respect to quality of the resultant materials in simulations can lead to new materials and further understanding of the dynamics of their formation.

Similar ideas are also applicable in the area of self-assembling proteins in biological systems [65]. Computational models exist for the formation of structures such as the filaments formed from silicatines.¹⁷ Investigating how the properties of the filaments depend on the quality of the intermediate fractal sheet that self-assembles is expected to be a fruitful area of research. Quality metrics for constituents that are not simple particles can be developed based on the ideas presented here.

The analysis of real experimental systems with respect to local pattern quality is becoming a reality with the advent of ways to visualise surface monolayers using ultrafast electron crystallography [76]. The combination of experimental tests of structural properties once the actual pattern features and local quality is known can lead to significant improvements in computational models and hence the design of self-assembling systems.

¹⁷Thanks to Gunjan Thakur at UCSB for suggesting this application.

Chapter 6

Conclusions and Future Directions

In this thesis we have explored ways to overcome the difficulty of the large number of dimensions in computational models of molecular systems. This requires both a low-dimensional representation of the system that captures the essential dynamics of transition and efficient computational methods.

The first half of the thesis emphasised the importance of exploiting knowledge of the reaction dynamics in small molecular systems to produce reduced-order models of the transition. Dynamical systems theory reveals that invariant manifold tubes mediate transport through the transition state in the ionization of a Rydberg atom. The intersections of these tubes with a Poincaré surface of section yield a low-dimensional model from which transition rates can be calculated.

A coarse-level description of a chemical system requires an appropriate coarse variable that allows for a reduction of the number of dimensions in the model and captures the essential dynamics. The gyration radii of a system are such coarse variables. The internal centrifugal force, which couples the kinematics of the bath particles to the gyration radii, can be as important as the potential force between particles in chemical reactions. The competition between these two forces yield an effective potential that provides a simple model of the system.

The low-dimensional models of transition found by using knowledge of these two mechanisms of reaction were shown to accurately capture the system's transition behaviour. In the case of the Rydberg atom, we are confident in our results as they agree with those obtained by an independent method. In the methane and oxygen dissociation example the model's effective potential explains the observed distributions.

An efficient computational tool is also necessary for the calculation of transition prob-

abilities. The set-oriented method, as implemented in the GAIO software package, is such a tool. It can efficiently find global structures while maintaining accuracy by local refinement in the interesting regions of phase space. An exciting future direction is combining the coarse gyration radii description of a system with the set-oriented method to calculate reaction rates. The isomerization of an Argon-6 cluster is a well-studied system [88] and is a good test-bed for this combined approach. The study of transitions in larger molecular systems will become tractable by using the gyration radii as coarse variables to describe the system and computing reaction rates via the set-oriented method.

In the second half of the thesis, we developed quality metrics that quantify the ordered patterns in lattices and enables lattices to be compared with respect to quality via a single number. The metrics were shown to focus on different aspects of lattices: the alignment, the local density, and the inter-particle distances. We discussed situations in which each metric would be applicable, as well as the versatility of the Defect Measure. An important feature of these metrics is that they provide information about a global attribute based on local information.

We investigated various applications of these quality assessment tools: their use in the generation of potentials for the self-assembly of particles, finding the phase diagram of krypton atoms on a graphite substrate, and identifying global defects and clusters. Other possible applications are in computational models that seek to understand collisions at the atomic level, and with further development, in models of self-assembling proteins in biological systems.

The efficient computational tool that, when combined with the ability to represent the quality of a lattice with one number, enabled us to find optimal potentials for the self-assembly of particles into a lattice was the trend optimization scheme presented in Section 5.1.4. This powerful method, which is ideal for the noisy, highly variable, expensive-to-evaluate objective functions that we have here, was shown to perform two orders of magnitude better than a standard simulated annealing procedure. The strength of this scheme, and the feature that leads to robust solutions, is the upper and lower envelopes that characterize the variability in the objective function. These envelopes identify regions of the domain that may be excluded from the refined search area and provide a stopping criterion for the algorithm. Difficult optimization problems that have very noisy (or discontinuous),

highly variable objective functions that are expensive to evaluate, yet exhibit a smooth trend, can easily be handled by this scheme.

Appendix A

Lattices Used for Comparison in Section 4.3

The test lattices in Section 4.3 that were used for comparing the quality assessment metrics developed in that chapter, were generated from two different types of potentials: Lattices 1–13 were generated using a Rechtsman-style potential [75] and lattices 14–20 were generated using a polynomial potential (as explained in the baseline geometric method of Section 5.1.1.1). These potentials are explained in Section 5.1.

The differences between the lattices are due to differences in the parameter a_2 of the Rechtsman potential, the initial density, the time allowed for cooling, or the initial conditions. The values used are shown in Table A. If two lattices appear to have the same parameters, density and cooling time, then they differ only the initial conditions of the system.

The lattices may be grouped according to {1, 2}, {3}, {4, 5}, {6–9}, {10}, {11}, {12}, {13}, {14–16}, {17–19}, {20}.

Lattice index	Type of potential	Parameter a_2	Initial Density	Cooling Time
1	R	2.49	0.5977	2.5
2	R	2.49	0.5977	3.0
3	R	1.2	0.6897	3.0
4	R	1.49	0.6897	3.0
5	R	1.49	0.6897	3.0
6	R	2.49	0.6897	2.95
7	R	2.49	0.6897	3.0
8	R	2.49	0.6897	3.0
9	R	2.49	0.6897	3.0
10	R	2.49	0.8234	3.0
11	R	1.48	0.6897	3.0
12	R	1.52	0.6897	3.0
13	R	2.49	0.6903	3.0
14	P	–	0.7235	3.0
15	P	–	0.7235	3.0
16	P	–	0.7235	3.0
17	P	–	0.7698	3.0
18	P	–	0.7698	3.0
19	P	–	0.7698	3.0
20	P	–	0.7854	3.0

Table A.1: Differences between test lattices used for comparing metrics in Section 4.3. R denotes a Rechtsman-style potential and P denotes a polynomial potential. The cooling time is in units of 10^5 time steps.

Bibliography

- [1] Ahn, J., D. N. Hutchinson, C. Rangan, and P. H. Bucksbaum [2001], Quantum Phase Retrieval of a Rydberg Wave Packet Using a Half-cycle Pulse, *Phys. Rev. Lett.* **86**, 1179.
- [2] Alexandrow, N. and M. Hussaini, editors [1997], *Multidisciplinary Design Optimization*. SIAM, Philadelphia.
- [3] Ashcroft, N. W. and N. D. Mermin [1976], *Solid State Physics*. Holt, Rinehart and Winston, New York.
- [4] Audet, C., V. Béchar, and S. L. Digabel [2006], *Nonsmooth Optimization Through Mesh Adaptive Direct Search and Variable Neighborhood Search*, Technical Report G-2006-61, Les Cahiers du GERAD, École Polytechnique de Montréal, Canada.
- [5] Audet, C. and J. E. Dennis Jr. [2003], Analysis of Generalized Pattern Searches, *SIAM Journal on Optimization* **13**, 889.
- [6] Audet, C. and J. E. Dennis Jr. [2004], *Mesh Adaptive Direct Search Algorithms for Constrained Optimization*, Technical Report G-2004-04, Les Cahiers du GERAD, École Polytechnique de Montréal, Canada.
- [7] Beyer, H.-G. and B. Sendhoff [2007], Robust Optimization—A Comprehensive Survey, *Computer Methods in Applied Mechanics and Engineering* **196**, 3190.
- [8] Boje, E. [2000], Finding Nonconvex Hulls of QFT Templates, *J. Dynamic Systems, Measurement and Control* **122**, 230.
- [9] Booker, A. J., J. E. Dennis Jr., P. D. Frank, D. B. Serafini, V. Torzon, and M. W. Trosset [1999], A Rigorous Framework for Optimization of Expensive Functions by Surrogates, *Structural Optimization* **17**, 1.

- [10] Branscome, C. and D. Mavris [2004], A Regression Confidence Band Approach to Global Optimization. In *10th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference*, number AIAA 2004-4463, Albany, New York.
- [11] Butler, D. M., J. A. Litzinger, and G. A. Stewart [1980], Completion of the Phase Diagram for the Monolayer Regime of the Krypton-Graphite Adsorption System, *Phys. Rev. Lett.* **44**, 466.
- [12] Chapuisat, X. [1992], Principal-axis hyperspherical description of N -particle systems: Quantum-mechanical treatment, *Phys. Rev. A* **45**, 4277.
- [13] Chen, B., M. Eddaoudi, S. T. Hyde, M. O’Keeffe, and O. M. Yaghi [2001], Interwoven Metal-Organic Framework on a Periodic Minimal Surface with Extra-Large Pores, *Science* **291**, 1021.
- [14] Conley, C. C. [1968], Low Energy Transit Orbits in the Restricted Three-body Problem, *SIAM J. Appl. Math.* **16**, 732–746.
- [15] Contopoulos, G. and K. Efsthathiou [2004], Escapes and Recurrence in a Simple Hamiltonian System, *Celestial Mechanics and Dynamical Astronomy* **88**, 163–183.
- [16] Costiner, S., P. du Toit, K. A. Grubits, and J. E. Marsden [2008], Fast Generation of Potentials for the Self Assembly of Particles, *in preparation*.
- [17] De-Hua, W. and L. Sheng-Lu [2004], Semiclassical Calculation of the Recurrence Spectra of He Rydberg Atom in Perpendicular Electric and Magnetic Fields, *Chinese Physics* **13**, 464.
- [18] De Leon, N. [1992], Cylindrical Manifolds and Reactive Island Kinetic Theory in the Time Domain, *J. Chem. Phys.* **96**, 285.
- [19] De Leon, N., M. A. Mehta, and R. Q. Topper [1991], Cylindrical Manifolds in Phase Space As Mediators of Chemical Reaction Dynamics and Kinetics I. Theory, *J. Chem. Phys.* **94**, 8310.
- [20] De Leon, N., M. A. Mehta, and R. Q. Topper [1991], Cylindrical Manifolds in Phase Space As Mediators of Chemical Reaction Dynamics and Kinetics II. Numerical con-

- siderations and applications to models with two degrees of freedom, *J. Chem. Phys.* **94**, 8329–8341.
- [21] Delaunay, B. [1934], Sur la sphère vide, *Izvestia Akademii Nauk SSSR, Otdelenie Matematicheskikh i Estestvennykh Nauk* **7**, 793.
- [22] Dellnitz, M., G. Froyland, and O. Junge [2001], The Algorithms Behind GAIO—Set Oriented Numerical Methods for Dynamical Systems. In Fiedler, B., editor, *Ergodic Theory, Analysis, and Efficient Simulation of Dynamical Systems*, pages 145–174. Springer.
- [23] Dellnitz, M., K. Grubits, J. E. Marsden, K. Padberg, and B. Thiere [2005], Set-oriented Computation of Transport Rates in 3-degree of Freedom Systems: The Rydberg Atom in Crossed Fields, *Regular and Chaotic Dynamics* **10**, 173.
- [24] Dellnitz, M. and A. Hohmann [1996], The Computation of Unstable Manifolds Using Subdivision and Continuation. In Broer, H. W., S. A. van Gils, I. Hoveijn, and F. Takens, editors, *Nonlinear Dynamical Systems and Chaos*, pages 449–459. Birkhäuser.
- [25] Dellnitz, M. and A. Hohmann [1997], A Subdivision Algorithm for the Computation of Unstable Manifolds and Global Attractors, *Numerische Mathematik* **75**, 293.
- [26] Dellnitz, M. and O. Junge [2002], Set Oriented Numerical Methods for Dynamical Systems. In Fiedler, B., G. Iooss, and N. Kopell, editors, *Handbook of Dynamical Systems II: Towards Applications*, pages 221–264. World Scientific.
- [27] Dellnitz, M., O. Junge, W.-S. Koon, F. Lekien, M. W. Lo, J. E. Marsden, K. Padberg, R. Preis, S. D. Ross, and B. Thiere [2005], Transport in Dynamical Astronomy and Multibody Problems, *International Journal of Bifurcation and Chaos* **15**, 699.
- [28] Eckart, C. [1934], The Kinetic Energy of Polyatomic Molecules, *Phys. Rev.* **46**, 383.
- [29] Federov, M. V. and M. Y. Ivanov [1990], Coherence and Interference in a Rydberg Atom in a Strong Laser Field: Excitation, Ionization, and Emission of Light, *J. Opt. Soc. Am. B.* **7**, 569.
- [30] Frenkel, D. and B. Smit [1996], *Understanding Molecular Simulations: From Algorithms to Applications*. Academic Press.

- [31] Gabern, F., W. S. Koon, J. E. Marsden, and S. D. Ross [2005], Theory and Computation of Non-RRKM Lifetime Distributions and Rates in Chemical Systems with Three Or More Degrees of Freedom, *Physica D* **211**, 391.
- [32] Gelb, L. D. and E. A. Müller [2002], Location of Phase Equilibria by Temperature-quench Molecular Dynamics Simulations, *Fluid Phase Equilibria* **203**, 1.
- [33] Gilbert, R. G. and S. C. Smith [1990], *Theory of Unimolecular and Recombination Reactions*. Blackwell Scientific Publications, Oxford.
- [34] Grubits, K. A. and J. E. Marsden [2008], Lattice Quality Assessment Tools and Their Applications, *in preparation*.
- [35] Guckenheimer, J. and P. Holmes [1983], *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*. Springer Verlag, New York.
- [36] Guichardet, A. [1984], On Rotation and Vibration Motions of Molecules, *Ann. Inst. H. Poincaré* **40**, 329.
- [37] Gutmann, H.-M. [2001], A Radial Basis Function Method for Global Optimization, *Journal of Global Optimization* **19**, 201.
- [38] Hastie, T., R. Tibshirani, and J. H. Friedman [2001], *The Elements of Statistical Learning*. Springer.
- [39] Held, H., J. Schlichter, G. Raithel, and H. Walther [1998], Observation of Level Statistics and Heisenberg-time Orbits in Diamagnetic Rydberg Spectra, *Europhysics Letters* **43**, 392.
- [40] Hooke, R. and T. A. Jeeves [1961], Direct Search Solution of Numerical and Statistical Problems, *J. Assoc. Comp. Mach.* **8**, 212.
- [41] Horst, R., P. M. Pardalos, and N. V. Thoai [2000], *Introduction to Global Optimization*. Kluwer Academic Publishers, 2nd edition.
- [42] Huyer, W. and A. Neumaier [1999], Global Optimization by Multilevel Coordinate Search, *Journal of Global Optimization* **14**, 331.

- [43] Hyun, S. and S. Torquato [2001], Designing Composite Microstructures with Targeted Properties, *J. Mater. Res.* **16**, 280.
- [44] Jaffé, C., D. Farrelly, and T. Uzer [1999], Transition State in Atomic Physics, *Physical Review A* **60**, 3833.
- [45] Jagla, E. A. [1999], Minimum Energy Configurations of Repelling Particles in Two Dimensions, *J. Chem. Phys.* **110**, 451.
- [46] Jones, D. R., C. D. Perttunen, and B. E. Stuckman [1993], Lipschitzian Optimization Without the Lipschitz Constant, *Journal of Optimization Theory and Application* **79**, 157.
- [47] Jones, D. R., M. Schonlau, and W. J. Welch [1998], Efficient Global Optimization of Expensive Black-box Functions, *Journal of Global Optimization* **13**, 455.
- [48] Junge, O. [1999], *Mengenorientierte Methoden Zur Numerischen Analyse Dynamischer Systeme*, PhD thesis, University of Paderborn.
- [49] Junge, O., J. E. Marsden, and I. Mezić [2004], Uncertainty in the Dynamics of Conservative Maps, *Proc. CDC* **43**, 2225–2230.
- [50] Kittel, C. [2005], *Introduction to Solid State Physics*. John Wiley & Sons, Inc., New York, 8th edition.
- [51] Koon, W. S., M. W. Lo, J. E. Marsden, and S. D. Ross [2000], Heteroclinic Connections Between Periodic Orbits and Resonance Transitions in Celestial Mechanics, *Chaos* **10**, 427.
- [52] Kuppermann, A. [1997], Reactive Scattering with Row-Orthonormal Hyperspherical Coordinates. 2. Transformation Properties and Hamiltonian for Tetraatomic Systems, *J. Phys. Chem.* **101**, 6368.
- [53] Lahrer, Y. [1974], Triple Point of the First Monomolecular Layer of Krypton Adsorbed on the Cleavage Face of Graphite. Influence of the Structure of this Surface on the Structure of the First Adsorbed Layer of Rare Gases, *Journal of the Chemical Society, Faraday Transactions 1: Physical Chemistry in Condensed Phases* **70**, 320.

- [54] Littlejohn, R. G. and M. Reinsch [1997], Gauge Fields in the Separation of Rotations and Internal Motions in the n -body Problem, *Rev. Mod. Phys.* **69**, 213.
- [55] Lobaskin, V., M. Brunner, C. Bechinger, and H. H. von Grünberg [2003], On the Nature of Long-range Contributions to Pair Interactions Between Charged Colloids in Two Dimensions, *J. Phys.: Condens. Matter* **15**, 6693.
- [56] Lyubartsev, A. P. and A. Laaksonen [1995], Calculation of Effective Interaction Potentials from Radial Distribution Functions: A Reverse Monte Carlo Approach, *Phys. Rev. E* **52**, 3730.
- [57] Marmet, L., H. Held, G. Raithel, J. A. Yeazell, and H. Walther [1994], Observation of Quasi-Landau Wave Packets, *Phys. Rev. Lett.* **72**, 3779.
- [58] Marsden, A. L., M. Wang, and J. E. Dennis Jr. [2003], *Annual Research Briefs*, chapter Constrained Acoustic Shape Optimization using the Surrogate Management Framework, page 399. NASA Center for Turbulence Research.
- [59] Marsden, J. E., R. Montgomery, and T. S. Ratiu [1984], Gauged Lie-Poisson Structures, *Contemp. Math.* **28**, 101.
- [60] Marsden, J. E. and T. S. Ratiu [1999], *Introduction to Mechanics and Symmetry*. Springer, New York.
- [61] Marsden, J. E., T. S. Ratiu, and J. Scheurle [2000], Reduction Theory and the Lagrange-Routh equations, *J. Math. Phys.* **41**, 3379.
- [62] Marston, C. C. and N. De Leon [1989], Reactive Islands As Essential Mediators of Unimolecular Conformational Isomerization: A Dynamical Study of 3-phospholene, *J. Chem. Phys.* **91**, 3392–3404.
- [63] McGehee, R. [1969], *Some Homoclinic Orbits for the Restricted Three-body Problem*, PhD thesis, University of Wisconsin, Madison.
- [64] Mermin, N. D. [1979], The Topological Theory of Defects in Ordered Media, *Rev. Mod. Phys.* **51**, 591.
- [65] Mershin, A., B. Cook, L. Kaiser, and S. Zhang [2005], A Classic Assembly of Nanobio-materials, *Nature Biotechnology* **23**, 1379.

- [66] Mezić, I. and T. Runolfsson [2004], Uncertainty Analysis of Complex Dynamical Systems, *Proc. Amer. Control Conference 2004* **3**, 2659.
- [67] Mezić, I. and S. Wiggins [1999], A Method for Visualization of Invariant Sets of Dynamical Systems Based on the Ergodic Partition, *Chaos* **9**, 213–218.
- [68] Nadler, B., S. Lafon, R. C. Coifman, and I. G. Kevrekidis [2006], Diffusion Maps, Spectral Clustering and the Reaction Coordinates of Dynamical Systems, *Appl. Comp. Harm. Anal.* **21**, 113.
- [69] Neumaier, A. [2004], Complete Search in Continuous Global Optimization and Constraint Satisfaction. In Iserles, A., editor, *Acta Numerica 2004*, pages 271–369. Cambridge University Press.
- [70] Padberg, K. [2005], *Numerical Analysis of Transport in Dynamical Systems*, PhD thesis, University of Paderborn.
- [71] Plimpton, S. J. [1995], Fast Parallel Algorithms for Short-range Molecular Dynamics, *J. Comp. Phys.* **117**, 1.
- [72] Raithel, G., M. Fauth, and H. Walther [1991], Quasi-Landau Resonances in the Spectra of Rubidium Rydberg Atoms in Crossed Electric and Magnetic Fields, *Phys. Rev. A* **44**, 1898.
- [73] Raithel, G. and H. Walther [1994], Ionization Energy of Rubidium Rydberg Atoms in Strong Crossed Electric and Magnetic Fields, *Phys. Rev. A* **49**, 1646.
- [74] Rechtsman, M., F. Stillinger, and S. Torquato [2005], Optimized Interactions for Targeted Self-assembly: Application To a Honeycomb Lattice, *Phys. Rev. Lett.* **95**, 228301.
- [75] Rechtsman, M., F. Stillinger, and S. Torquato [2006], Designed Interaction Potentials Via Inverse Methods for Self-assembly, *Phys. Rev. E* **73**, 011406.
- [76] Ruan, C.-Y., D.-S. Yang, and A. H. Zewail [2004], An Iterative Scheme to Derive Pair Potentials from Structure Factors and Its Application to Liquid Mercury, *J. Am. Chem. Soc.* **126**, 12797.
- [77] Salsburg, Z. W., R. W. Zwanzig, and J. G. Kirkwood [1953], Molecular Distribution Functions in a One-dimensional Fluid, *J. Chem. Phys* **21**, 1098.

- [78] Sasena, M. J., P. Papalambros, and P. Goovaerts [2002], Global Optimization of Problems with Disconnected Feasible Regions via Surrogate Modelling. In *9th AIAA/ISSMO Symposium on Multidisciplinary Analysis and Optimization*, number AIAA-2002-5573, Atlanta, Georgia.
- [79] Tachibana, A. and T. Iwai [1986], Complete Molecular Hamiltonian Based on the Born-Oppenheimer Adiabatic Approximation, *Phys. Rev. A* **33**, 2262.
- [80] Theil, F. [2006], A Proof of Crystallization in Two Dimensions, *Comm. Math. Phys* **262**, 209.
- [81] Torczon, V. [1995], Pattern Search Methods for Nonlinear Optimization, *SIAG/OPT Views and News* **6**, 7.
- [82] Torczon, V. and M. W. Trosset [1998], From Evolutionary Operation to Parallel Direct Search: Pattern Search Algorithms for Numerical Optimization, *Computing Science and Statistics* **29**, 396.
- [83] Toth, G. [2003], An Iterative Scheme to Derive Pair Potentials from Structure Factors and Its Application to Liquid Mercury, *J. Chem. Phys.* **118**, 3949.
- [84] Truhlar, D. G., B. C. Garrett, and S. J. Klippenstein [1996], Current Status of Transition-state Theory, *J. Phys. Chem.* **100**, 2771.
- [85] Uzer, T., C. Jaffé, J. Palacián, P. Yanguas, and S. Wiggins [2002], The Geometry of Reaction Dynamics, *Nonlinearity* **15**, 957.
- [86] Voronoi, G. [1907], Nouvelles applications des paramètres continus à la théorie des formes quadratiques, *Journal für die Reine und Angewandte Mathematik* **133**, 97.
- [87] Wojcik, A. and R. Parzynski [1995], Dark-State Effect in Rydberg-Atom Stabilization, *J. Opt. Soc. Am. B.* **12**, 369.
- [88] Yanao, T., W.-S. Koon, J. E. Marsden, and I. G. Kevrekidis [2007], Gyration-Radius Dynamics in Structural Transitions of Atomic Clusters, *J. Chem. Phys.* **126**, 1.
- [89] Yanao, T. and K. Takatsuka [2003], Collective Coordinates and an Accompanying Metric Force in Structural Isomerization Dynamics of Molecules, *Phys. Rev. A* **68**, 032714.

- [90] Yanao, T. and K. Takatsuka [2004], Kinematic Effects Associated with Molecular Frames in Structural Isomerization Dynamics of Clusters, *J. Chem. Phys.* **120**, 8924–8936.