Quantum Drude Oscillators for Accurate Many-body Intermolecular Forces

Andrew Jones

A thesis submitted in fulfilment of the requirements
for the degree of Doctor of Philosophy
to the
University of Edinburgh
July 2010

Abstract

One of the important early applications of Quantum Mechanics was to explain the Van-der-Waal's $1/R^6$ potential that is observed experimentally between two neutral species, such as noble gas atoms, in terms of correlated uncertainty between interacting dipoles, an effect that does not occur in the classical limit [London-Eisenschitz,1930]. When many-body correlations and higher-multipole interactions are taken into account they yield additional many-body and higher-multipole dispersion terms.

Dispersion energies are closely related to electrostatic interactions and polarisation [Hirschfelder-Curtiss-Bird,1954]. Hydrogen bonding, the dominant force in water, is an example of an electrostatic effect, which is also strongly modified by polarisation effects. The behaviour of ions is also strongly influenced by polarisation. Where hydrogen bonding is disrupted, dispersion tends to act as a more constant cohesive force. It is the only attractive force that exists between hydrophobes, for example. Thus all three are important for understanding the detailed behaviour of water, and effects that happen in water, such as the solvation of ions, hydrophobic de-wetting, and thus biological nano-structures.

Current molecular simulation methods rarely go beyond pair-wise potentials, and so lose the rich detail of many-body polarisation and dispersion that would permit a force field to be transferable between different environments. Empirical force-fields fitted in the gas phase, which is dominated by two-body interactions, generally do not perform well in the condensed (many-body) phases. The leading omitted dispersion term is the Axilrod-Teller-Muto 3-body potential, which does not feature in standard biophysical force-fields. Polarization is also usually ommitted, but it is sometimes included in next-generation force-fields following seminal work by Cochran [1971]. In practice, many-body forces are approximated using two-body potentials fitted to reflect bulk behaviour, but these are not transferable because they do not reproduce detailed behaviour well, resulting in spurious results near inhomogeneities, such as solvated hydrophobes and ions, surfaces and interfaces.

The Quantum Drude Oscillator model (QDO) unifies many-body, multipole polarisation and dispersion, intrinsically treating them on an equal footing, potentially

leading to simpler, more accurate, and more transferable force fields when it is applied in molecular simulations. The Drude Oscillator is simply a model atom wherein a single pseudoelectron is bound harmonically to a single pseudonucleus, that interacts via damped coulomb interactions [Drude,1900].

Path Integral [Feynman-Hibbs,1965] Molecular Dynamics (PIMD) can, in principle, provide an exact treatment for moving molecules at finite temperature on the Born-Oppenheimer surface due to their pseudo-electrons. PIMD can be applied to large systems, as it scales like $N\log(N)$, with multiplicative prefactor P that can be effectively parallelized away on modern supercomputers. There are other ways to treat dispersion, but all are computationally intensive and cannot be applied to large systems. These include, for example, Density Functional Theory provides an existence proof that a functional exists to include dispersion, but we dont know the functional. We outline the existing methods, and then present new density matrices to improve the discretisation of the path integral.

Diffusion Monte Carlo (DMC), first proposed by Fermi, allows the fast computation of high-accuracy energies for static nuclear configurations, making it a useful method for model development, such as fitting repulsion potentials, but there is no straightforward way to generate forces. We derived new methods and trial wavefunctions for DMC, allowing the computation of energies for much larger systems to high accuracy.

A Quantum Drude model of Xenon, fit in the gas-phase, was simulated in the condensed-phase using both DMC and PIMD. The new DMC methods allowed for calculation of the bulk modulus and lattice constant of FCC-solid Xenon. Both were in excellent agreement with experiment even though this model was fitted in the gas-phase, demonstrating the power of Quantum Drudes to build transferable models by capturing many-body effects. We also used the Xenon model to test the new PIMD methods.

Finally, we present the outline of a new QDO model of water, including QDO parameters fitted to the polarisabilities and dispersion coefficients of water.

Declaration

Except where otherwise stated, the research undertaken in this thesis was the unaided work of the author. Where the work was done in collaboration with others, a significant contribution was made by the author.

Andrew P. Jones
July 2009

Acknowledgements

I would like to record my thanks to all those who have helped me in the course of this work. In particular, the following:

Prof. Glenn J. Martyna

Prof. Jason Crain

Prof. Martin Müser

Dr. Troy Whitfield

Contents

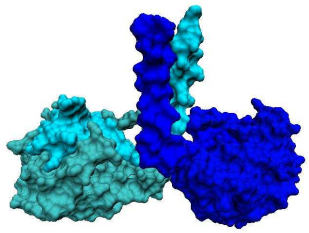
List of Figures

List of Tables

Chapter 1

Introduction

Biology at the molecular level is a world of nano-machines and nano-structures, including such wonders as kinesin the walking molecule [Asb05], the flagellum [JA91] and ATP-synthase rotary motors [SLW99], the myosin ratchet motor [CEO92]; bi-layer membranes, liposomes, vesicles, gates, ion-pumps and organelles; DNA, RNAs, and the associated proteins that open, copy, cut, repair and otherwise operate on the encoded information; ribosomes and chaperonins that assemble new machines; proteases and lysosomes that recycle old ones [Goo09, TT99]. All of these are built out of organic molecules (plus a handful of other atom types), but the mechanisms of molecular biology involve much more than just organic chemistry. For example, when amino acids are joined into proteins by chemical reactions, that is only the first step in a multi-level process of folding into its secondary and tertiary (and even quaternary) structure [BT91]. Folding mostly proceeds with no reactions [BT91]. Likewise the chemical reaction of ATP with a molecular machine introduces internal tensions to the machine, but understanding how it changes shape, or does work against the environment, requires additional physics [SD10].



Molecular Machine: 'feet' of the kinesin walking motor (3KIN.pdb)

To grasp deeply the mechanisms of biology at the molecular level, we need accurate working models of the systems of interest. Simple macroscale models, using measurable bulk/continuum properties of materials, are often not applicable to the nanoscale/molecular level, where the relevant physics can be different. For example, inertia is negligible for molecular machines trying to move through water (very low Reynolds number) [Goo04]. Since many biological processes are more efficient than

human technology [HB05], it makes sense for us to learn from these systems about the physics they use [Goo04, Vin09]. In structural molecular biology, it is a priori difficult to predict either structure or the physics of function, from simple inputs such as sequence or atomic coordinates, respectively [BT91, ch. 16]. Without detailed insight into their physics, many biomolecular structures would remain black boxes, and thus our biological design methods would be limited to trial-and-error, similar to the effects of Darwinian processes [BBL08, GEFS10, PISW49, BNR67, KSS+96].

Modern drug-discovery methods are focussed on screening for ligands [ES90, AAA+10, SD03], and/or rationally designing them [BBL+95, BGBS05, Tia10], so as to bind and interfere with a particular target, but they often also have unpleasant side-effects when they accidentally target other molecules [MW05]. Also, if they act without taking heed of the design or normal operation of the system, for instance crudely switching a particular target completely on/off, they can be vulnerable to, or even cause, secondary problems [MW05]. The former problem suggests that creating specificity in biomolecular engineering is not just about learning how to strengthen a particular interaction or binding, but also understanding the principles by which unwanted interactions can be suppressed; as the number of desired interactions is usually a small subset of all possible interactions, the latter is perhaps even more important [BGBS05]. For example, the behaviour of the familiar globular proteins is strongly atypical of the behaviour of long polypeptide chains in general [Axe04]; the 'natural' behaviour of proteins, if they are significantly mutated without selection, or physically misshapen by heat, is to unfold and stick to other proteins in a mesh of fibrils [MD00, SD03], as in a fried egg. Only after suppressing this 'natural', generic behaviour of proteins do they become the globular and free-floating modules that can be programmed with highly specific tasks.

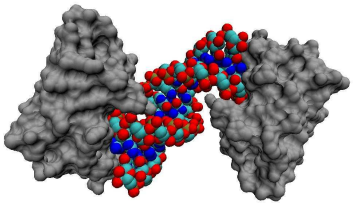
Getting out of the 'black-box' regime into real engineering, to be able to rationally design new systems, or augment existing ones, requires models; we need to understand in detail what the molecules are doing in the same way an engineer not only understands a bridge or aircraft by modelling its parts (mechanical engineering → biomolecular engineering), but also ensures that they work together to form a functional whole (systems engineering → systems biology).

If chemistry (reactions) provide the rivets that form permanent bonds between parts, and the energy to do work, it is remarkable just how much of the structure is self-assembled by physical forces. Many parts fall into place with a minimum of manipulation, according to preprogrammed patterns of *chemical complementarity* and

hydrophobicity [Goo09].

Examples of complementarity are the attraction between oppositely charged ions, or between oppositely polarised groups, which includes the especially strong hydrogen-bond [AdP02]. Most of the secondary structure of proteins is due to hydrogen-bonding (plus the occasional sulphur-sulphur covalent-bond 'rivet' [BT91]).

Hydrophobicity is caused by the lack of and thus disruption of charged and polar interactions, especially the disruption of hydrogen bonding in water. Because water molecules are particularly good at forming hydrogen bonds, water tends to segregate itself from large hydrophobic masses, causing them to 'de-wet' and subsequently form membranes (in the case of lipids) or fold into globular shapes (in the case of proteins) [LCW99, ZHMB04, BWZ09, AdP02]. Hydrophobicity is an extremely important feature of the design of biological nanostructures; if the stability of these structures were dominated by strong bonds buried deep inside them, that would be fine for its functional lifetime, but would make it very difficult to disassemble, recycle or digest [TG02]. Proteins, for example, are able to fold and unfold in response to small changes in pH. Lipid membranes are essentially fluid and can be shaped, joined and pinched-off by the cell[Goo09, TT99].



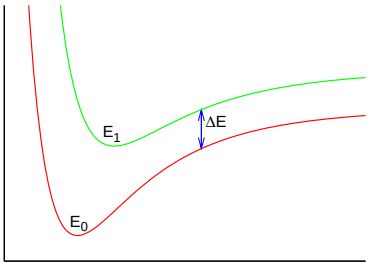
Helicase bound to double-stranded RNA fragment (3LRN.pdb)

The effects of complementary and hydrophobic attractions are strongly mediated by the shape of the molecules involved. If two molecules fit together as 'lock and key' or lego bricks, then the the attractive forces between their different parts can have a greater combined effect [Goo09, TT99]. In this way biological molecules can be programmed to bind with a very specific set of other molecules. Also shape can cause interesting things to happen by 'induced fit' [TT99]. This means one or both of the molecules changes shape in response to other, as a hand reshapes a glove when it enters. The deformation can be used to create a stress that results in a reaction of a substrate, or to change the shape of an enzyme so that another active site is activated or inactivated. The shape and squishiness of a molecule is described by its intermolecular repulsion potential. The balance of long-range attractions and short-range repulsion allows stable structures to be formed, that is, minima in the energy surface.

1.1 Computer Simulation of Condensed Matter on the Born-Oppenheimer Surface

Computer simulations of condensed matter provide an important bridge between analytical theory and experiment, the other two pillars of modern science[1]. Simulation allows us to 'see' what is going on at a microscopic level via a model[FS96]. Simulation can for example elucidate the microscopic origins of effects seen by experiment, and thus help in developing an analytical explanation for them, or in other cases simulation can demonstrate that a particular microscopic theory has implications that do not agree with experiment. It also allows us to explore 'alternate realities' (for example what would happen if certain kinds of forces did not exist), and other 'unphysical' tactics to gain detailed insight into the mechanisms of the real world [ZHMB04]. It also gave us new insight into the laws of nature as algorithms [FS96]. In practice, we cannot put precise fundamental physics into simulations. Instead we have to construct simplified model systems, and then test them to destruction to make sure that they reproduce all the most relevant phenomena of the real system they represent, with a minimum of ad-hoc or post-hoc corrections.

For the intermolecular effects mentioned above, there are no low-lying electronic excited states and therefore no 'surface-hopping' [TP71, PT05]; all can be treated within the Born-Oppenheimer approximation [BO27]. The Born-Oppenheimer surface is the ground state electronic energy solved as a function of the instantaneous nuclear configuration, upon which the nuclei move (We are not interested in excited state BO surfaces here). If the nuclei are treated classically, then they move under the influence of forces derived from said surface, using Newton's equation of motion $F = ma$. This is a very good approximation at room temperature, particularly if one is interested in physical as opposed to chemical processes.



B.O. limit if $kT \ll \Delta E$

Exact solution of the ground state Born-Oppenheimer surface is often not possible, analytically or numerically, for the large complex systems of interest, so approximations are required to model the BO surface accurately. We would like a method that is not as complicated as solving the full electronic structure calculation (which is prohibitively

[1] http://www.wtec.org/sbes/SBES-GlobalFinalReport.pdf

computationally intensive), but that captures the most important forces, and at a computational cost that scales linearly with atom number so that large-scale and long-time simulations can be performed.

1.2 Deconstructing the Born-Oppenheimer Energy Surface

Consider two molecules far apart. The Born-Oppenheimer energy and electron-density of two isolated molecules is a natural reference, which we then perturb to gain insight into how they interact. At long range, the forces between the molecules are predominantly caused by the elecrostatic force, modified by the electronic response of a molecule to the electrostatic interactions [HCB54, Buc67, BU70, Sto96]. This includes interactions between multipoles (a bare charge is the 0^{th}-multipole), but also polarisation, and dispersion, which arise from the flexibility of the charge distribution in a molecule as a stable, unreactive, typically closed-shell moietie, as we now discuss in detail below.

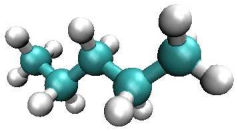
To a first approximation a molecule can be treated as a static charge distribution, which interacts with other molecules via electrostatic forces. This can be modelled by a multipole-expansion, or by a distributed set of point charges (see for instances the classic works of Hirschfelder-Curtiss-Bird [HCB54] and Stone [Sto96]).

Also, the charge distribution within a molecule can be 'stretched'/perturbed out of its equilibrium distribution, under the influence of Coulomb forces external to the molecule, which is *polarisation* [Buc67]. The shifted charge distribution of a molecule changes the pattern of Coulomb forces which it exerts on other molecules, leading to polarisation effects that can propagate over many molecules; polarisation is a many-body phenomenon[DO58, Coc71, Buc67, BU70].

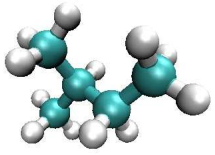
Even where there are no external Coulomb forces or permanent charges or polar molecules to cause polarisation in a collection of molecules, there are electronic *dispersion* forces [HCB54]. From here-on-in, we will refer to electronic dispersion simply as 'dispersion'. Dispersion is the dominant type of force between non-polar molecules and molecular residues, including noble gases (especially heavy atoms) and hydrocarbons (especially benzene-like groups, alkenes and alkanes) [AdP02]. As such it is crucial for understanding hydrophobicicity. Dispersion is sometimes described as being due to instantaneous fluctuations, but this is misleading; the mean (or

measurable) charge distribution shows no change over time [HCB54]. It is more correct to think of it as a purely quantum- mechanical coupling/correlation of the uncertainty of the charge distribution between molecules, caused by Coulomb interactions. The couplings occur because quantum uncertainty applies to the system as a whole; not just individual molecules or even pairs. The dominant terms are those between pairs of molecules, starting with the familiar $-C_6/r^6$ potential, which can be corrected by $-C_8/r^8$ and $-C_{10}/r^{10}$ terms, and then terms that involve triplets of molecules, such as the Axelrod-Teller-Muto $C_9/(r^3)^3$ potential [HCB54, Fon61, Sto96, AT43]. And so on. Note, for non-spherical molecules, dispersion will of course be at least somewhat anisotropic [Buc67, Sto96].

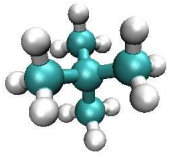
All molecules experience dispersion attractions, including both hydrophiles and hydrophobes [Sto96]. For hydrophobes, many-body dispersion is the dominant kind of attraction [AdP02]. For example dispersion accounts for the differences in boiling temperature between pentane, isopentane and neopentane, three hydrophobic molecules having the same weight. For hydrophils, multipolar interactions and many-body polarization are dominant but dispersion still accounts for a reasonable fraction of the energy. Dispersion tends to be much less directional, however. Therefore to understand complementarity and hydrophobicity requires a model that treats charge interactions, polarisation and dispersion well, and one that goes beyond simple pairwise interactions. Here, we consider systems in which modeling charge transfer [RI91] (on the BO surface) is not important and point the reader to interesting recent work [NM09, CM09] on the topic.



pentane



isopentane



neopentane

Consider the two molecules again. As they are moved together, the asymptotic long-range behaviour breaks down, and the electronic structure of the isolated molecules becomes a less reliable reference [SO96]. The repulsive interactions between molecules arise from increased nuclear repulsion, Pauli exclusion and in general complex changes in electronic structure. It has been found that repulsion can be decomposed, to a good approximation, into pairwise terms centered on the atoms of the molecules [Sto96].

Advanced modelling often uses non-isotropic repulsion, for example there is Madden's seminal work on ionic solids and liquids [MPL$^+$01, AM03b], which follows work by Stone and others [SP88, Sto96]. As described above, accurate repulsion is important as it is crucial to correctly define the size, shape and 'squishiness' of a molecule, and therefore the stable energetic minima of its interactions with other molecules.

1.3 Current Force Fields, and Limits on their Predictions

Most current biomolecular simulation methods use relatively simple two-body potentials that capture repulsion and two-body dispersion fairly well [PC03]. They are often also fitted to incorporate environmental effects in a 'mean-field' approach. The standard approach, the Lennard-Jones potential [LJ24], simply uses isotropic repulsion of the form r^{-12} and a dispersion attraction of the form r^{-6} with fixed coefficients, plus a fixed charge Coulomb interaction; for example the force-fields used by CHARMM, AMBER, GROMOS, [BBO$^+$83, CCB$^+$95, SHT$^+$99] for which this is the main approach, as was the classic work of Stillinger and Rahman on liquid water. This mean-field method works well for state points around which the parmeters were fit especially for a one-component system in bulk, without interfaces and other heterogeneities, or wherever else polarisation and other many-body effects can be treated as small perturbations to a mean-field model.

Explicit polarisation and many-body dispersion are usually omitted from standard simulations [PC03], mainly because they are costly in terms of man-hours and computer-hours to implement properly. Polarisation requires a model to describe a complex and mobile charge distribution in a compact way. This means including an expansion of variable multipoles in the description of a molecule, as well as correctly parameterising and calculating the many interactions between different terms on different molecules. Many-body terms, including both polarisation and dispersion terms, are complicated even just to write down, and utilising explicit n-body terms can increase computation costs from at most N^2 to N^n (where N is the number of interacting particles). Multipolar polarization can be reduced to an iterative N^2 method and implemented with $N \log N$ scaling, but many-body dispersion cannot be properly decomposed like this without approximations. However, there are some cases where polarisation and many-body dispersion become very significant, and change the behaviour of a system profoundly.

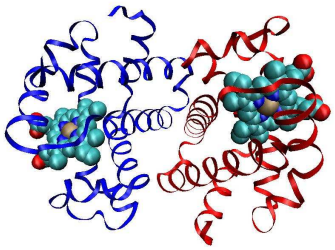
For example, the dipole-moment of water is roughly 1.85D in the gas phase, where each molecule is relatively isolated most of the time. However in condensed environments each molecule makes hydrogen bonds with other molecules resulting in strong polarisation effects. These polarisation effects propagate to other hydrogen bonds, leading to a total dipole-moment of roughly 2.6-3D [BF33, SR74, CIKP03, BXJ98] for water in such environments. The standard method for getting around this, is to use different models for water; one with a small dipole-moment, for the gas phase, and another with a large dipole-moment, for the condensed phase. This works pretty well for reproducing many of the properties of liquid, solid and gas phase water, but not all. The triple point of water has remained elusive, for example. Such mean-field models are also unable to respond to diverse environments where there is an intermediate amount of polarisation, such as at low density, near surfaces and hydrophobic masses, or even near other charged or polarisable groups, where the average dipole-moment would be different. As such, these models are not transferable.

Polarisation also influences the behaviour of ions in water, including the coordination number (the number of water molecules in the innermost shell of neighbours) [SKW90]. It also determines an ion's preference for the bulk or surface [HPDS05, KLF$^+$00].

An example of the importance of dispersion is the condensed-phase noble gases, for which experiments show that the condensed phase occupies a volume several percent larger than that predicted from the gas-phase two-body potential in simulation [MW70], due to many-body dispersion. This prompted the use of different models, one for gas-phase and another for the condensed phase, fitted to reproduce the total binding energy and the lattice constant. However, these condensed-phase models had other deficiencies. Barron and Domb [BD55] noticed that two-body dispersion alone predicts an HCP structure for heavy noble-gas atoms in the solid state, in contrast to the experimentally observed preference for FCC. This difference was attributed to *non-central* (that is, many-body, non-two-body) forces, showing that the many-body quality is somehow just as important as the magnitude of the correction. The models also predicted incorrect elastic properties, showing that many-body dispersion is also important for those as well, and cannot be reduced to effective two-body potentials. The surface tension is also inaccurate when many-body dispersion is omitted [LBP74, Bar93]. Because the underlying physics of many-body dispersion is ubiquitous, the same effects could influence many other cases, although it would be most marked in non-polar/hydrophobic contexts, where no other forces are significant.

In all these cases, the reason those non-trivial corrections are required, is that important physics is missing. But that physics is also likely to shape the detailed behaviour of the system, making it is difficult to determine whether the detailed behaviour is correct. When a mean-field model encounters a situation which is significantly different from the one in which it is fitted, we can expect that will need to be parameterised, and that many not be easy; we are forced either to investigate each detailed context empirically, or reintroduce the omitted physics. This is very relevant to biological systems, where varied contexts are the norm. This includes intermediate densities of fluids, and various anisotropic contexts such as interfaces.

Many important problems in molecular biology involve mixtures of hydrophilic and hydrophobic molecules, and an interplay of polarisation and dispersion interactions. In such cases, it is often very difficult to predict or explain the detailed behaviour, as found with methanol-water mixtures [DCP$^+$02]. One such important area which combines the need for accurate polarisation and dispersion in a strongly heterogeneous environment, is the complicated balance involved in hydrophobic de-wetting [LCW99]. Hydrophobic de-wetting is probably the most important factor in the folding and aggregation of proteins, and in the formation of micelles and membranes, so it is absolutely essential for biological situations that we can understand it and model it fairly accurately. Small hydrophobic molecules can happily coexist with the hydrogen-bond network that exists in water, but large hydrophobic molecules and large aggregates of hydrophobes tend to disrupt that network, leaving dangling OH-groups. This creates a free-energy cost for water molecules in contact with the hydrophobes, leading to drying/de-wetting [BWZ09]. De-wetting is determined by the fine balance between the surface tension in water, which is due to strong but directional hydrogen-bonding, which tends to pull water out of hydrophobic regions into regions where it can form more hydrogen bonds, the Van-Der-Waals forces for which the dominant terms are always attractive, and other interactions which tend to disrupt and slow the de-wetting process [BWZ09, ZHMB04]. The surface tension is modified by the degree of polarisation of the water molecule, and the detail of the Van-der-Waals forces will require many-body interactions and polarisation modifications. The problem of de-wetting really calls for a unified treatment of all these forces.



Haemoglobin folds, with
2 Haem groups (3EOK.pdb)

1.4 Advanced Force Fields

There exist a variety of approaches for including many-body polarisation and/or dispersion [SD76]. One method is to use classical point dipole oscillators, a model which captures many-body polarisation in the dipole limit, but not many-body dispersion, or many-body multipole polarisation (that is, no quadrupolar- octopolar-, or higher-, polarisation). These models have had great success in explaining water, as dipole-polarisation generates much of the observed adjustment to the dipole-moment in condensed water [SK88], as well as charge-screening, and determines the behaviour of ions in water [SKW90]. They also have had great success treating liquid ammonia [DMK94].

Advanced (post mean field) force field development began with Dick and Overhauser which was quickly follwed by Cochran's seminal work [DO58, Coc71]. Both used a spherical shell model for polarisable atoms, with a rigid shell of charge representing outer electrons, harmonically attached to an oppositely charged core, representing the nucleus and remaining electrons. This has the advantage that the spherically distributed charge prevents singularities in the interaction with other nuclei, and the overlap density prescribed a method of estimating the exchange interaction.

Recently, Lamoureux-Mackerell-Roux [LADMR03] have discovered that the gas phase dipole-polarisability of water is too large to explain condensed phase behaviour. They suggested that possibly the real electronic distribution is spatially constrained by the neighbouring molecules, restricting its polarisability; an effect which a point-dipole polarisability does not capture. A similar effect is discovered by Jungwirth and Tobias for the polarisability of Cl^- ions in an aqueous environment [JT02].

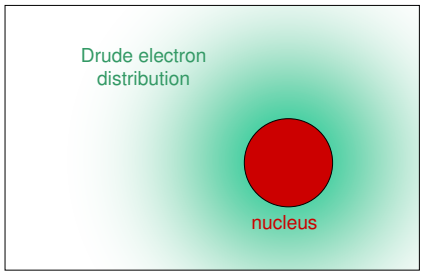
Batista, Xantheas, and Jónsson added more point-multipole polarisabilities to study some phases of ice. They found that quadrupole polarisability is also important [BXJ98] It has been shown that that quadrupole polarisations are necessary to properly understand the behaviour of the Silver (Ag) atom in condensed phase [WCCM96]. Thus it would be useful to have a general method that incorporates quadrupole and higher polarisation should they arise.

Important studies have been undertaken in which some of the 3-body terms are incorporated explicitly. Normally this would be limited to the Axilrod-Teller-Muto term [AT43]), a term involving triplets of fluctuating dipoles. However, not only is that term complicated, it is only the first term in an ever more complicated series of 3-body

and 4-body terms. Even two-body dispersion involves non-trivial terms beyond C_{10}. Some of the three-body terms have been calculated by Bell and implemented by Doran-Zucker [Bel70, DZ71], who showed that they help to explain the HCP/FCC anomaly in crystalline noble-gas condensates, but these are terms in an asymmpotic series that require non-trivial damping functions. Given the above discussion, many-body polarisation and dispersion are quite important to describe heterogeneous environments of interest today. Therefore, it is important to develop an accurate model along with methods to solve it that are computationally inexpensive.

1.5 The Quantum Drude Model as a potential High-accuracy Force-field

The Quantum Drude Oscillator is a model atom in which a pseudo-electron is bound to its pseudo-nucleus by a harmonic potential [HCB54, WM06]. A Quantum Drude interacts with other Quantum Drudes and other charged particles via Coulomb potentials, in a similar way to electrons in closed shells. As it is quantum, it is not a point particle, but has a spatial distribution of charge which can be shifted and thus polarised to all orders (dipole, quadrupole, and so on). Correlated quantum uncertainty produces dispersion effects between Quantum Drude atoms [EL30]. Since Drude electrons are distinguishable particles, exchange is neglected (a good approximation at long range) which allows for order N solutions as there is no Fermi-sign problem.



Quantum Drude atom

As such, the QDO model provides an excellent framework for modelling the intermolecular long- and intermediate-range interactions relevant to biomolecules. Electostatics, polarisation and dispersion, are addressed in a unified way that reflects the underlying force that cause all three kinds of interactions, and leads to simplified, unified, much more general models than currently exist.

Drude Oscillators were first used to explain the refraction of light, as a classical model [Dru00], and then in a quantum model with only minor changes [HCB54, pg.881], and also as a simple model for calculating pairwise London/van-der-Waals dispersion [HCB54, pg.956] [BK57]. Cao and Berne [CB92a] used a dipole-limit Quantum-Drude

model to study many-body dispersion and polarisation in rare-gas clusters, and showed that dispersion expands the volume of such clusters slightly.

Quantum Drude Oscillators approximate the behaviour of real atoms quite well, as we will show later (sec. 4.3). Therefore, we propose to use them to build models of those atoms. For this purpose, QDOs have 3 major advantages.

First, by explicitly keeping some of the quantum mechanical nature of the electronic structure, which allows many-body dispersion and polarisation to be treated on an equal footing, the Quantum Drude model simplifies the actual model building considerably. One need only fit the properties of individual species from polarity, polarisation and dispersion measurements in the gas phase, yet expect interactions between mixed species to be broadly accurate, without having to fit parameters for every possible interacting pair of species. One can expect to obtain not just the usual electrostatic interactions, but also accurate Van-der-Waals forces which come 'for free' (they do not need to be added explicitly). One can also expect the induced polarisations to be correctly modified by the surrounding environment without any further effort.

Second, one samples polarisation/dispersion accurately to all orders of many-body multi-polar interactions, that were previously difficult and computationally expensive to compute.

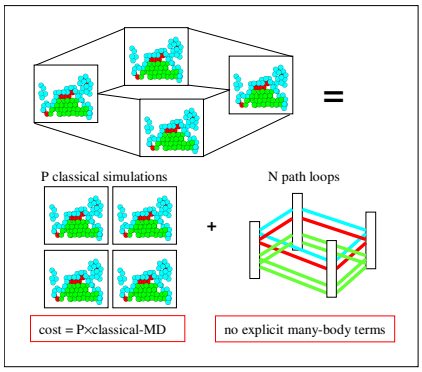
Third, short range repulsions can be treated as empirical two-body atom-centered terms, in a way that is consistent with the current state of the art.

1.6 Simulation of QDO's via DMC and PIMD

Diffusion Monte Carlo (DMC) [GS71, And75, KW86, HJR94, Cep95, UNR93] provides a means of quickly generating numerically exact energies for static configurations of QD nuclei and charges, which is useful for example when fitting parameters at the model-building stage. DMC represents the wavefunction by a set of walkers that evolve such that they sample the ground state wavefunction or the ground state wavefunction times a trial function. As such, it does not sample the true probability distribution $|\Psi_0|^2$, preventing any straighforward use of the Hellman-Feynman theorem, and so there is no straightforward way to calculate the forces on the nuclei. Therefore, we also need a fast and simple method to evolve the nuclei with accurate forces.

Quantum Drude Oscillators can also be simulated using Path Integral Molecular

Dynamics, which uses the Feynman Path Integral formulation of Quantum Mechanics [Fey72]. Although in the classical limit one can simply write down probabilities related to the local energy, in quantum mechanics one must sum over amplitudes to construct probabilities, which leads to an integral over the action of all possible paths passing through that state. This can be further abstracted to say that the weight of a particular path depends on its action, and that we can sample the state-space of quantum paths in a manner analogous to classical states. Since the pseudo-electrons of the QDO model are distinguishable particles, there is no Fermi sign problem and the sampling of the paths is straight forward, and of order $N \log N \times P$ in computational cost, where P is the discretisation-number of the path, and N the number of atoms. We approximate the path to a chain of 'beads' with spring-like connections between them, and then perform classical-style molecular dynamics on the chain. Prosaically, the beads are like pseudo-classical 'parallel universes', and the springs represent 'interactions' between them. In the classical limit, the springs tighten and the beads collapse onto one another, so the springs represent quantum uncertainty at work.



PIMD pictorial

PIMD samples the free energy as an estimate of the ground-state Born-Oppenheimer energy. This assumes that the temperature is low enough for the real electronic structure to be close to the BO-surface, which is valid for the range of temperatures and pressures we are interested in.

The great advantage of PIMD is that it is based on classical Molecular Dynamics which allows us to use all the wealth of experience that exists already [FS96]. Using an adiabatic separation, we will show how it is possible to simultaneously evolve the QDO and nuclei such that the nuclei move on the ground state BO surface of the QDO, or at least an approximation that can be systematically improved by increasing the 'bead' number P.

1.7 Road Map

Here we present a broad overview of the structure and flow-of-the-argument in this thesis.

In chapters 2-4, we begin by describing the Quantum Drude model and its properties. The Quantum Drude model is introduced in chapter 2, by comparing it to the classical Drude model, describing its Hamiltonian and its damped coulomb interaction potentials.

We explore the Quantum *Dipole* Oscillator, which is the dipole-limit of the Quantum Drude, in chapter 3, as a gentle introduction to the many-body properties of Quantum Drudes. We use diagrams to illuminate the meaning of 'correlated uncertainty'. This simple model also permits analytic solutions of wavefunctions and energies, in terms of a simple expansion (in contrast to the complex diagrammatic expansion tackled in chapter that follows). We show how terms in the energy expansion correspond to many-body dispersion energies.

In chapter 4, we explore in more depth the properties of the full Quantum Drude model. We derive exact expressions for its multipole-polarisabilities and two-body multipole-dispersion. We explore the relationships between these quantities and the Quantum Drude's parameters, in order to test how close the behaviour of real atoms is to the ideal Quantum Drude, and derive some new parameter-fitting methods for a Quantum Drude to best mimic an arbitrary atom, in terms of its polarisabilites and dispersion coefficients. We introduce a new diagrammatic expansion that allows visualisation of many-body terms in the energy, including many-body polarisation terms and many-body dispersion terms. Finally we sketch possibilities for models with multiple Quantum Drudes per molecule.

Having motivated the Quantum Drude model by examining its properties, we move on to describe and develop the methods used for simulating it. Chapters 5-7 are devoted to studying QDOs using the Diffusion Monte Carlo (DMC), a method for generating accurate energies for static configurations of the nuclei, which we use primarily for fitting models. In chapter 5 we present the basic theory and methods of DMC, including the use of trial wavefunctions for importance sampling, then present improved diffusion-drift propagators and a new population-conserving scheme with reduced bias, that satisfies detailed balance unlike the traditional method.

Then in chapter 6 we present several new trial wavefunctions, including one that is very accurate for the dipole-limit QDO, and another that is very acccurate for the full model QDO (Gaussian on-site potentials modified by off-site Coulomb potentials). Combined with the tools developed in chapter 5, these wavefunctions allow us to make DMC more stable for large systems, and therefore allow us to obtain accurate energies

for a Quantum Drude model in a large system, such as the FCC-solid Xenon we study in chapter 7.

In chapter 7, we describe the Quantum Drude model of Xenon, developed by Martyna and Whitfield, as a useful test-case. We calculate the radial Born-Oppenheimer surface for the Xenon dimer simulated via DMC, and compare it to first-order perturbation theory results for various wavefunctions, in particular demonstrating the high-accuracy of the 'Coulomb' wavefunction developed in the chapter before. We simulate the Xenon FCC crystal in the dipole-limit, and compare to analytic results, as a test of the machinery, then simulate Xenon FCC crystal in the full Quantum Drude model, for which we calculate a lattice-constant and bulk modulus that are in remarkable agreement to experiment, as a demonstration of the transferability of this gas-phase model.

The next few chapters 8-10, are devoted to the study of QDO models using the more complex, but much more powerful Path Integral Molecular Dynamics method (PIMD), Chapter 8 introduces the concepts and machinery of PIMD which is a method for moving atoms on a good approximation to the Born-Oppenheimer surface, which can be systematically corrected by increasing the discretisation number. We start with the concept of 'high-temperature' approximate density matrices, then describe the complex machinery required for PIMD simulations.

Chapter 9, discusses the need for improved discretisations of the path-integrals, and some potential paths to solution, to motivate the search for improved 'high-temperature' approximate density matrices. Then we present a new dipole-approximation density matrix for use in PIMD, along with its requisite energy-estimators, pressure estimator and forces. We also sketch some approaches to porting the high-quality on-site plus external Coulomb trial wavefunction into a density matrix for PIMD.

In chapter 10, we use the new dipole-approximation density matrix to simulate Quantum Drude model Xenon in PIMD, presenting various tests, for the Xenon dimer and the Xenon FCC-crystal. We demonstrate that the energy is converged with reasonable values for the timestep, RESPA-number for multiple time-stepping and faux-mass for achieving an adiabatic separation, and show that performing PIMD for a Quantum Drude model Xenon dimer, gives a very close approximation to the exact ground-state Born-Oppenheimer energy potential calculated earlier via DMC, at the relatively high (Drude) temperature of $\hbar\omega/kT = 10.0$.

Finally, we present the general outline of a new model of water, the ultimate goal of this project, in chapter 11. We present Quantum Drude parameters that best reproduce its polarisability properties, and sketch how we will complete the rest of the fitting.

1.8 Summary of Original Contributions

- Chapter 4 is new material on the interaction properties of QDOs.
- Chapter 5: sections 5.1 and 5.2 describe existing DMC theory; sections 5.3 (accurate diffusion-drift propagators) and 5.4 (population-conserving DMC) are new.
- Chapter 6 describes new trial wavefunctions.
- Chapter 7 derives Martyna and Whitfield's QDO model for Xenon; then describes original work testing the new wavefunctions and studying Xenon using DMC.
- Chapter 9 describes new density matrices for PIMD.
- Chapter 10 describes original work testing the new density matrices by studing Xenon using PIMD.
- Chapter 11 is mostly a sketch of future work, but section 11.2.1 is new, and describes a rational parameterisation for a QDO-based model of water.

Chapter 2

The Quantum Drude Oscillator

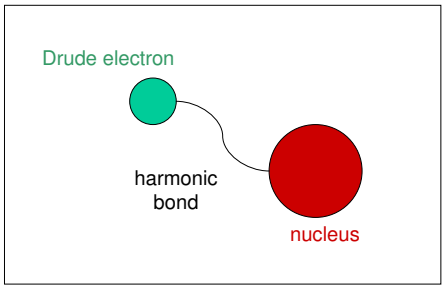
We introduce the Quantum Drude Oscillator model of an atom, comparing and contrasting it with the classical Drude model. We begin by discussing the model's quantum behaviour. We present the base harmonic oscillator Hamiltonian, and its interactions via damped Coulomb interactions.

2.1 The Model

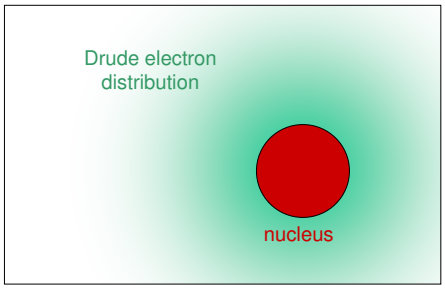
In 1900, Drude used a classical polarisable model to explain the refractive index of light [Dru00]. In the Classical Drude Oscillator model of an atom or molecule, a classical charged particle is attached to a centre (for example an atomic nucleus) by a simple harmonic bond whose force constant is related to the dipole polarizability. In other words, it is a pseudo-electron on a spring. For the Classical Drude model, only the minimised energy is meaningful; there is no spatial distribution of charge and it sits at the point in space which minimises the energy.

An external electric field can shift the equilibrium position of the charge, to produce a dipole, but as the position of a classical particle is precisely defined at any instant, this point-distribution is too simple to allow for quadrupole, octopole and higher multipole polarisations. That is, although two spatially separated charges technically have higher-order multipoles which depend on the choice of origin, these are not independent variables and the dipole always dominates in the Classical model.

The Quantum-Drude Oscillator is also localised by a harmonic potential, with the difference that the whole system, possibly containing many Quantum-Drudes and their coulombic interactions, is described by a Schrödinger wave equation. The pseudo-kinetic ∇^2 operator introduces Heisenberg uncertainty to the position of the particle, which means that the spatial distribution for the particle is "smeared-out" even at zero temperature. For real electrons and for Quantum-Drudes, this yields a charge



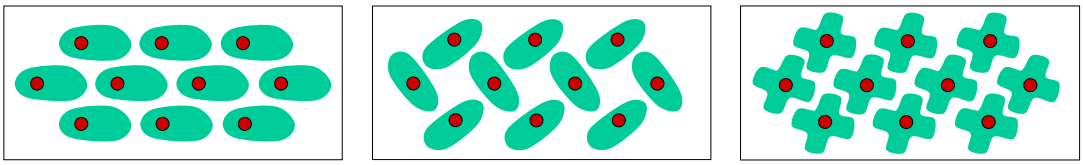
(a) Classical-Drudes have a well-defined position in space.



(b) Quantum-Drude position-states are inherently uncertain/smeared-out.

distribution rather than a point particle, which profoundly changes the behaviour. It is possible to have non-trivial charge distributions, and thus quadrupole, octopole and all the higher multipole polarisations, and when it interacts with other Quantum-Drudes, that also gives rise to a surprising effect, electronic dispersion, which does not exist for polarisable charge distributions in the classical limit.

Multipole many-body polarisation illustrated using deformable charge distributions.



Dipole Polarisation | Quadrupole Polarisation | Octupole Polarisation

2.1.1 Regularisation (damping) and repulsion terms

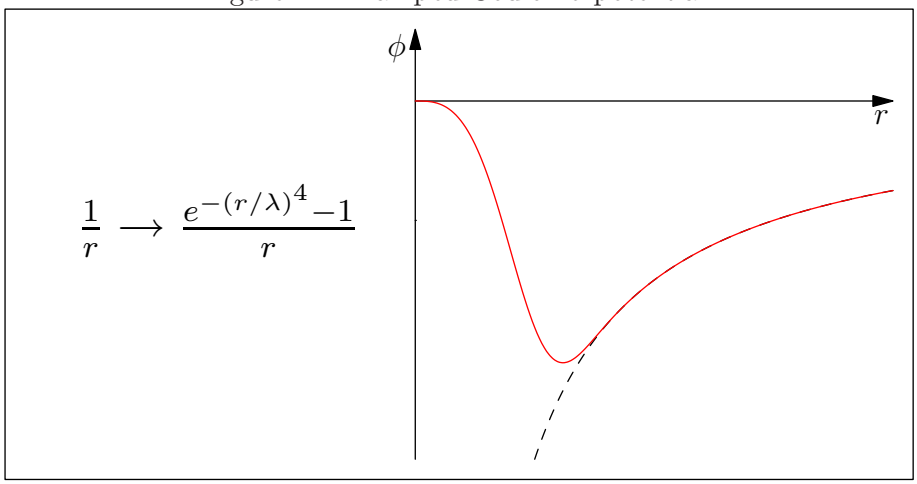
At very close range, real atoms are not well modelled by Coulomb effects alone. Electronic distributions are influenced by Pauli-exchange at close range, which resists overlap between the electrons of neighbouring atoms, pushing the atoms apart. Also, nuclear charge repulsion increases because the nuclei are localized while the neutralizing electron cloud is dispersed.

The Quantum-Drude model does not reproduce the true physics at short range. The simple coulombic potential between a Quantum-Drude and a neighbouring nucleus would be divergently attractive, causing that nucleus to 'yank on the spring'; this would be an unphysical artifact of the model. Additionally, the Quantum-Drude particle should not be thought of as a valence electron; it is only intended to represent the total shift in charge distribution from all the electronic orbitals. Thus it should never get too close to the centre of another charged point particle, such as a nucleus or another electron, or at least if it does, the Coulombic potential must not be permitted to diverge.

Regularisation of the Coulomb potential at small separations (often referred to as 'damping') neutralises these unphysical divergences as well as generating some repulsion between the atoms, and the remainder of the repulsive potential can be corrected by a sum of exponential potentials that act only at short interatomic separations.

For example, when designing a model of Xenon, Martyna and Whitfield [WM06] found that the coulombic attraction between the Quantum-Drude and a neighbouring atom is suitably damped in the following way:

Figure 2.1: Damped Coulomb potential

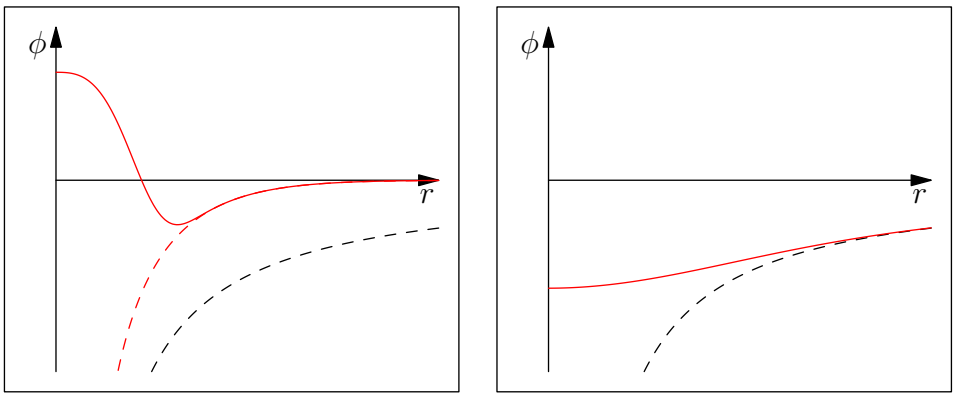


This potential produces repulsive forces within the damped range. The damped potential between two Quantum-Drude particles was chosen to have a similar form for the sake of simplicity, (although with a tighter damping radius and thus a higher repulsive barrier). Because of the damping form, the repulsive forces switch to unphysical attractive at very close range. This was thought not to matter, as was proved by the good results obtained. It is possible that the model could be improved by damping that potential to a flat plateau instead. We leave that possibility to future work.

2.1.2 Reciprocal-space Energy Sum

At long range, the potential is purely coulombic (the damping is all short-range), and can be treated exactly the same way as in classical simulations, the reciprocal Ewald sum, as described by Leeuw, Perram and Smith[dLPS80]. In small simulation boxes with periodic boundary conditions, it is sometimes necessary to include a few repetitions of the box in order to maintain high accuracy in calculating the real-space contribution. For large simulation boxes, the Particle Mesh Ewald technique[EPB$^+$95] and the Fast Multipole Method [Rok85, GR87] can be used to increase efficiency.

To perform Ewald summation, the potential must be split into short-range and long-range components. Because the damping is short-range, we can simply add the difference to the usual short-range component. It also makes sense to leave the long-range component unchanged because it is calculated via its Fourier transform, which would be non-trivial to calculate.



Short-range
(damping added)

Long-range
(unchanged)

Figure 2.2: Short-range and long-range components to the damped Coulomb potential. The black dotted line shows the full $1/r$ potential. The red dotted line shows what the short-range potential would be without damping.

2.2 The Hamiltonian

The Hamiltonian for a Quantum-Drude without interactions, in Cartesian coordinates is

$$\hat{H} = -\frac{\hbar^2}{2m}\left[\nabla_x^2 + \nabla_y^2 + \nabla_z^2\right] + \frac{m\omega^2}{2}\left[x^2 + y^2 + z^2\right].$$

The ground state wavefunction is simply:

$$\begin{aligned}\psi_0 &= \left(\frac{m\omega}{\pi\hbar}\right)^{\frac{3}{4}} \exp\left\{-\frac{m\omega}{2\hbar}\left[x^2 + y^2 + z^2\right]\right\} \\ &= \left(\frac{m\omega}{\pi\hbar}\right)^{\frac{3}{4}} \exp\left\{-\frac{m\omega}{2\hbar}\left[r^2\right]\right\}.\end{aligned}$$

In order to explore external interactions, it is convenient to use spherical polar coordinates. Following Flugge [Flu71] into spherical polar coordinates (and replacing

m with μ to avoid confusion with the magnetic quantum number m) yields

$$\hat{H} = -\frac{\hbar^2}{2\mu}\left[\frac{1}{r}\frac{\partial^2 r}{\partial r^2} - \frac{\hat{L}^2}{\hbar^2 r^2}\right] + \frac{\mu\omega^2}{2}r^2,$$

where

$$\hat{L}^2 = -\hbar^2\left[\frac{\partial^2}{\partial\theta^2} + \cot\theta\frac{\partial}{\partial\theta} + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right].$$

Because the $\hat{L}^2$ operator contains all the θ-, ϕ-dependence, it can be separated from the r-dependence, and thus also commutes with the Hamiltonian. In the r-dependent part, we can also simplify by cancelling the r-multiplicand in front of the ∂^2, which gives us

$$\psi_{klm}(r,\theta,\phi) = \frac{1}{r}\mathcal{R}_{kl}(r)\mathcal{Y}_{lm}(\theta,\phi),$$

where

$$-\frac{\hbar^2}{2\mu}\left[\frac{\partial^2}{\partial r^2} - \frac{l(l+1)}{r^2}\right]\mathcal{R}_{kl}(r) + \frac{\mu\omega^2}{2}r^2\mathcal{R}_{kl}(r) = E_{kl}\mathcal{R}_{kl}(r).$$

The eigenvectors and eigenvalues can be written as [Flu71]

$$\begin{aligned}\mathcal{R}_{kl}(r) &= N_{kl}\ r^{l+1}\ L_k^{(l+\frac{1}{2})}\left(\tfrac{\mu\omega}{\hbar}r^2\right)\ \exp\left(-\tfrac{\mu\omega}{2\hbar}r^2\right),\\ N_{kl} &= \left(\frac{\mu^3\omega^3}{4\pi\hbar^3}\right)^{1/4}\left(\frac{2^{k+2l+3}k!\left(\frac{\mu\omega}{2\hbar}\right)^l}{(2k+2l+1)!!}\right)^{1/2},\end{aligned}$$

and

$$E_{klm} = \hbar\omega\left(2k+l+\frac{3}{2}\right),$$

respectively. Here, the $L_k^{(l+\frac{1}{2})}$ are the generalized Laguerre polynomials.

Chapter 3

Quantum *Dipole* Oscillators: the dipole-limit QDO

We explore a simplified model of the Quantum Drude Oscillator in order to produce some exact solutions; the dipole-limit, or Quantum Dipole Oscillator. In the process we gain some insight into the more complex full Quantum model. We also use the dipole-limit for testing the machinery of our software, and its analytic form makes it useful for producing the first iteration of trial wave-functions and density matrices for increasing the efficiency of DMC and PIMD respectively.

First we calculate the exact solution, wavefunction and interacion energy for a dimer of interacting quantum dipole-oscillators in terms of its normal modes, using 1-dimensional oscillators to illustrate the meaning of correlated uncertainty, the effect that gives rise to dispersion. We then extend the normal-mode approach to an arbitrary assemblage of quantum dipole-oscillators using rotations to the diagonal basis to find the normal modes (we can also do this for actual cases using matrix diagonalisation, although this scales as N^3 and is not a method one would wish to use in large systems). Finally, we derive an matrix expansion in terms of the dipole-interaction tensor $\alpha\mathbf{T}$, generating the many-body dipole limit dispersion interactions.

These are not new results, as the dimer result was derived by Hirschfelder, Curtiss and Bird [HCB54], and the many-body limit was originally done by Cao and Berne [CB92a].

3.1 The Dipole Approximation for Neutral Atoms

The QDO model is approximated by replacing its spatially oscillating drude particle, with a harmonically oscillating point dipole-moment $\mu = Q_{\mathrm{D}}x$.
The Hamiltonian for a system of such oscillating point dipoles is therefore

$$\hat{H} \quad = \quad -\tfrac{\hbar^2}{2m}\nabla^2 + \tfrac{1}{2}m\omega^2 r^2 + \phi_{\mathrm{interaction}} - \tfrac{3}{2}\hbar\omega,$$

where this time the interaction energy is as follows (note μ now means dipole, not mass):

$$\phi_{\mathrm{interaction}} \quad = \quad -\tfrac{1}{2}\mu_{i\alpha}\mathbf{T}_{i\alpha j\beta}\mu_{j\beta}, \qquad \text{where} \quad \mu_i \equiv Q_{\mathrm{D}}r_i.$$

The tensor $\mathbf{T}$ is defined as follows:

$$\begin{aligned} T_{i\alpha j\beta} &= \nabla_{i\alpha}\nabla_{j\beta}\frac{1}{|R_{ij}|} \\ &= \frac{1}{|R|^5}\left(3R_\alpha R_\beta - R^2\delta_{\alpha\beta}\right), \qquad \text{with} \quad R = R_{ij} = (R_i - R_j). \end{aligned}$$

3.1.1 Damping and Periodicity of the Potential in the Dipole Limit

Aguado and Madden[AM03a] derive an Ewald sum for multipoles, which can include just dipole moments or also quadrupole moments. They are derived by taking appropriate derivatives of the normal Ewald sum, following Smith [Smi82, Smi98].

3.2 Single Quantum Dipole Oscillator

For an uncoupled dipole oscillator, the Hamiltonian is

$$H = \frac{-\hbar^2\omega^2\alpha}{2}\nabla_\mu^2 + \frac{1}{2}\frac{\mu^2}{\alpha},$$

where α is polarisability (units L^3), μ is a dipole moment ($Q \times L$), ω is frequency ($1/T$). This is the familiar harmonic oscillator Hamiltonian

$$H = \frac{-\hbar^2}{2m}\nabla_\mu^2 + \frac{1}{2}m\omega^2\mu^2,$$

with an effective mass, $m_{\text{eff}} = \frac{1}{\alpha\omega^2}$, and frequency $\omega_{\text{eff}}^2 = \frac{1}{\alpha m_{\text{eff}}} = \omega^2$,
for which the solution is

$$\begin{aligned}\psi &= \left(\frac{m\omega}{\pi\hbar}\right)^{\frac{1}{4}} e^{-\frac{m\omega}{2\hbar}\mu^2} = \left(\frac{1}{\pi\hbar\omega\alpha}\right)^{\frac{1}{4}} e^{-\frac{1}{2\hbar\omega\alpha}\mu^2},\\ \rho &= \psi^*\psi = \sqrt{\frac{m\omega}{\pi\hbar}} e^{-\frac{m\omega}{\hbar}\mu^2},\end{aligned}$$

with ground-state energy:

$$E_0 = \frac{3}{2}\hbar\omega \quad \text{in 3 dimensions.}$$

3.3 Coupled Pair of Dipole Oscillator

For two coupled dipole oscillators A and B, having the same frequency $\omega = \omega_0$, the Hamiltonian contains an additional interaction term:

$$H = H_{\text{A}} + H_{\text{B}} - \mu_{\text{A}}\mathbf{T}^{\text{AB}}\mu_{\text{B}},$$

or using the same substitution,

$$H = -\tfrac{\hbar^2}{2m}\left(\nabla^2_{\mu_{\text{A}}} + \nabla^2_{\mu_{\text{B}}}\right) + \tfrac{1}{2}m\omega_0^2(\mu_{\text{A}}^2 + \mu_{\text{B}}^2) - \mu_{\text{A}}\mathbf{T}^{\text{AB}}\mu_{\text{B}}.$$

The dipole tensor $\mathbf{T}^{\text{AB}}$ is defined:

$$T_{ij}^{\text{AB}} \equiv \nabla_i\nabla_j\left(\frac{1}{|R_{\text{AB}}|}\right) = \frac{3R_iR_j}{R^5} - \frac{\delta_{ij}}{R^3}.$$

3.3.1 Analytical Solution using Normal Modes

Following Hirschfelder-Curtiss-Bird [HCB54], we use the diagonal basis, where $\vec{R} = (0, 0, R)$,

$$T = \frac{1}{R^3}\begin{pmatrix} -1 & 0 & 0\\ 0 & -1 & 0\\ 0 & 0 & +2\end{pmatrix}.$$

Because T is diagonal, the Hamiltonian can be split into components:

$$H_i = -\tfrac{\hbar^2}{2m}\nabla^2_{\mu_{\text{A}},i} - \tfrac{\hbar^2}{2m}\nabla^2_{\mu_{\text{B}},i} + \tfrac{1}{2}m\omega_0^2\mu^2_{\text{A},i} + \tfrac{1}{2}m\omega_0^2\mu^2_{\text{B},i} - \mu_{\text{A},i}T_{ii}\mu_{\text{A},i}\ .$$

Now let $\nu_{(+)} = \frac{1}{\sqrt{2}}(\mu_{\rm A} + \mu_{\rm B})$, $\nu_{(-)} = \frac{1}{\sqrt{2}}(\mu_{\rm A} - \mu_{\rm B})$:

$$\nabla^2_{\nu_{(+)}} + \nabla^2_{\nu_{(-)}} = \nabla^2_{\mu_{\rm A}} + \nabla^2_{\mu_{\rm B}},$$
$$\nu^2_{\pm,i} = \tfrac{1}{2}(\mu^2_{{\rm A},i} + \mu^2_{{\rm B},i}) \pm \mu_{{\rm A},i}.\mu_{{\rm B},i} \ ,$$

rewrite the Hamiltonian,

$$H_{\pm,i} = \ -\tfrac{\hbar^2}{2m}\nabla^2_{\nu_+,i} + \tfrac{1}{2}\left(m\omega_0^2 \mp T_{ii}\right)\nu^2_{\pm,i} \ ,$$

and, recalling that $\alpha = \frac{1}{m\omega_0^2}$,

$$\begin{aligned} H_{\pm,i} &= -\tfrac{\hbar^2}{2m}\nabla^2_{\nu_\pm,i} + \tfrac{1}{2}m\omega^2_{\pm,i}\nu^2_{\pm,i} \\ E_{\pm,i} &= \tfrac{1}{2}\hbar\omega_{\pm,i} \ , \end{aligned}$$

with the normal modes:

$$\begin{aligned} \omega_{(+),i} &= \omega_0\sqrt{1+\alpha T_{ii}}, \qquad \text{(more energy, tighter oscillation)}, \\ \omega_{(-),i} &= \omega_0\sqrt{1-\alpha T_{ii}}, \qquad \text{(less energy, wider oscillation)}. \end{aligned}$$

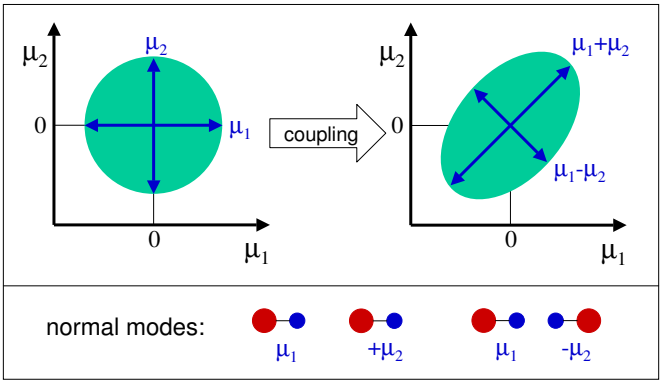


Figure 3.1: The normal modes of a pair of QDO's in 1 dimension

3.3.2 Dispersion Energy for a Pair of Dipole Oscillators

In the diagonal basis, we can do a Taylor expansion:

$$\begin{aligned} E &= \sum_{\pm,i} \tfrac{1}{2}\hbar\omega_{\pm,i} = \sum_{\pm,i} \tfrac{1}{2}\hbar\omega_0 \sqrt{1 \pm \alpha T_{ii}} \\ &= \sum_i \tfrac{1}{2}\hbar\omega_0 \left(1 - \tfrac{\alpha}{2}T_{ii} - \tfrac{\alpha^2}{8}T_{ii}^2\right) + \dots \\ &\quad + \sum_i \tfrac{1}{2}\hbar\omega_0 \left(1 + \tfrac{\alpha}{2}T_{ii} - \tfrac{\alpha^2}{8}T_{ii}^2\right) + \dots \\ &= 3\hbar\omega_0 - \tfrac{\alpha^2}{8}\hbar\omega \sum_i T_{ii}^2 + \dots \end{aligned}$$

The first term is the unperturbed energy of 6 harmonic oscillators (3 dim $\times$ 2 particles), and the second term is the interaction energy. In this simple case,

$$\sum_i T_i^2 = \frac{1}{R^6}\left((-1)^2 + (-1)^2 + (+2)^2\right) = \frac{6}{R^6}\,,$$

so the interaction energy is:

$$E_{dipole-dipole} \simeq -\frac{3}{4}\frac{\alpha^2}{R^6} = -\frac{C_6}{R^6}\,.$$

3.4 Many Coupled Dipole Oscillators

Following Cao and Berne [CB92a] we now investigate dispersion in the dipole-limit for many-body systems. For N dipoles distributed in space, the total interaction term comes from the sum of all the dipole pairs (making sure none are counted twice). Let A,B denote the A^{th},B^{th} dipole.

$$\begin{aligned} H' &= -\sum_{\mathrm{A>B}}^{\mathrm{N}} \mu_{\mathrm{A}} \mathbf{T}^{\mathrm{AB}} \mu_{\mathrm{B}} = -\tfrac{1}{2} \sum_{\mathrm{AB}}^{\mathrm{N}} \mu_{\mathrm{A}} \mathbf{T}^{\mathrm{AB}} \mu_{\mathrm{B}}, \\ H &= \sum_{A}^{\mathrm{N}} \left(-\tfrac{\hbar^2}{2m} \nabla^2_{\mu_{\mathrm{A}}} + \tfrac{1}{2} m \omega_0^2 \mu_{\mathrm{A}}^2 \right) - \tfrac{1}{2} \sum_{\mathrm{AB}}^{\mathrm{N}} \mu_{\mathrm{A}} \mathbf{T}^{\mathrm{AB}} \mu_{\mathrm{B}}. \end{aligned}$$

3.4.1 Analytical Solution using Normal Modes

If we treat all the 3-vectors μ_{A} as a single 3N-vector μ and the 3×3-tensors $\mathbf{T}^{\mathrm{AB}}$ as a 3N×3N-tensor $\mathbf{T}$, then we can write,

$$H = -\tfrac{\hbar^2}{2m} \nabla^2_{\mu} + \tfrac{1}{2} m \omega_0^2 \mu^2 - \tfrac{1}{2} \mu^{\mathrm{T}} \mathbf{T} \mu.$$

Tensor $\mathbf{T}$ is easy to resolve if diagonalised by a unitary transformation S to the eigenmode basis $\boldsymbol{\lambda}$.

$$\begin{aligned} S\mu &= \nu, & \mu &= S^{\mathrm{T}} \nu, \\ \mu^{\mathrm{T}} S^{\mathrm{T}} &= \nu, & \mu &= \nu^{\mathrm{T}} S, \\ S \mathbf{T} S^{\mathrm{T}} &= \boldsymbol{\lambda}, & \mathbf{T} &= S^{\mathrm{T}} \boldsymbol{\lambda} S, \end{aligned}$$

$$(i) \quad \mu^2 = \quad \mu^{\mathrm{T}} \mu = \nu^{\mathrm{T}} S S^{\mathrm{T}} \nu = \nu^{\mathrm{T}} \nu = \nu^2,$$

$$(ii) \quad \nabla_{\mu i} = \quad \frac{d}{d\mu_i} = \frac{\partial \nu_j}{\partial \mu_i} \frac{d}{d\nu_j} = S^{\mathrm{T}} \nabla_{\nu},$$

$$\nabla^2_{\mu} = \quad \nabla^{\mathrm{T}}_{\mu} \nabla_{\mu} = \nabla^{\mathrm{T}}_{\nu} S S^{\mathrm{T}} \nabla_{\nu} = S S^{\mathrm{T}} \nabla^{\mathrm{T}}_{\nu} \nabla_{\nu} = \nabla^{\mathrm{T}}_{\nu} \nabla_{\nu} = \nabla^2_{\nu},$$

$$(iii) \quad \mu^{\mathrm{T}} \mathbf{T} \mu = \quad \nu^{\mathrm{T}} \boldsymbol{\lambda} \nu.$$

Because the dipole tensor is now diagonal, the hamiltonian can be split into 3N components labelled by i:

$$\begin{aligned} H = &\quad \textstyle\sum_i \left(-\frac{\hbar^2}{2m}\nabla^2_{\nu i} + \frac{1}{2}m\omega_0^2\nu_i^2\right) - \frac{1}{2}\sum_i \nu_i\lambda_i\mu_i, \\ H = &\quad \textstyle\sum_i \left(-\frac{\hbar^2}{2m}\nabla^2_{\nu i} + \frac{1}{2}m\omega_i^2\nu_i^2\right), \\ &\quad \text{where } \omega_i = \omega_0\sqrt{1-\alpha\lambda_i}, \quad \alpha = (m\omega_0^2)^{-1}. \end{aligned}$$

So we can solve 3N independent wavefunctions:

$$H_i\psi_i = \left(-\frac{\hbar^2}{2m}\nabla^2_{\nu i} + \frac{1}{2}m\omega_i^2\nu_i^2\right)\psi_i \qquad \text{(no sum)}.$$

Omitting the normalisation factor gives the following ground state:

$$\begin{aligned} \psi_i(\nu_i) &= e^{-\frac{m\omega_i}{2\hbar}\nu_i^2}, \\ \Psi(\nu) &= \prod_i^{3N}\psi_i(\nu_i) = \exp\left(-\frac{m}{2\hbar}\sum_i \omega_i\nu_i^2\right). \end{aligned}$$

3.4.2 Wavefunction from Dipole-tensor Expansion

The above wavefunction is exact (but only in the dipole-limit of the QDO) if it is possible to diagonalise the dipole tensor, but in practice this will be too difficult and/or too computationally-expensive to be worth it, especially as we are really interested in the full model, where higher multipole terms dwarf the higher many-body dipole terms. Instead we take the exponent down and expand it in a Taylor series:

$$\begin{aligned} \sum_i \omega_i\nu_i^2 &= \omega_0\sum_i(1-\alpha\lambda_{ii})^{1/2}\nu_i^2 \\ &\simeq \omega_0\sum_i\left(1-\tfrac{\alpha}{2}\lambda_{ii} - \tfrac{\alpha^2}{8}\lambda_{ii}^2 + \mathcal{O}(\lambda^3)\right)\nu_i^2 \\ &\simeq \omega_0\left(\sum_i\nu_i^2 - \tfrac{\alpha}{2}\sum_i\nu_i\lambda_{ii}\nu_i - \tfrac{\alpha^2}{8}\sum_i\nu_i\lambda_{ii}\lambda_{ii}\nu_i + \mathcal{O}(\lambda^3)\right). \end{aligned}$$

Then we take the expansion back to the original basis μ, in which it is non-diagonal:

$$\begin{aligned}
\sum_i \nu_i^2 &= \nu^{\mathrm{T}}\nu = \mu^{\mathrm{T}}\mu = \sum_i \mu_i^2 \ , \\
\sum_i \nu_i \lambda_{ii} \nu_i &= \nu^{\mathrm{T}}\boldsymbol{\lambda}\nu = \mu^{\mathrm{T}}\mathbf{T}\mu = \sum_{ij} \mu_i T_{ij} \mu_j \ , \\
\sum_i \nu_i \lambda_{ii}\lambda_{ii} \nu_i &= \nu^{\mathrm{T}}\boldsymbol{\lambda}\boldsymbol{\lambda}\nu = \mu^{\mathrm{T}}\mathbf{T}\mathbf{T}\mu = \sum_{ijk} \mu_i T_{ij} T_{jk} \mu_k \ .
\end{aligned}$$

The third term is currently $\mathcal{O}\left(N^3\right)$ but because the tensor is symmetric ($T = T^{\mathrm{T}}$), it has a symmetry that can be exploited to reduce the calculation to $\mathcal{O}\left(N^2\right)$:

$$\begin{aligned}
\mu^{\mathrm{T}}\mathbf{T}\mathbf{T}\mu &= \mu^{\mathrm{T}}\mathbf{T}^{\mathrm{T}}\mathbf{T}\mu = (\mathbf{T}\mu)^{\mathrm{T}}(\mathbf{T}\mu) \\
&= (\mathbf{T}\mu)^2 = \sum_i \left(\sum_j T_{ij}\mu_j \right)^2 \ ,
\end{aligned}$$

$$\sum_i \omega_i \nu_i^2 \simeq \omega_0 \left(\mu^{\mathrm{T}}\mu - \tfrac{\alpha}{2}\mu^{\mathrm{T}}\mathbf{T}\mu - \tfrac{\alpha}{8}\mu^{\mathrm{T}}\mathbf{T}^2\mu \right) ,$$

$$\begin{aligned}
\Psi &\simeq e^{-\frac{m\omega_0}{2\hbar}\mu^2} e^{+\frac{\alpha m\omega_0}{4\hbar}\mu\mathbf{T}\mu} e^{+\frac{\alpha^2 m\omega_0}{16\hbar}\mu\mathbf{T}^2\mu} \ldots \\
&= e^{-\frac{1}{2\alpha\hbar\omega_0}\mu^2} e^{+\frac{1}{4\hbar\omega_0}\mu\mathbf{T}\mu} e^{+\frac{\alpha}{16\hbar\omega_0}\mu\mathbf{T}^2\mu} \ldots \ .
\end{aligned}$$

The first term is the unperturbed harmonic wavefunction ($\mu_i\mu_i$). The second term represents the effect of direct dipole-dipole interactions. The third term is the effect of interactions between two dipoles propagated by a third dipole.
They are terms from a more complex expansion which we will explain in the next chapter.

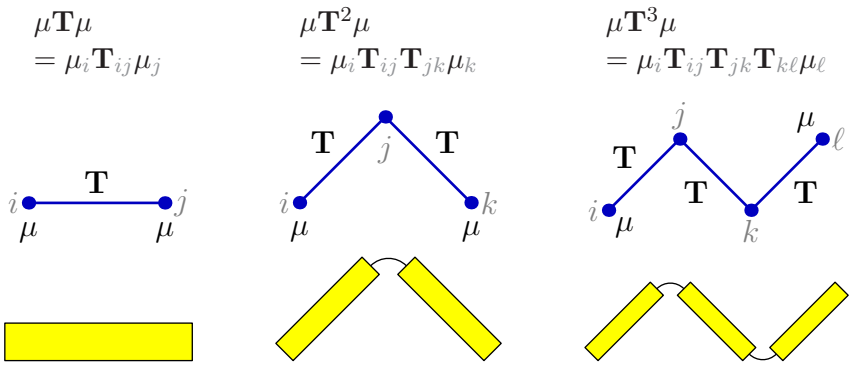


Figure 3.2: Feynman-style Propagator diagrams

3.4.3 Dispersion Energy from Dipole-tensor Expansion

In the diagonal or normal mode basis:

$$H\Psi \quad = \quad \sum_i H_i \psi_i = \tfrac{1}{2}\hbar \sum_i \omega_i = \tfrac{1}{2}\hbar\omega_0 \sum_i (1 - \alpha\lambda_{ii})^{\frac{1}{2}} .$$

Approximate again with a Taylor expansion:

$$(1 - \alpha\lambda_{ii})^{1/2} = 1 - \tfrac{\alpha}{2}\lambda_{ii} - \tfrac{\alpha^2}{8}\lambda_{ii}^2 + O(\lambda^3),$$

$$\begin{aligned} H\Psi \quad &\simeq \quad \tfrac{1}{2}\hbar\omega_0 \left(3N - \tfrac{\alpha}{2} \sum_i \lambda_{ii} - \tfrac{\alpha^2}{8} \sum_i \lambda_{ii}^2 \right) \\ &= \quad \tfrac{3N}{2}\hbar\omega_0 + \tfrac{1}{2}\hbar\omega_0 \left(-\tfrac{\alpha}{2}\mathbf{Tr}(\boldsymbol{\lambda}) - \tfrac{\alpha^2}{8}\mathbf{Tr}(\boldsymbol{\lambda}^2) \right) . \end{aligned}$$

We also know that traces are invariant under unitary transformations as $\mathbf{Tr}(M') = \mathbf{Tr}(SMS^{\mathrm{T}}) = \mathbf{Tr}(MS^{\mathrm{T}}S) = \mathbf{Tr}(M)$, and we know that $\mathbf{Tr}(\mathbf{T}) = 0$, as dipoles do not interact with themselves.

$$H\Psi \quad \simeq \quad \tfrac{3N}{2}\hbar\omega_0 - \tfrac{\alpha^2}{16}\hbar\omega_0 \mathbf{Tr}(\mathbf{T}^2) - \tfrac{\alpha^3}{32}\hbar\omega_0 \mathbf{Tr}(\mathbf{T}^3) \ldots .$$

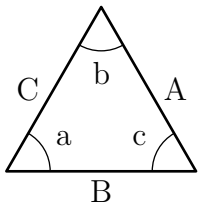
The first term is the unperturbed energy of the harmonic oscillators. In practice this term will often be dropped as it is not part of the interaction energy.

The second term is the sum of interaction energies between pairs of dipoles.

$$\begin{aligned} E^{(2)} &= -\tfrac{3\alpha^2\hbar\omega_0}{4R^6}, \\ C_6 &= -\tfrac{3}{4}\alpha^2\hbar\omega_0. \end{aligned}$$

The third term is the sum of interaction energies around triplets of dipoles. This is the Axilrod-Teller-Muto term [AT43].

$$\begin{aligned} E^{(3)} &= \tfrac{9}{16}\alpha^3\hbar\omega_0 \left[\frac{1+3\cos a\cos b\cos c}{A^3B^3C^3}\right], \\ C_9 &= \tfrac{3}{16}\alpha^3\hbar\omega_0 \qquad (\times\ 3 \text{ atoms}). \end{aligned}$$



The trace terms can also be represented diagrammatically, and represent a subset of the energy-expansion terms shown in the following chapter.

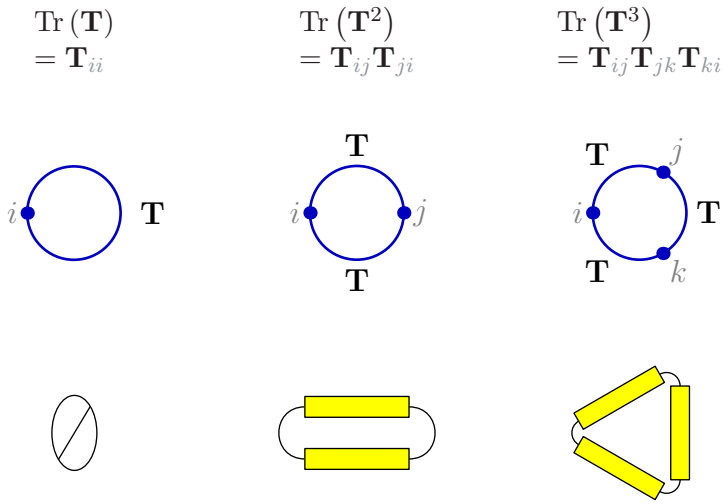


Figure 3.3: Feynman-style Bubble diagrams

Because $\mathbf{Tr}(\mathbf{T}) = 0$, it does not appear in the energy, but $\mathbf{Tr}(\mathbf{T}^2)$ does appear.
If a higher order expansion of the wavefunction was used, higher order traces could appear, and would represent explicit 3-body and many-body terms in the ground-state energy. Thus this expansion allows us to generate the coefficients of the 4-body, 5-body, 6-body, ... many-body dipole-limit dispersion terms for the model.

Calculating the Higher-Order Traces

To calculate these higher order traces, one would first need to construct the complete tensor $\mathbf{T}$, from the individual dipole-interaction tensors T_{ij} that exists between two molecules i and j.

$$\mathbf{T} = \begin{pmatrix} 0 & \mathrm{T}_{12} & \mathrm{T}_{13} & \mathrm{T}_{14} \\ \mathrm{T}_{21} & 0 & \mathrm{T}_{23} & \mathrm{T}_{24} \\ \mathrm{T}_{31} & \mathrm{T}_{32} & 0 & \mathrm{T}_{34} \\ \mathrm{T}_{41} & \mathrm{T}_{42} & \mathrm{T}_{43} & 0 \end{pmatrix}, \qquad \text{where} \quad \mathrm{T}_{ij} = \mathrm{T}_{ji} = \nabla_i \nabla_j \left(1/r_{ij}\right).$$

The diagonal elements will be zero unless the system is periodic, in which case a dipole can interact with its own image. Periodic systems, however have to be calculated by Fourier transform, and the $\mathbf{T}$ tensor can be calculated using complex numbers, in which case each derivative becomes a complex derivative, so that each entry in the $\mathbf{T}$ matrix itself becomes a 2×2 matrix.

In either case, if there are N particles in the system, the total storage required for the $\mathbf{T}$-tensor is of order N^2. When multiplying the matrices, there are $\mathcal{O}\left(N^2\right)$ new components to calculate, and each requires an contraction over a dummy index, which is $\mathcal{O}\left(N\right)$, giving a total cost of $\mathcal{O}\left(N^3\right)$ operations. The simplest way to multiply the $\mathbf{T}$-tensors is by repeatedly multiplying the original tensor.

$$\left[\mathbf{T}^{n+1}\right]_{ij} = \sum_k \mathbf{T}_{ik} \left[\mathbf{T}^n\right]_{kj}.$$

In terms of storage space, this requires a total of 3 matrices. The trace can be calculated separately at each power:

$$\mathrm{Tr}\left(\mathbf{T}^n\right) = \sum_i \left[\mathbf{T}^n\right]_{ii}.$$

Chapter 4

Properties of the Full Quantum Drude model

We work through properties of the full Quantum Drude in more detail, in order to demonstrate that it is capable of realistically modelling the actual properties of atoms and molecules. We begin by deriving its polarisabilities and dispersion coefficients, including mixed-species dispersion, in terms of the basic paramters of the model. The polarisability derivation has not been published specifically for Quantum Drudes before. We derive some previously unpublished 'combining rules'; unit-less ratios between polarisabilities and dispersion coefficients quantities that are exact for Quantum Drudes, irrespective of parameterisation, to which we can compare real atoms, as a test of how well Quantum Drudes actually reproduce their behaviour. We present a completely new scheme for visualising the many-body, multipolar interactions of the Quantum Drude, including many-body multipolar polarisation, many-body multipolar dispersion, and terms that are intermediate between the two. Finally, we discuss some possibilities for augmenting the model, by using multiple Quantum-Drudes per molecule, mentioning issues of symmetry, and then adding interactions between Quantum-Drudes within a molecule that might yield correct hyperpolarisation.

4.1 Polarisabilities of a Quantum Drude

The following derivation has not been published before, to my knowledge. In order to develop an analytic expression for the polarizabilities of the Quantum Drude oscillator model, the perturbing effect of a test charge at a large distance R (distinguish from

$\mathcal{R}(r)$) away on the z-axis, is examined.

$$H'(r,\theta) \;=\; Qq\sum_{l=1}^{\infty}\left(\frac{r^l}{R^{l+1}}\right)\left(\frac{4\pi}{2l+1}\right)^{1/2}\mathcal{Y}_{l0}(\theta).$$

Using the fact that $\mathcal{Y}_{00}(\theta) = 1/\sqrt{4\pi}$, the perturbation to the ground state energy is zero to 1st order.

The 2nd order perturbation to the energy is, similarly,

$$\begin{aligned} E^{(2)} &= -\sum_{kl}\frac{|\langle 00|H'|kl\rangle|^2}{(2k+l)\,\hbar\omega} \\ &= -\sum_{kl}\frac{\left|\int_r\int_\Omega \mathcal{R}_{00}\mathcal{Y}_{00}\left[\sum_{l'}\left(\frac{r^{l'}}{R^{l'+1}}\right)\left(\frac{4\pi}{2l'+1}\right)^{1/2}\mathcal{Y}_{l'0}\right]\mathcal{R}_{kl}\mathcal{Y}_{l0}\right|^2}{(2k+l)\,\hbar\omega}, \end{aligned}$$

which simplifies because $\mathcal{Y}_{00} = 1/\sqrt{4\pi}$ and $\int_\Omega \mathcal{Y}_{l0}\mathcal{Y}_{l'0} \equiv \langle l|l'\rangle \equiv \delta_{ll'}$:

$$E^{(2)} \;=\; \sum_{kl} -\left[\frac{Q^2q^2}{R^{2l+2}(2l+1)}\right]\frac{\left[\int_0^\infty \mathrm{d}r\, r^l\mathcal{R}_{00}\mathcal{R}_{kl}\right]^2}{(2k+l)\,\hbar\omega}.$$

The $k = 0$ modes have another property which allows a simplification.

$$\left[\frac{\mathcal{R}_{0l}}{N_{0l}}\right] = r^l\left[\frac{\mathcal{R}_{00}}{N_{00}}\right] \qquad \Longrightarrow \qquad r^l\mathcal{R}_{00} = C_l\mathcal{R}_{0l}, \qquad \text{where} \quad C_l = \frac{N_{00}}{N_{0l}}.$$

The $k \neq 0$ and $l \neq l'$ modes do not contribute because of orthogonality

$$\langle \mathcal{R}_{kl}|\mathcal{R}_{k'l'}\rangle = \delta_{kk'll'}\ .$$

This gives us

$$\begin{aligned} E^{(2)} &= \sum_l E_l^{(2)} \\ E_l^{(2)} &= -\left[\frac{Q^2q^2}{R^{2l+2}(2l+1)}\right]\frac{\left[C_l\int_0^\infty \mathrm{d}r\,\mathcal{R}_{0l}\mathcal{R}_{0l}\right]^2}{l\hbar\omega} \\ &= -\left[\frac{Q^2q^2C_l^2}{R^{2l+2}(2l+1)l\hbar\omega}\right], \qquad \text{and} \quad C_l^2 = (2l+1)!!\left(\tfrac{\hbar}{2\mu\omega}\right)^l, \\ &= -\left[\frac{q^2}{\mu\omega^2}\right]\left[\frac{Q^2(2l-1)!!}{2R^{2l+2}l}\right]\left[\tfrac{\hbar}{2\mu\omega}\right]^{l-1}, \end{aligned}$$

where we are temporarily using μ for the Drude mass instead of m in order to avoid confusion with the magnetic quantum number.
The polarisabilities are defined as follows:

$$E_l^{(2)} = -\left[\frac{Q^2\alpha_l}{2R^{2l+2}}\right],$$

which implies,

$$\begin{aligned}\alpha_l &= \left[\frac{q^2}{\mu\omega^2}\right]\left[\frac{(2l-1)!!}{l}\right]\left[\tfrac{\hbar}{2\mu\omega}\right]^{l-1},\\ \alpha_l &= \alpha_1\left[\frac{(2l-1)!!}{l}\right]\left[\tfrac{\hbar}{2\mu\omega}\right]^{l-1}.\end{aligned}$$

In the classical limit, $\lim \hbar \to 0$, only the dipole polarizability is non-zero:

$$\alpha_1 \equiv \frac{q^2}{\mu\omega^2}, \qquad \alpha_l \to 0 \quad \forall l > 1.$$

We can also write down some explicit relationships between the polarisabilities.

$$\begin{aligned}\alpha_1 &= \frac{q^2}{\mu\omega^2},\\ \alpha_2 &= \tfrac{3}{4}\left(\tfrac{\hbar}{\mu\omega}\right)\alpha_1,\\ \alpha_3 &= \tfrac{5}{4}\left(\tfrac{\hbar}{\mu\omega}\right)^2\alpha_1 = \tfrac{5}{3}\tfrac{\hbar}{\mu\omega}\alpha_2.\end{aligned}$$

4.2 Dispersion between Quantum Drudes

This derivation is essentially the same as was published by Fontana[Fon61], where he estimates the dispersion energy for noble gas atoms, using the approximation that those atoms respond like harmonic oscillators. Similar approaches have been used by others [HCB54, Sto96].

We can write down the interaction Hamiltonian in an shorthand expansion of multipoles Q and multipole-interaction tensors T.

$$H'_{\mathrm{AB}} \equiv \sum_{m,n} Q_m^{\mathrm{A}}\, Q_n^{\mathrm{B}}\, T_{mn}^{\mathrm{AB}}.$$

It is useful to use spherical coordinate basis again, where the two sets of tensors Q_m^{A} become essentially two spherical vectors representing fluctuating charge distributions on A and B, and the tensors T_{mn}^{AB} become essentially a single interaction-tensor that depends only on the positions and orientations of molecules A and B. [Sto96].

Next, we plug this into perturbation theory as before. The derivation is similar to that for the polarisation; the k-sum will be truncated as before, so it is omitted, and there is an implicit sum over the magnetic quantum-number m, to make up the total number of components, but we omit this also for clarity. This time the interaction energy introduces coupling between two quantum-drude ground-state wavefunctions.

$$E^{(2)} = \sum_{l_{\mathrm{A}} l_{\mathrm{B}}} \frac{|\langle 00|H'_{\mathrm{AB}}|l_{\mathrm{A}} l_{\mathrm{B}}\rangle|^2}{l_{\mathrm{A}}\hbar\omega_{\mathrm{A}} + l_{\mathrm{B}}\hbar\omega_{\mathrm{B}}}.$$

Each set of multipole l-values picks out its counterpart in the wavefunction. The T-tensor components can be pulled out as constants.

$$\begin{aligned} E^{(2)} &= \sum_{l_{\mathrm{A}} l_{\mathrm{B}}} \frac{\left[T_{l_{\mathrm{A}} l_{\mathrm{B}}}^{\mathrm{AB}} \langle 0|Q_{l_{\mathrm{A}}}^{\mathrm{A}}|l_{\mathrm{A}}\rangle\langle 0|Q_{l_{\mathrm{B}}}^{\mathrm{B}}|l_{\mathrm{B}}\rangle\right]^2}{l_{\mathrm{A}}\hbar\omega_{\mathrm{A}} + l_{\mathrm{B}}\hbar\omega_{\mathrm{B}}} \\ &= \sum_{l_{\mathrm{A}} l_{\mathrm{B}}} T_{l_{\mathrm{A}} l_{\mathrm{B}}}^{\mathrm{AB}} T_{l_{\mathrm{A}} l_{\mathrm{B}}}^{\mathrm{AB}} \frac{\left|\langle 0|Q_{l_{\mathrm{A}}}^{\mathrm{A}}|l_{\mathrm{A}}\rangle\right|^2 \left|\langle 0|Q_{l_{\mathrm{B}}}^{\mathrm{B}}|l_{\mathrm{B}}\rangle\right|^2}{l_{\mathrm{A}}\hbar\omega_{\mathrm{A}} + l_{\mathrm{B}}\hbar\omega_{\mathrm{B}}}. \end{aligned}$$

The wavefunction projections can be resolved in the same way as for the polarisability, and can be simplified in terms of the two atoms' multipole polarisabilities.

$$\begin{aligned} E^{(2)} &= \sum_{l_{\mathrm{A}} l_{\mathrm{B}}} T_{l_{\mathrm{A}} l_{\mathrm{B}}}^{\mathrm{AB}} T_{l_{\mathrm{A}} l_{\mathrm{B}}}^{\mathrm{AB}} \frac{C_{l_{\mathrm{A}}}^2}{(2l_{\mathrm{A}}+1)} \frac{C_{l_{\mathrm{B}}}^2}{(2l_{\mathrm{B}}+1)} \frac{1}{l_{\mathrm{A}}\hbar\omega_{\mathrm{A}} + l_{\mathrm{B}}\hbar\omega_{\mathrm{B}}} \\ &= \sum_{l_{\mathrm{A}} l_{\mathrm{B}}} T_{l_{\mathrm{A}} l_{\mathrm{B}}}^{\mathrm{AB}} T_{l_{\mathrm{A}} l_{\mathrm{B}}}^{\mathrm{AB}} \left[\frac{q_{\mathrm{A}}^2 q_{\mathrm{B}}^2 (2l_{\mathrm{A}}-1)!!(2l_{\mathrm{B}}-1)!!}{l_{\mathrm{A}}\hbar\omega_{\mathrm{A}} + l_{\mathrm{B}}\hbar\omega_{\mathrm{B}}}\right] \left[\frac{\hbar}{2\mu_{\mathrm{A}}\omega_{\mathrm{A}}}\right]^{l_{\mathrm{A}}} \left[\frac{\hbar}{2\mu_{\mathrm{B}}\omega_{\mathrm{B}}}\right]^{l_{\mathrm{B}}} \\ &= \sum_{l_{\mathrm{A}} l_{\mathrm{B}}} T_{l_{\mathrm{A}} l_{\mathrm{B}}}^{\mathrm{AB}} T_{l_{\mathrm{A}} l_{\mathrm{B}}}^{\mathrm{AB}} \ \alpha_{l_{\mathrm{A}}}^{\mathrm{A}} \ \alpha_{l_{\mathrm{B}}}^{\mathrm{B}} \ \left[\frac{\hbar}{4} \frac{l_{\mathrm{A}} l_{\mathrm{B}} \omega_{\mathrm{A}} \omega_{\mathrm{B}}}{(l_{\mathrm{A}}\omega_{\mathrm{A}} + l_{\mathrm{B}}\omega_{\mathrm{B}})}\right]. \end{aligned}$$

The α terms should strictly be tensors in the magnetic components, but Quantum Drude Oscillators are isotropic, so we take them to be simple constants. The implied sum becomes a sum over T-tensor components-squared. In cartesian coordinates these traces can also be evaluated using a recursive relation, but in polar coordinates, it is simplest to sum the components explicitly. The components are listed in Stone's book [Sto96], but here we simply list the results. Any terms with $l_{\mathrm{A}} = 0$ or $l_{\mathrm{B}} = 0$ will

vanish.

$$|T_{1,1}|^2 = 6R^{-6}, \qquad |T_{1,2}|^2 = 15R^{-8},$$
$$|T_{1,3}|^2 = 21R^{-10}, \qquad |T_{2,2}|^2 = 70R^{-10}.$$

Inserting those values,

$$\begin{aligned}
E^{(2)} &= \left[6R^{-6}\right]\alpha_1^{\mathrm{A}}\alpha_1^{\mathrm{B}}\left[\frac{\hbar}{4}\frac{\omega_{\mathrm{A}}\omega_{\mathrm{B}}}{(\omega_{\mathrm{A}}+\omega_{\mathrm{B}})}\right] \\
&\quad + \left[15R^{-8}\right]\alpha_1^{\mathrm{A}}\alpha_2^{\mathrm{B}}\left[\frac{\hbar}{4}\frac{2\omega_{\mathrm{A}}\omega_{\mathrm{B}}}{(\omega_{\mathrm{A}}+2\omega_{\mathrm{B}})}\right] + \left[15R^{-8}\right]\alpha_2^{\mathrm{A}}\alpha_1^{\mathrm{B}}\left[\frac{\hbar}{4}\frac{2\omega_{\mathrm{A}}\omega_{\mathrm{B}}}{(2\omega_{\mathrm{A}}+\omega_{\mathrm{B}})}\right] \\
&\quad + \left[21R^{-10}\right]\alpha_1^{\mathrm{A}}\alpha_3^{\mathrm{B}}\left[\frac{\hbar}{4}\frac{3\omega_{\mathrm{A}}\omega_{\mathrm{B}}}{(\omega_{\mathrm{A}}+3\omega_{\mathrm{B}})}\right] + \left[21R^{-10}\right]\alpha_3^{\mathrm{A}}\alpha_1^{\mathrm{B}}\left[\frac{\hbar}{4}\frac{3\omega_{\mathrm{A}}\omega_{\mathrm{B}}}{(3\omega_{\mathrm{A}}+\omega_{\mathrm{B}})}\right] \\
&\quad + \left[70R^{-10}\right]\alpha_2^{\mathrm{A}}\alpha_2^{\mathrm{B}}\left[\frac{\hbar}{4}\frac{4\omega_{\mathrm{A}}\omega_{\mathrm{B}}}{(2\omega_{\mathrm{A}}+2\omega_{\mathrm{B}})}\right] \qquad + \mathcal{O}\left(R^{-12}\right), \\
E^{(2)} &\equiv C_6R^{-6} + C_8R^{-8} + C_{10}R^{-10} + \mathcal{O}\left(R^{-12}\right).
\end{aligned}$$

The two-body dispersion coefficients C_6, C_8, C_{10} etcetera can be extracted:

$$\begin{aligned}
C_6^{\mathrm{AB}} &= \tfrac{3}{2}\alpha_1^{\mathrm{A}}\alpha_1^{\mathrm{B}}\frac{\hbar\omega_{\mathrm{A}}\omega_{\mathrm{B}}}{(\omega_{\mathrm{A}}+\omega_{\mathrm{B}})}, \\
C_8^{\mathrm{AB}} &= \tfrac{15}{2}\left[\alpha_1^{\mathrm{A}}\alpha_2^{\mathrm{B}}\frac{\hbar\omega_{\mathrm{A}}\omega_{\mathrm{B}}}{(\omega_{\mathrm{A}}+2\omega_{\mathrm{B}})} + \alpha_2^{\mathrm{A}}\alpha_1^{\mathrm{B}}\frac{\hbar\omega_{\mathrm{A}}\omega_{\mathrm{B}}}{(2\omega_{\mathrm{A}}+\omega_{\mathrm{B}})}\right], \\
C_{10}^{\mathrm{AB}} &= \left[21\alpha_1^{\mathrm{A}}\alpha_3^{\mathrm{B}}\frac{\hbar\omega_{\mathrm{A}}\omega_{\mathrm{B}}}{(\omega_{\mathrm{A}}+3\omega_{\mathrm{B}})} + 21\alpha_3^{\mathrm{A}}\alpha_1^{\mathrm{B}}\frac{\hbar\omega_{\mathrm{A}}\omega_{\mathrm{B}}}{(3\omega_{\mathrm{A}}+\omega_{\mathrm{B}})} + 70\alpha_2^{\mathrm{A}}\alpha_2^{\mathrm{B}}\frac{\hbar\omega_{\mathrm{A}}\omega_{\mathrm{B}}}{(2\omega_{\mathrm{A}}+2\omega_{\mathrm{B}})}\right].
\end{aligned}$$

If the drude particles are also identical: $\omega_{\mathrm{A}} = \omega_{\mathrm{B}} = \omega$, $\alpha^{\mathrm{A}} = \alpha^{\mathrm{B}} = \alpha$, then:

$$\begin{aligned}
C_6 &= \tfrac{3}{4}\alpha_1\alpha_1\hbar\omega, \\
C_8 &= 5\alpha_1\alpha_2\hbar\omega, \\
C_{10} &= \left[\tfrac{21}{2}\alpha_1\alpha_3 + \tfrac{70}{4}\alpha_2\alpha_2\right]\hbar\omega.
\end{aligned}$$

It is possible also to calculate 3-body dispersion energies using spherical polar methods, but this is much more complicated. For now, we simply take the 3-body result in the dipole limit, which we derive in the next chapter, which is the Axilrod-Teller-Muto term:

$$C_9 = \tfrac{3}{16}\alpha_1^3\hbar\omega.$$

By substituting in our previous polarisability relations, we can write down some

relationships that are more explicit in terms of the underlying Quantum-Drude parameters q, m and ω.

$$\begin{aligned} C_8 &= \tfrac{15}{4}\left(\tfrac{\hbar}{\mu\omega}\right)\alpha_1\alpha_1\hbar\omega \\ &= 5\left(\tfrac{\hbar}{\mu\omega}\right)\times C_6, \\ C_{10} &= \left[\tfrac{105}{8}+\tfrac{315}{32}\right]\alpha_1\alpha_1\left(\tfrac{\hbar}{\mu\omega}\right)^2\hbar\omega \\ &= \tfrac{735}{32}\alpha_1\alpha_1\left(\tfrac{\hbar}{\mu\omega}\right)^2\hbar\omega \\ &= \tfrac{245}{8}\left(\tfrac{\hbar}{\mu\omega}\right)^2\times C_6, \end{aligned}$$

$$C_9 = \tfrac{1}{4}\alpha_1\times C_6.$$

4.3 How Realistic are Quantum Drudes?

The attraction of the Quantum Drude is that it is a simplified model for describing polarisation and dispersion. But its usefulness depends on its ability to mimic the real physical behaviour of real atoms. Here we describe a series of tests (which have not been previously published) that compare the behaviour of real atoms to that of the ideal Quantum Drude model. This is an important first test of the Quantum Drude model using independent theoretical and experimental data. We find that Quantum Drudes are in fact a very good model of the polarisation and dispersion behaviour of real atoms. Later, in chapter 11 we will find this even extends also to the strongly anisotropic water molecule, so there is good reason to expect that it will be good at describing many other molecular groups as well.

In particular, we show that the Quantum Drude model is good at reproducing the dispersion-interaction behaviour between mixed species of atoms. This means that Quantum Drude parameters fitted for single species will automatically generate correct long-range interactions when they interact with arbitrary new neighbours, without having to parameterise each possible pair, or triplet (The standard approach to date had been to parameterise every possible pair potential). This is an excellent example of the model's ability to support simple models that are widely applicable and transferable.

We derive a set of relations, between some of the properties of atoms relevant to polarisation and dispersion, including the dipole-polarisability (α_1), the quadrupole-polarisability (α_2), the octopole-polarisability (α_3), the two-body single species dispersion coefficients (C_6, C_8 and C_{10}), the three-body single species dispersion

coefficient (C_9), and the leading mixed species dispersion coefficient (C_6^{AB}).

Each relation is derived such that it yields a dimensionless constant, and it is chosen in such a way that that constant is exactly equal to 1 for the Quantum Drude model, independent of its choice of parameters (m, ω, q). Each relation can then be calculated for a set of real atoms, as a measure of how much real atoms deviate from the ideal behaviour of Quantum Drudes. We found that agreement is good, for both noble gases and alkali metal atoms, the atoms which we have independent data to compare.

4.3.1 Combining Rules

The expressions we just derived for the polarisabilities α_n and dispersion parameters $C_{(2n+2)}$ can be combined into four new dimensionless constants, or combining rules. Here are the parameters:

$$\begin{aligned} \alpha_1 &= \frac{q^2}{\mu\omega^2}, & C_6 &= \tfrac{3}{4}\alpha_1\alpha_1\hbar\omega, \\ \alpha_2 &= \tfrac{3}{4}\left(\tfrac{\hbar}{\mu\omega}\right)\alpha_1, & C_8 &= 5\left(\tfrac{\hbar}{\mu\omega}\right)\times C_6, \\ \alpha_3 &= \tfrac{5}{4}\left(\tfrac{\hbar}{\mu\omega}\right)^2\alpha_1, & C_{10} &= \tfrac{245}{8}\left(\tfrac{\hbar}{\mu\omega}\right)^2\times C_6, \\ & & C_9 &= \tfrac{1}{4}\alpha_1\times C_6. \end{aligned} \tag{4.1}$$

From these it is straightforward to derive three of the combining rules:

$$\sqrt{\tfrac{20}{9}}\frac{\alpha_2}{\sqrt{\alpha_1\alpha_3}} = 1 \qquad \text{(polarisability combining rule)},$$

$$\sqrt{\tfrac{49}{40}}\frac{C_8}{\sqrt{C_6C_{10}}} = 1 \qquad \text{(dispersion combining rule)},$$

$$\frac{C_6\alpha_1}{4C_9} = 1 \qquad \text{(3-body dispersion rule)}.$$

The case of mixed-species dispersion is more tricky.
Substitute the following;

$$C_6^{\mathrm{AA}} = \tfrac{3}{4}\alpha_1^{\mathrm{A}}\alpha_1^{\mathrm{A}}\hbar\omega_{\mathrm{A}} \implies \qquad \hbar\omega_{\mathrm{A}} = \frac{4C_6^{\mathrm{AA}}}{3\alpha_1^{\mathrm{A}}\alpha_1^{\mathrm{A}}},$$

into the expression for mixed-species dispersion,

$$C_6^{AB} = \tfrac{3}{2}\alpha_1^A\alpha_1^B\frac{\hbar\omega_A\omega_B}{(\omega_A+\omega_B)} = \tfrac{3}{2}\alpha_1^A\alpha_1^B\frac{\left(\frac{4C_6^{AA}}{3\alpha_1^A\alpha_1^A}\right)\left(\frac{4C_6^{BB}}{3\alpha_1^B\alpha_1^B}\right)}{\left(\frac{4C_6^{AA}}{3\alpha_1^A\alpha_1^A}\right)+\left(\frac{4C_6^{BB}}{3\alpha_1^B\alpha_1^B}\right)}$$
$$= \frac{2C_6^{AA}C_6^{BB}\alpha_1^A\alpha_1^B}{C_6^{AA}\alpha_1^B\alpha_1^B+C_6^{BB}\alpha_1^A\alpha_1^A},$$

to yield the mixed-species dispersion rule:

$$\frac{1}{C_6^{AB}}\left[\frac{2C_6^{AA}C_6^{BB}\alpha_1^A\alpha_1^B}{C_6^{AA}\alpha_1^B\alpha_1^B+C_6^{BB}\alpha_1^A\alpha_1^A}\right]=1.$$

Below, we test these rules for some a set of group I atoms (alkali metals) and group VIII (noble gases), and calculate Δ, the divergence in the behaviour of the real atom (or atom pair in the case of mixed dispersion), from the ideal Quantum Drude case. Upper and lower bounds for α_1, α_2, α_3, C_6, C_8, C_{10}, and C_9 were obtained from two papers by Certain et al [TNC76, SC85], which used a combination of theory and experiment. [1] We find that almost all atoms show less than 10% deviation from ideal behaviour, and some are much better than that.

Table 4.1: Polarisability Combining Rule

Atom	α_1	α_2	α_3	rule	Δ	err
H	4.5	15	131.25	0.920	-8.0%	0.2%
He	1.38	2.44	10.6	0.952	-4.8%	0.2%
Ne	2.66	6.42	34.27	1.003	0.3%	0.2%
Ar	11.1	50.21	531.3	0.975	-2.5%	0.2%
Kr	16.7	95.55	1260	0.982	-1.8%	0.2%
Xe	27.3	212.6	3602	1.011	1.1%	0.2%
Li	164	1383	36800	0.839	-16.1%	0.2%
Na	161	1799	51170	0.934	-6.6%	0.2%
K	291	4597	150200	1.037	3.7%	0.2%
Rb	322	5979	212700	1.077	7.7%	0.2%
Cs	409	9478	339900	1.198	19.8%	0.2%

[1] We have omitted the combining rules involving dispersion coefficients of sodium(Na) because their methods resulted in large uncertainties for those values, rendering the result of the combining rule somewhat meaningless (that is, they were consistent with the combining rules, but within an error-bar of more than 40%).

Table 4.2: Dispersion Combining Rule

Atom	C_6	C_8	C_{10}	rule	Δ	err
H	6.49 ± 0.02	124.5 ± 0.5	3280 ± 10	0.944	-5.6%	0.7%
He	1.46 ± 0.02	14.05 ± 0.15	182.5 ± 1.5	0.954	-4.6%	1.9%
Ne	6.88 ± 0.4	76 ± 20.5	1173 ± 347	0.937	-6.3%	41.8%
Ar	67.2 ± 3.6	1530 ± 350	47900 ± 13k	0.944	-5.6%	36.9%
Kr	133 ± 9	3740 ± 800	139500 ± 30500	0.961	-3.9%	34.3%
Xe	298.5 ± 26.5	11400 ± 2500	551500 ± 123500	0.983	-1.7%	36.9%
Li	1385 ± 5	80400 ± 1500	6775k ± 275k	0.919	-8.1%	3.7%
K	4k ± 30	392k ± 8k	45800k ± 1800k	1.014	1.4%	4.4%
Rb	4690 ± 50	530k ± 14k	69100k ± 2700k	1.029	2.9%	5.2%
Cs	6665 ± 35	904k ± 46k	131500k ± 8500k	1.069	6.9%	9.2%

Table 4.3: 3-Body Dispersion Combining Rule

Atom	α_1	C_6	C_9	rule	Δ	err
H	4.5	6.49 ± 0.02	7.21 ± 0.01	1.013	1.3%	0.6%
He	1.38	1.46 ± 0.02	0.49 ± 0.002	1.016	1.6%	1.5%
Ne	2.66	6.88 ± 0.4	4.25 ± 0.14	1.044	4.4%	9.4%
Ar	11.1	67.2 ± 3.6	176 ± 4	1.036	3.6%	8.0%
Kr	16.7	133 ± 9	523.5 ± 13	1.036	3.6%	9.6%
Xe	27.3	298.5 ± 26.5	1850 ± 60	1.067	6.7%	13.0%
Li	164	1385 ± 5	56600 ± 0	1.003	0.3%	0.5%
K	291	4k ± 30	287k ± 0	1.014	1.4%	0.9%
Rb	322	4690 ± 50	365k ± 0	1.034	3.4%	1.2%
Cs	409	6665 ± 35	662k ± 0	1.029	2.9%	0.6%

Table 4.4: Mixed Dispersion Combining Rule

Atom		rule	Δ	err
He	Ne	1.009	0.9%	2.6%
He	Ar	1.004	0.4%	3.0%
He	Kr	0.996	-0.4%	4.4%
He	Xe	0.987	-1.3%	5.9%
Ne	Ar	0.997	-0.3%	6.3%
Ne	Kr	0.990	-1.0%	7.6%
Ne	Xe	0.975	-2.5%	9.7%
Ar	Kr	1.000	0.0%	6.0%
Ar	Xe	0.994	-0.6%	7.0%
Kr	Xe	0.998	-0.2%	8.0%
H	Li	0.988	-1.2%	0.7%
H	K	0.954	-4.6%	2.5%
H	Rb	0.942	-5.8%	3.1%
H	Cs	0.952	-4.8%	2.1%
Li	K	1.001	0.1%	0.4%
Li	Rb	1.001	0.1%	0.4%
Li	Cs	0.998	-0.2%	0.3%
K	Rb	1.000	0.0%	0.9%
K	Cs	0.999	-0.1%	0.6%
Rb	Cs	1.001	0.1%	0.6%

Atom		rule	Δ	err
H	He	0.995	-0.5%	0.4%
H	Ne	1.011	1.1%	1.2%
H	Ar	1.008	0.8%	1.5%
H	Kr	1.011	1.1%	2.1%
H	Xe	1.020	2.0%	2.9%
He	Li	0.971	-2.9%	1.7%
He	K	0.898	-10.2%	4.5%
He	Rb	0.879	-12.1%	6.3%
He	Cs	0.914	-8.6%	4.9%
Ne	Li	0.972	-2.8%	2.2%
Ne	K	0.884	-11.6%	5.6%
Ne	Rb	0.860	-14.0%	8.1%
Ne	Cs	0.909	-9.1%	7.4%
Ar	Li	0.985	-1.5%	1.7%
Ar	K	0.921	-7.9%	3.9%
Ar	Rb	0.902	-9.8%	5.7%
Ar	Cs	0.931	-6.9%	4.4%
Kr	Li	0.983	-1.7%	1.5%
Kr	K	0.926	-7.4%	3.9%
Kr	Rb	0.909	-9.1%	5.4%
Kr	Cs	0.934	-6.6%	4.0%
Xe	Li	1.002	0.2%	1.5%
Xe	K	0.947	-5.3%	3.9%
Xe	Rb	0.929	-7.1%	5.3%
Xe	Cs	0.952	-4.8%	4.0%

4.3.2 Parameter-Fitting Rules

The same equations that relate Quantum-Drude parameters to their polarisabilities and dispersion parameters, can be rearranged into a form that allows us to choose parameters that satisfy those relations. These also have not been published before.

As real atoms behave similar to, but not precisely like Quantum-Drudes only a subset of the relations can be precisely satisfied at the same time. Any given parameterisation is always therefore an approximation for real atoms. These are the three relevant relations, again:

$$\alpha_1 = \frac{q^2}{\mu\omega^2}, \qquad C_6 = \tfrac{3}{4}\alpha_1\alpha_1\hbar\omega,$$
$$\alpha_2 = \tfrac{3}{4}\left(\tfrac{\hbar}{\mu\omega}\right)\alpha_1, \qquad C_8 = 5\left(\tfrac{\hbar}{\mu\omega}\right) \times C_6.$$

The easiest parameter to find is ω, using the dipole dispersion relation,

$$\omega = \frac{1}{\hbar}\frac{4C_6}{3\alpha_1^2}.$$

Next, m can be found using *either* a polarisation relation, or a dispersion relation,

$$m = \frac{\hbar}{\omega}\frac{3\alpha_1}{4\alpha_2} \qquad \text{or} \qquad m = \frac{\hbar}{\omega}\frac{5C_6}{C_8}.$$

Finally, q can be found from the dipole polarisability,

$$q = \pm\sqrt{m\omega^2\alpha_1}.$$

As the mobile part of an atom's charge distribution are the negative electrons, one expects that sign of q will also be negative.

4.4 Complex Interactions of the Quantum Drude

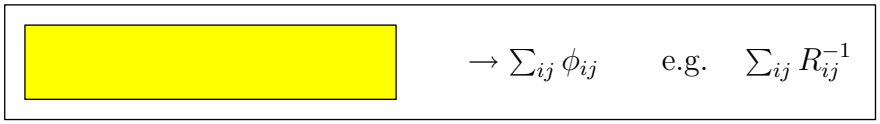
A method for generating successive many-body and multipole interactions of the Quantum Drude, is described towards the end of the DMC chapter (section 6.2.3 on pg. 107), where it comes from iterative perturbations of the Jastrow trial wavefunction of an assembly of Quantum Drudes. The detailed derivation is left until then, in order to maintain the flow of argument in this chapter.

The method is specific to the harmonic nature of Quantum Drudes, and as such, is not general to all quantum polarisable charge distributions, but for Quantum Drudes it is a convenient and quite powerful method. It is new and has not previously been published.

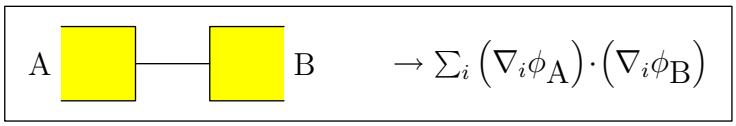
Below, in fig. 4.2 we will show a representation of the quantum energy terms in a diagrammatic expansion to fourth order. The prefactors are complicated sums that would be critical for a calculation, but here we leave them out, in order to give an more intuitive feel for the expansion. First we need to explain the rules of the diagram (next page).

4.4.1 Diagram Rules for Quantum Drudes

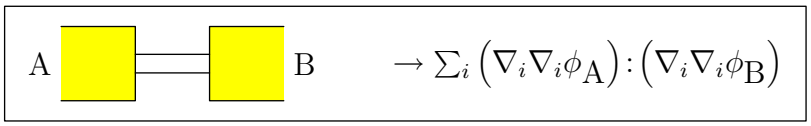
Each yellow box represents an coulombic interaction between two molecules, with an implied summation index at each end.



Each connecting line represents a 'contraction' that comes iteratively out of the 'kinetic' $\left(\nabla^2\right)$ operator in the Schrödinger equation. As $\nabla^2 R^{-1} = 0$, so this has to be between two different instances of the coulomb interaction.

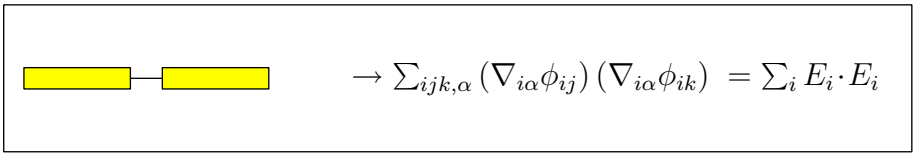


Two lines coming out of a box at the same end indicates that both contractions share the same particle index, but different 'geometric' indices. For example:

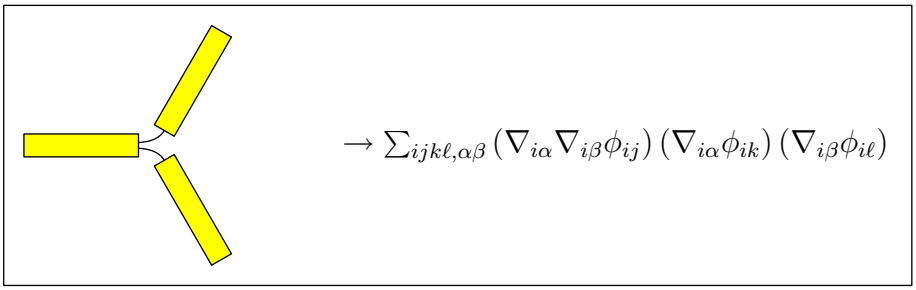


Single-connected box ends therefore represent dipole-interactions. Double and triple connections represent quadrupole and octopole interactions respectively. Unconnected ends represent monopole interactions; these represent the sources of any permanent fields. Finally, the power in R can be easily calculated as (-1) per box (R^{-1} coulomb interactions) plus (-2) per connector (∇^2 contractions).

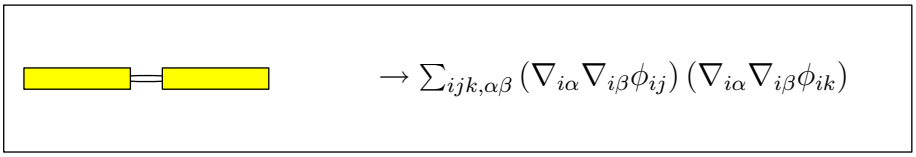
Figure 4.1: QDO example interaction-diagrams



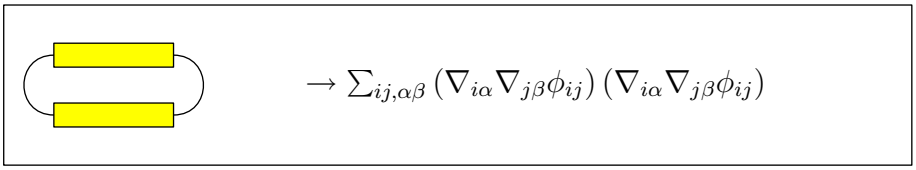
(a) R^{-4} (α_1 or $\alpha_{\alpha\beta}$) dipole-dipole polarisation.



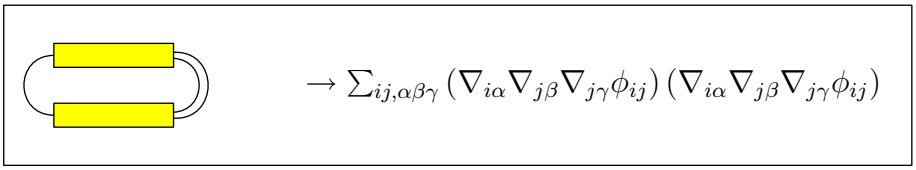
(b) R^{-7} ($B_{\alpha\beta\gamma\delta}$) dipole-dipole-quadrupole polarisation.



(c) R^{-6} (α_2 or $C_{\alpha\beta\gamma\delta}$) quadrupole polarisation.



(d) R^{-6} (C_6) dipole-dipole dispersion.



(e) R^{-8} (C_8) dipole-quadrupole dispersion.

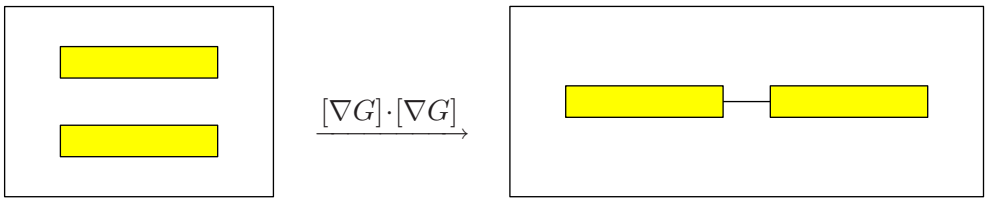
4.4.2 Iterative Energy Expansion

The diagrammatic energy expansion in figure 4.2 was generated by the iterative method described later in the DMC chapter (section 6.2.3 on pg. 107). If prefactors are neglected, that method simplifies to:

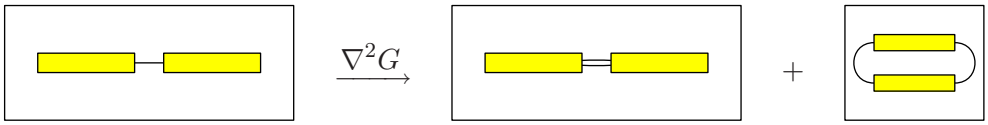
$$\begin{aligned} G_{n,m} \quad &\approx \tfrac{1}{2}\sum_{\nu=1}^{n-1}\sum_{\mu=0}^{m}[\nabla G_{\nu,\mu}]\cdot[\nabla G_{n-\nu,m-\mu}] \quad \text{(terms to left \& those above them)} \\ &\quad + \tfrac{\alpha}{2q^2}\hbar\omega\nabla^2 G_{n,m-1} \qquad\qquad\qquad\qquad\quad \text{(terms above)}, \end{aligned}$$

and the energies $E_{n,m}$ are simply the constant components of the $G_{n,m}$ that can be removed at each step as constants of integration.

The terms created via $[\nabla G]\cdot[\nabla G]$ involved joining two existing diagrams with a new line to form a new diagram (which must be legal).



The terms created via $\nabla^2 G$ add a new connector to an existing diagram to form a new diagram (which also must be legal). This can create loops, which create dispersion terms, akin to "bubble diagrams" which represent the self-interactions of a field in particle physics (in our case the quantum field is made up of discrete Quantum Drude Oscillators that can be moved in space and interact non-locally); or parallel connectors, for example double connectors which represent quadrupole-quadrupole interactions, or triple connectors which represent octopole-octopole interactions.



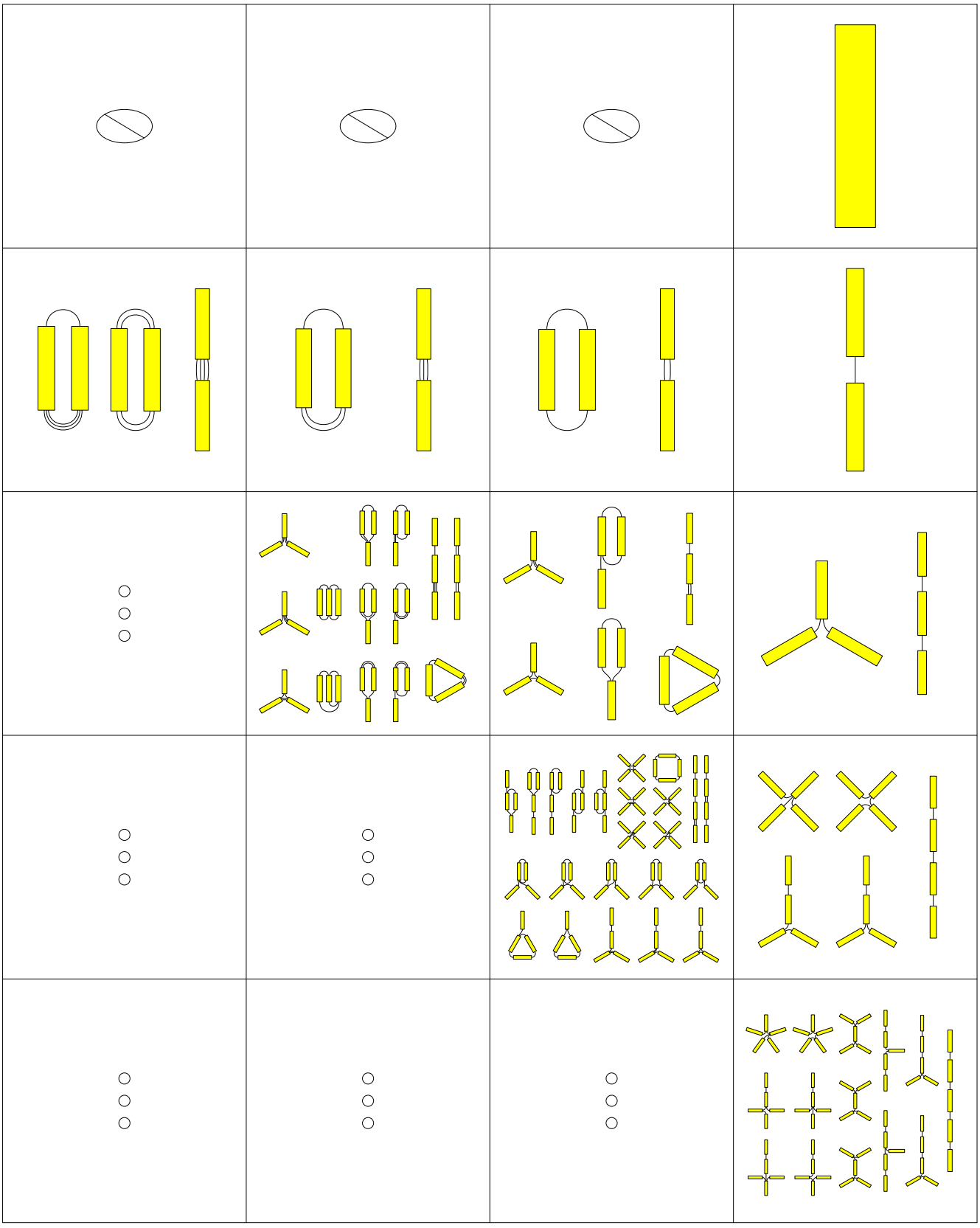


Figure 4.2: QDO many-body, multipole, full diagrammatic energy-expansion.

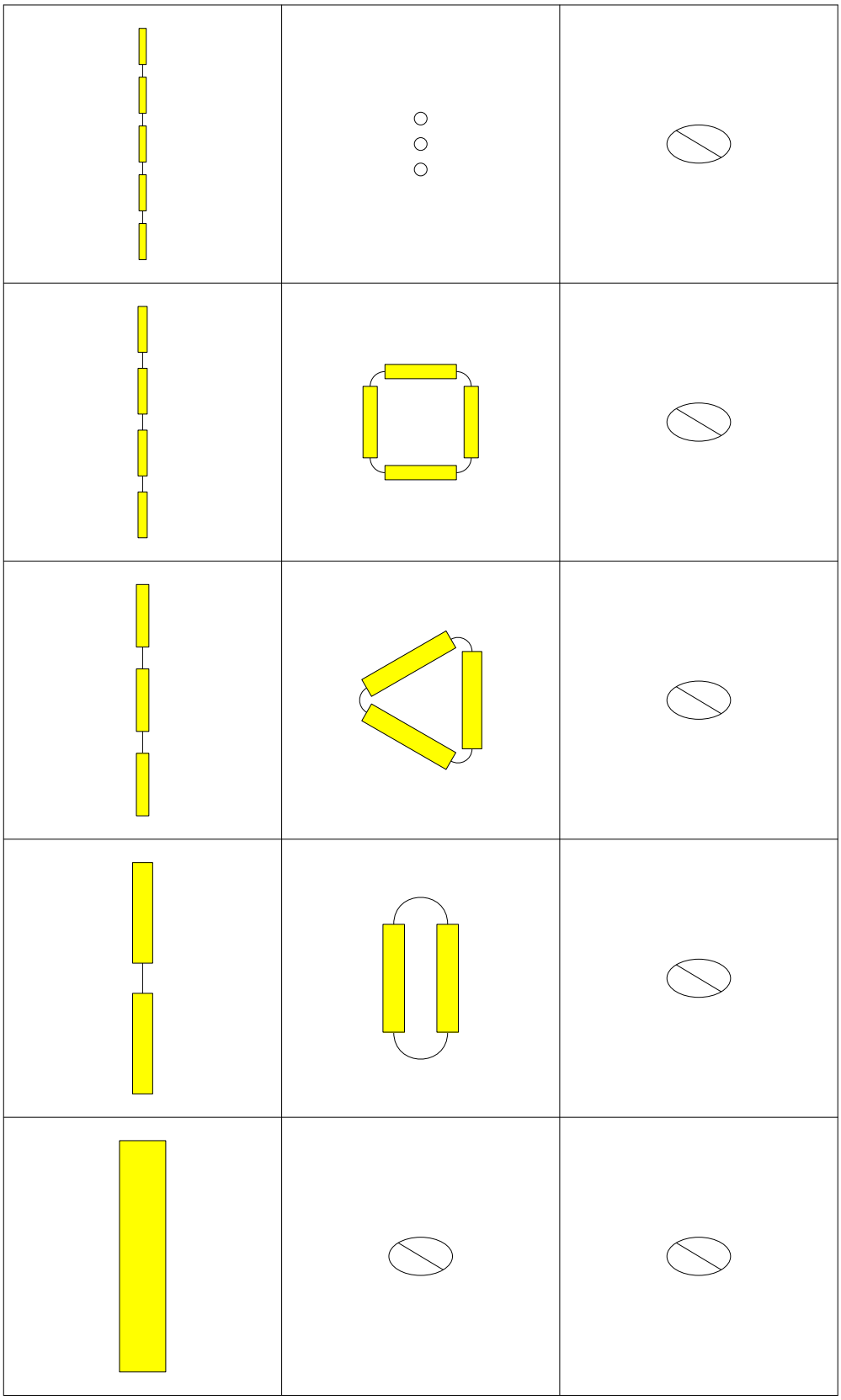


Figure 4.3: QDO many-body, dipole-limit diagrams.

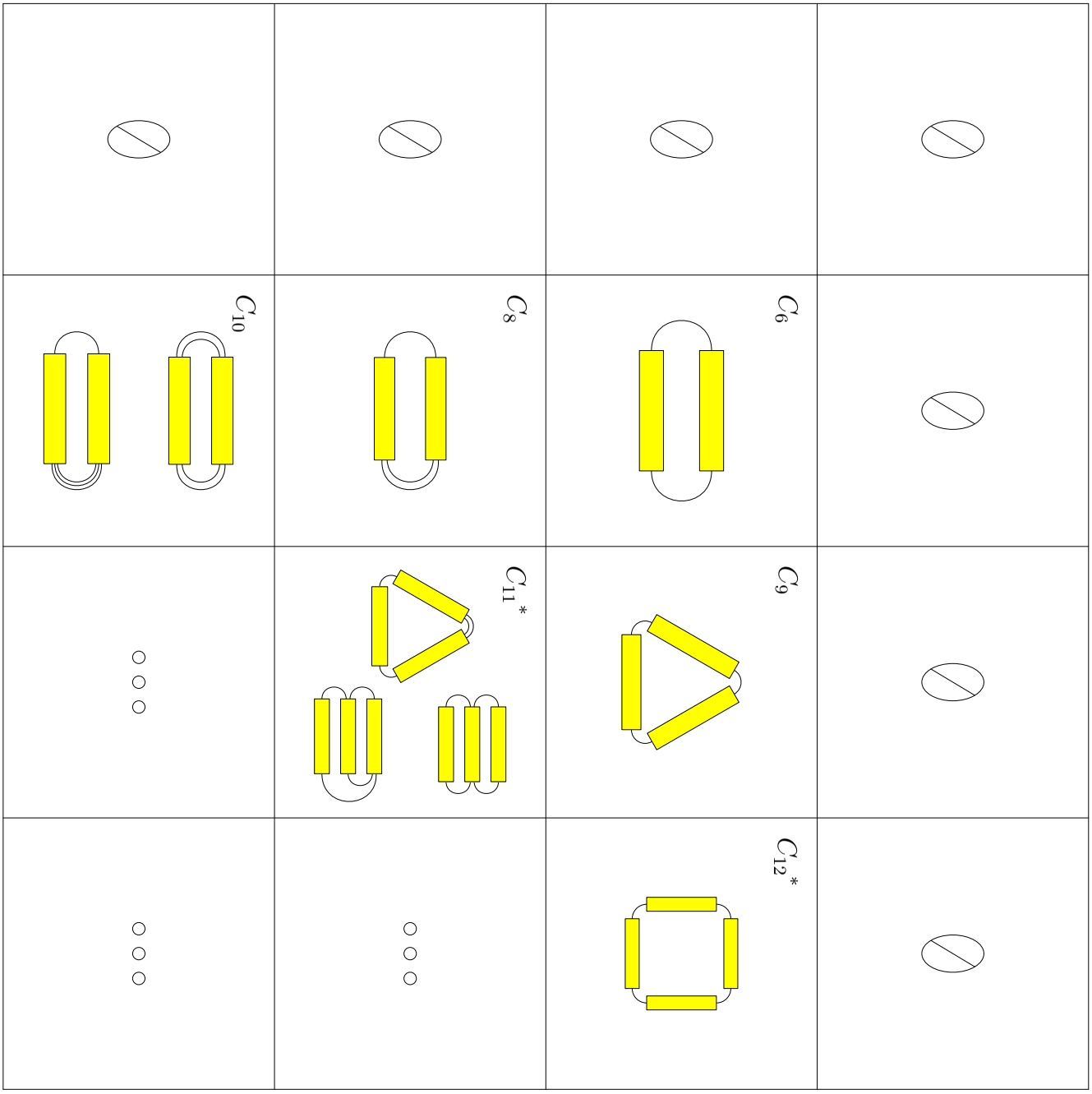


Figure 4.4: QDO many-body, dispersion-only diagrams.

4.4.3 What's in the Expansion?

It is important to note that these diagrams represent the interactions experienced by *Quantum Drudes*. Real atoms and molecules could have even more interactions, as we will discuss below.

In the dipole limit of the Quantum Drude model (Fig. 4.3), there is a very simple series of dipole interactions, due to the simple quadratic nature of the perturbing potential. These are the dipole-limit terms we showed the previous chapter. In the *classical* limit of the Drude model , only the dipole limit exists (only α_1 is non-zero; $\alpha_2 = \alpha_3 = 0$), but there is also no dispersion ($C_6 = C_8 = C_{10} = C_9 = 0$). Thus the classical model captures dipole polarisation, which has a recursive effect (many-body), but nothing else

But the full Quantum Drude model (fig. 4.2) is much richer, and captures many more of the kind of interactions that would be present between real molecules. As we showed above, real atoms behave very much like Quantum Drudes, even if in a less idealised way (there is more on the limitations of the Quantum Drude model in the next section). The expansion allows us to see just how many such interactions there are as one seeks each new level of accuracy in perturbation theory.

In environments where there are no permanent fields, the Quantum Drude model captures many-body dispersion, including all multipoles (fig. 4.4). Where permanent fields do exist , the picture is far more complicated (fig. 4.2); not only do permanent fields "polarise the dispersion loops", but they also create hybrid polarisation-dispersion terms that are difficult to describe.

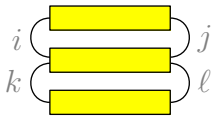
It should now be clear how great an advantage the Quantum Drude model has *if* it can be treated in a way that allows us to avoid having to write down and then compute all of these terms explicitly.

C_{11} would be an ambiguous coefficient

Looking at the dispersion diagrams (fig. 4.4), specifically at the 3rd row, 3rd column, we see that there are three different diagrams that would lay claim to a coefficient named C_{11}. The first is the a dipole-dipole-quadrupole 3-body dispersion term, closely related to the Axilrod-Teller-Muto term, which has been calculated by Bell[Bel70]. However, there are also two different 2-body terms. They are more difficult to name but, together, they relate to the dipole-dipole-quadrupole hyperpolarisability B in the same way that the dipole dispersion C_6 relates to the dipole polarisability α (more detail of these quantities follows in the next section). From these dispersion diagrams, one can see that there is a pragmatic reason to halt the two body dispersion expansion at C_{10}; the next term is C_{11}, not C_{12}. If C_{11} is to be neglected, then there is no point in including C_{12} and higher contributions, including 4-body terms, and so the expansion typically ends with C_{10}.

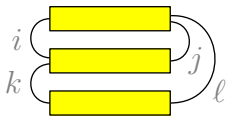
The method which we described in section 4.2 uses 2nd-order perturbation theory, and generates C_6, C_8, C_{10} $\ldots C_{2n+4}$. However, these C_{11} terms would require careful 3rd-order perturbation theory. Further contractions would produce C_{13}, C_{15} and so on. One advantage of the Quantum Drude diagrammatic expansion is thus that it allows us to see more easily which terms could exist. Further terms in the expansion would expose fourth- and nth-order terms. The 3rd and 4th order two-body terms had been calculated for the hydrogen atom by Mitroy and Bromley from work previously done by Ovsiannikov, Guilyarovskia and Lopatko[MB05, MO05, OM06, OGL88]. The method presented here is simpler and hence more useful for understanding the QDO model, which is the topic of this thesis.

The two diagrams are translated into maths as follows:



$$= \left[\nabla_i \nabla_j \left(\tfrac{1}{R}\right)\right] \left[\nabla_i \nabla_j \nabla_k \nabla_\ell \left(\tfrac{1}{R}\right)\right] \left[\nabla_k \nabla_\ell \left(\tfrac{1}{R}\right)\right]$$

$$\begin{aligned} = \quad & \frac{1}{R^{19}} \left[3R_iR_j - R^2\delta_{ij}\right] \\ & \left[105R_iR_jR_jR_\ell - 15R^2 \left(R_iR_j\delta_{k\ell} + R_iR_k\delta_{j\ell} + R_iR_\ell\delta_{jk}\right.\right. \\ & \left.\left. + R_jR_k\delta_{i\ell} + R_jR_\ell\delta_{ik} + R_kR_\ell\delta_{ij}\right) + 3R^4 \left(\delta_{ij}\delta_{k\ell} + \delta_{ik}\delta_{j\ell} + \delta_{i\ell}\delta_{jk}\right)\right] \\ & \left[3R_kR_\ell - R^2\delta_{k\ell}\right] \qquad = \quad \frac{216}{R^{11}} \end{aligned}$$



$$= \left[\nabla_i \nabla_j \nabla_\ell \left(\tfrac{1}{R}\right)\right] \left[\nabla_i \nabla_j \nabla_k \left(\tfrac{1}{R}\right)\right] \left[\nabla_k \nabla_\ell \left(\tfrac{1}{R}\right)\right]$$

$$\begin{aligned} = \quad & \frac{1}{R^{19}} \left[-15R_iR_jR_\ell + 3R^2 \left(R_i\delta_{j\ell} + R_j\delta_{i\ell} + R_\ell\delta_{ij}\right)\right] \\ & \left[-15R_iR_jR_k + 3R^2 \left(R_i\delta_{jk} + R_j\delta_{ik} + R_k\delta_{ij}\right)\right] \\ & \left[3R_kR_\ell - R^2\delta_{k\ell}\right] \qquad = \quad \frac{72}{R^{11}} \end{aligned}$$

4.5 Limitations and Potential Improvements of the Quantum Drude model

We have just seen that the Quantum Drude model reproduces many of the polarisation and dispersion interactions of real molecules. It is interesting to now look (briefly) at which interactions are not present in the model, as well as sketching future work on the possibilities for including them.

A single Quantum Drude with 3 parameters has spherical symmetry. This means that there are only a few independent components in the polarisability tensors, and even of those it has, it is difficult to fit them to the properties of real atoms. By adding more than one Quantum Drude, with 3 new free parameters per Quantum Drude, it may be possible to improve the fit. For example, by adding a second Quantum Drude, that doesn't interact with the first, at the same site, to the Xenon model, it would be possible to fit more multipole/hyperpolarisabilities and dispersion coefficients.

For two Quantum Drudes on the same site, polarisabilities simply add, but dispersion also has cross-terms between the different Quantum Drudes. For example consider an atom with two quantum drudes A and B on the same site, then $\alpha = \alpha_A + \alpha_B$. Now imagine it is interacting with another identical atom, then its first dispersion coefficient is $C_6 = C_{6,AA} + C_{6,AB} + C_{6,BA} + C_{6,BB}$.

By fitting Quantum Drudes at different sites on the same molecule, it would be possible to model anisotropic responses, *provided that we allow the Quantum Drudes to interact*. The following table shows the number of components that are independent in various polarisability response tensors, for 4 different symmetry point groups. α denotes the dipole-polarisability (which we have been referring to as α_1. β and γ denote the double-dipole and triple-dipole polarisabilities respectively. $A_{\alpha,\beta\gamma}$ is the dipole-quadrupole polarisability, $B_{\alpha\beta,\gamma\delta}$ is the dipole-dipole-quadrupole polarisability, and $C_{\alpha\beta,\gamma\delta}$ is the quadrupole-quadrupole polarisability (which we have been referring to as α_2).

A Quantum Drude has spherical symmetry, but the water molecule (which we will certainly be interested in) has only C_{2v} symmetry and so the Quantum Drude, having only 1 independent component to the dipole polarisability, can in general only be an isotropic approximation to that of the anisotropic water molecule (but fortunately it is a reasonably good approximation). The polarisability of water requires 3 components: 1 for each axis in the molecule's basis. By contrast, the symmetry-less group C_1 requires

an extra 3 to define the dipole-polarisation basis relative to the molecule's basis; 6 in total. There are various ways of arranging 3 or more Quantum Drudes to match the C_{2v} symmetry of water, but an intermediate possibility is to place 2 Quantum Drudes on the frame of the molecule (for example on the O-H bonds). Such an arrangement would have $D_{\infty h}$ (roughly, cylindrical) symmetry. We include each of these symmetry groups in the table 4.5 below.

Table 4.5: The number of Independent Elements in Polarisabilities [Buc67]

Symmetry	$\alpha_{\alpha\beta}$	$\beta_{\alpha\beta\gamma}$	$\gamma_{\alpha\beta\gamma\delta}$	$A_{\alpha,\beta\gamma}$	$B_{\alpha\beta,\gamma\delta}$	$C_{\alpha\beta,\gamma\delta}$
Sphere	1	0	1	0	1	1
$D_{\infty h}$	2	0	3	0	4	3
C_{2v}	3	3	6	4	9	6
C_1	6	10	15	15	30	15

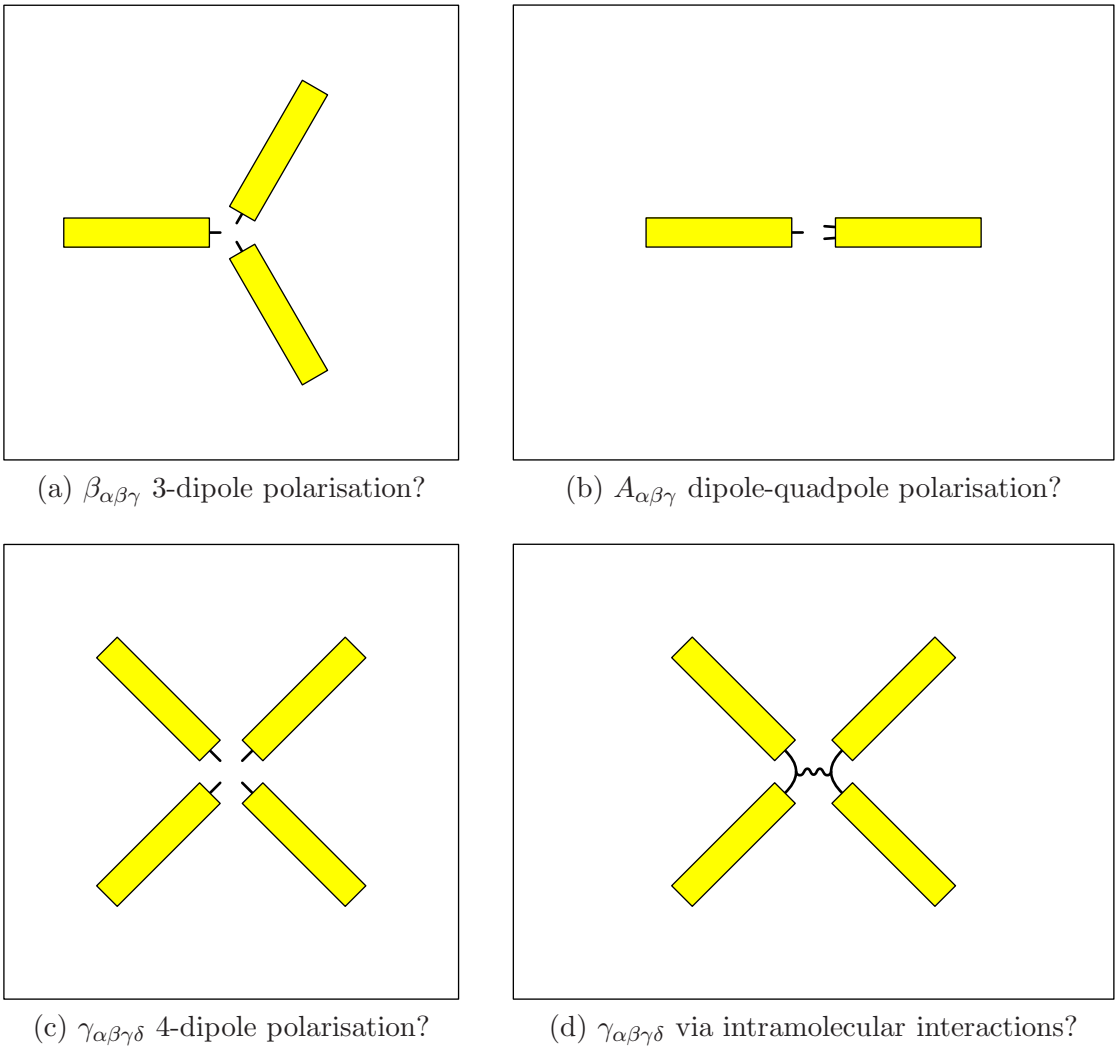
Note that some of the tensors disappear entirely for a spherically symmetric species like a single Quantum Drude. For example,the hyperpolarisability β disappears, as does the dipole-quadrupole polarisability A (although there could be a dipole-dipole-quadrupole polarisability B and a quadrupole-quadrupole polarisability C).

Importantly, the diagrammatic expansion technique correctly predicts that there are no diagrams that correspond to β or A (see fig. 4.2). It also predicts that there is no diagram for γ, while there are diagrams that correspond to α (fig. 4.1a), B (fig. 4.1b), and C (fig. 4.1c).

To see why, consider what the diagrams would have to look like: β (fig. 4.5a) is a dipole-dipole-dipole hyper polarisability. Its diagram would require 3 dipole field representations converging on one atom, which means 3 yellow bars each with a single black line coming out of it. Then A (fig. 4.5b), the dipole-quadrupole polarisabilty, implies 2 yellow boxes, one with a single line and the other having a double line. Clearly it is not possible to connect either of these in a single diagram using our rules. Finally γ (fig. 4.5c), the dipole-dipole-dipole-dipole hyperpolarisability implies 4 yellow boxes, each with a single line coming from it. This can be connected into two diagrams, but not one.

Thus although γ hyperpolarisability is obtainable for spherical symmetry in general, we can use the fact that it lacks a diagram to determine that it does not exist for single Quantum Drude particles (a new result), though it might exist for models involving

Figure 4.5: QDO forbidden interaction-diagrams



(a) $\beta_{\alpha\beta\gamma}$ 3-dipole polarisation?

(b) $A_{\alpha\beta\gamma}$ dipole-quadpole polarisation?

(c) $\gamma_{\alpha\beta\gamma\delta}$ 4-dipole polarisation?

(d) $\gamma_{\alpha\beta\gamma\delta}$ via intramolecular interactions?

more than one QDO per molecule. In fig. 4.5d, the four yellow boxes represent fourfold dipole-field interaction. Each 'normal' line between two of the yellow boxes represents an interaction like a dipole-dipole polarisation on one of the Quantum Drudes, and the additional squiggly line represents a new intramolecular interaction connecting two Quantum-Drudes, coupling their interactions, thus extending the rules of our 'Feynman-style' diagrams. This idea needs more development, which we leave for future work.

It will also take more analysis to discover which additional properties can be obtained using multiple quantum drudes in different symmetry arrangements. As a simplest case example, consider the possibility of having N non-interacting Quantum Drudes on the same site. The polarisabilites will add linearly as N, and the dispersion

coefficients will add like N^2 (for example polarisation $\alpha = \alpha_{\mathrm{A}} + \alpha_{\mathrm{B}}$, but dispersion $C_6 = C_{\mathrm{AA6}} + C_{\mathrm{BB6}} + 2C_{\mathrm{AB6}}$). However, the parameters (q, m, ω) need not add linearly, allowing for greater flexibility in the fitting. We will extend our QDO-specific diagrammatic method to these cases in future work.

Chapter 5

Diffusion Monte Carlo

The Diffusion Monte Carlo (DMC) method [GS71, And75, KW86, HJR94, Cep95, UNR93] is of great interest because it is a computationally efficient way to generate ground state energies for the Quantum Drude model at a fixed nuclear configuration, which we can use for model fitting. Therefore, we want to construct the most accurate, efficient and stable DMC method we possibly can.

We outline the essential ideas of DMC, then begin presenting our new work. We developed a 'flux-balancing' population operator, which is our attempt at a solution to a long-standing problem with the DMC method, the desirability of using a fixed population of walkers, vs the bias which that unfortunately produces. Our new method conserves the population by balancing the birth-rate with the death-rate at each timestep, but in a way that satisfies detailed-balance unlike the standard DMC technique which does not satisfy detailed-balance and exhibits unbounded population fluctuations. We also developed a much more precise diffuse-drift operator.

5.1 Introduction to DMC

At the heart of DMC is an elegant idea. Let $\tau = it$; the Schrödinger equation suggests an imaginary-time evolution operation

$$\mathrm{i}\frac{\partial \Psi}{\partial t} = -\hat{H}\Psi \implies \frac{\partial \Psi}{\partial \tau} = -\hat{H}\Psi \implies \Psi(\tau) = e^{-\tau \hat{H}}\Psi(0).$$

For clarity of presentation, imaginary time will be referred to simply as 'time'.

After a long time, the operator has the effect of promoting the states of lowest energy,

eventually causing the ground-state to dominate.

$$\begin{aligned}
\text{let} \quad \Psi(0) &= \sum_i c_i \Psi_i \qquad \text{be some arbitrary initialisation} \\
\Psi(\tau) &= e^{-\tau \hat{H}}\ \Psi(0) \\
&= \sum_i e^{-\tau E_i} c_i \Psi_i \\
&= e^{-\tau E_0} \left(c_0 \Psi_0 + \sum_{i>0} e^{-\tau \Delta_i} c_i \Psi_i \right), \qquad \text{where} \quad \Delta_i = E_i - E_0, \\
\Psi(\tau) &\rightarrow e^{-\tau E_0}\ c_0 \Psi_0 \qquad \text{as} \quad \tau \rightarrow \infty.
\end{aligned}$$

However, the long time solution either grows without bound ($E_0 < 0$) or vanishes ($E_0 > 0$).

Normalisation

If we include the normalisation into the operator, we get,

$$\Psi(\tau) = e^{-\tau\left(\hat{H} - E_0\right)}\ \Psi(0).$$

The ground-state energy E_0 is a constant which gives a stable long time solution, and it is useful to think of it as being part of the potential operator V.

Discretisation

The operator can be discretised into small time-steps $\tau = N\Delta\tau$:

$$\Psi(\tau) = \left[e^{-\Delta\tau\left(\hat{H} - E_0\right)} \right]^n \Psi(0).$$

Operator Splitting

In the following, we assume $\hat{H} = \hat{T} + V$. The discretised operator can in turn be approximately decomposed into two parts.

$$e^{-\tau \hat{H}} \approx e^{-\tau \hat{T}} e^{-\tau (V - E_0)} \qquad \left(= e^{-\tau\left(\hat{T} + V - E_0\right) + \frac{\tau^2}{2}[\hat{T}, V]} \right).$$

In practice, we use a symmetric decomposition, which eliminates the τ^2 errors (because [T,V]≡-[V,T]). This defines an approximate hamiltonian $\hat{\bar{H}}$ and its normalisation $\bar{E}$.

$$\begin{aligned} e^{-\tau\left(\hat{\bar{H}}-\bar{E}\right)} &\equiv e^{-\tau(V-E_0)/2}e^{-\tau\hat{T}}e^{-\tau(V-E_0)/2} \qquad \left(= e^{-\tau\left(\hat{T}+V-E_0\right)+\mathcal{O}\left(\tau^3\right)}\right) \\ &\approx e^{-\tau\left(\hat{H}-E_0\right)}+\mathcal{O}\left(\tau^3\right). \end{aligned}$$

The first is the diffusion operator;

$$e^{-\tau\hat{T}} = e^{\tau\frac{\hbar^2}{2m}\nabla^2} \implies \frac{\partial\Psi}{\partial\tau} = \frac{\hbar^2}{2m}\nabla^2\Psi,$$

which is the diffusion equation with constant $D = \frac{\hbar^2}{m}$, but note that this interpretation assumes that Ψ is positive definite. This limits it to wavefunctions that have no nodes. It can be adapted to other wavefunctions that have nodes, by breaking up the space into separate regions and using a fixed-node approximation.

The second is the logistic-growth/population-growth operator;

$$\begin{aligned} e^{-\tau V} \implies \qquad & \frac{\partial\log\Psi}{\partial\tau} = \bar{E}-V, \\ \text{or} \quad & \frac{\partial\Psi}{\partial\tau} = \left(\bar{E}-V\right)\Psi, \end{aligned}$$

which are equations for exponential biological population growth or decline, where the '$-V$' would be a measure of comparative biological-fitness. Thus the normalisation value $\bar{E}$ divides space into regions where the population grows and regions where it shrinks.

The approximate operator $\hat{\bar{H}}$ gives us an approximate wavefunction $\bar{\Psi}$ and an approximate energy $\bar{E}$. After a long time, it picks out the ground-state wavefunction, but with some $\mathcal{O}\left(\Delta\tau^3\right)$ error that comes from commmutators.

$$\begin{aligned} \bar{\Psi}(\tau) &= e^{-\tau\left(\hat{\bar{H}}-\bar{E}\right)}\bar{\Psi}(0) \\ &= \Psi(\tau)+\mathcal{O}\left(\Delta\tau^3\right) \\ &\to c_0\Psi_0+\mathcal{O}\left(\Delta\tau^3\right) \qquad \text{as} \quad \tau\to\infty. \end{aligned}$$

Thus the ground-state wavefunction can be represented by a Monte-Carlo process; a population of 'walkers' or 'samplers' which diffuse through phase-space, breeding in areas where the classical potential is low, and expiring where the classical potential is high.

5.1.1 DMC with Importance Sampling

Efficiency can be improved by importance sampling. This has the effect of changing the diffusion operator into a 'diffusion-with-flow' or 'diffusion-and-drift' operator. This moves some of the influence of the potential/breeding operator into this 'kinetic' operator. Broadly this means that walkers tend to drift from regions of high potential to regions of low potential. The remaining potential/breeding operator still plays a role in the fine detail of this new distribution, but becomes less important. This is good because the greatest inefficiences come from that operator. As a bonus, importance sampling provides an estimator of the energy that has lower variance.

The down-side of DMC, is that it explores only the wavefunction and not the density. This means that there is no Hellman-Feynman theorem, so the only quantity that can be measured with only finite timestep error is the energy (without advanced methods).

Use of Trial Wavefunctions

The density operator projects out the ground-state wavefunction Ψ_0, but it turns out we can improve efficiency by sampling a distribution more like the density function $\rho(x) = |\Psi_0(x)|^2$. We can do this by weighting the true groundstate wavefunction Ψ_0 with a trial function Ψ_T. In what follows, we will refer to Ψ_0 simply as Ψ. Instead of letting $f(x,\tau) \to \Psi(x)$, we let $f(x,\tau) \to \Psi_0(x)\Psi_\mathrm{T}(x)$, where Ψ_0 is the true ground state, and Ψ_T is an trial estimate of it [And75]. The trial wavefunction Ψ_T is often described as a Jastrow wavefunction, after Jastrow[Jas55], who first used a variational form, in first-order perturbation theory, with potential-like terms plus parameters to be tuned to minimise the upper bound (1st-order perturbation theory) of the ground state energy. DMC goes beyond first order perturbation theory, but still makes use of Jastrow's trial wavefunctions.

We can assume that Ψ_0 is positive definite. This is true for non-identical particles in the ground state, including Drude electrons because each is uniquely attached to an atom, so this assumption always holds for our work; Quantum-Drudes are *distinguishable* particles. We can then interpret $f(x) = \Psi_0\Psi_\mathrm{T}$ as a probability density

function.

$$\begin{aligned}\frac{\partial \Psi}{\partial \tau} &= \frac{1}{\Psi_{\mathrm{T}}}\frac{\partial f}{\partial \tau},\\ -\hat{H}\Psi &= -\hat{H}\left(\frac{f}{\Psi_{\mathrm{T}}}\right) = -\hat{T}\left(\frac{f}{\Psi_{\mathrm{T}}}\right) - \frac{1}{\Psi_{\mathrm{T}}}\phi f.\end{aligned}$$

Now we meed to evaluate the effect of the kinetic operator $\hat{T}$

$$\begin{aligned}\nabla\frac{f}{\Psi_{\mathrm{T}}} &= \frac{\nabla f}{\Psi_{\mathrm{T}}} - \frac{f\nabla\Psi_{\mathrm{T}}}{\Psi_{\mathrm{T}}^2},\\ \nabla^2\frac{f}{\Psi_{\mathrm{T}}} &= \frac{\nabla^2 f}{\Psi_{\mathrm{T}}} - \frac{2\nabla f\nabla\Psi_{\mathrm{T}}}{\Psi_{\mathrm{T}}^2} - \frac{f\nabla^2\Psi_{\mathrm{T}}}{\Psi_{\mathrm{T}}^2} + \frac{2f\nabla\Psi_{\mathrm{T}}\nabla\Psi_{\mathrm{T}}}{\Psi_{\mathrm{T}}^3},\\ -\hat{T}\frac{f}{\Psi_{\mathrm{T}}} &= \frac{\hbar^2}{2m}\left(\frac{\nabla^2 f}{\Psi_{\mathrm{T}}} - \frac{2\nabla f\nabla\Psi_{\mathrm{T}}}{\Psi_{\mathrm{T}}^2} - \frac{f\nabla^2\Psi_{\mathrm{T}}}{\Psi_{\mathrm{T}}^2} + \frac{2f\nabla\Psi_{\mathrm{T}}\nabla\Psi_{\mathrm{T}}}{\Psi_{\mathrm{T}}^3}\right).\end{aligned}$$

Then plug back in:

$$\begin{aligned}\frac{\partial f}{\partial \tau} &= -\Psi_{\mathrm{T}}\hat{H}\frac{f}{\Psi_{\mathrm{T}}}\\ &= \frac{\hbar^2}{2m}\left(\nabla^2 f - 2f\nabla\log\Psi_{\mathrm{T}}\nabla\log\Psi_{\mathrm{T}} - 2\nabla f\nabla\log\Psi_{\mathrm{T}} - \frac{f\nabla^2\Psi_{\mathrm{T}}}{\Psi_{\mathrm{T}}}\right) - \frac{f\phi\Psi_{\mathrm{T}}}{\Psi_{\mathrm{T}}}.\end{aligned}$$

Then note

$$\nabla.\left(\nabla - 2\nabla\log\Psi_{\mathrm{T}}\right)f = \nabla^2 f - 2f\nabla\log\Psi_{\mathrm{T}}\nabla\log\Psi_{\mathrm{T}} - 2\nabla f\nabla\log\Psi_{\mathrm{T}} - \frac{2f\nabla^2\Psi_{\mathrm{T}}}{\Psi_{\mathrm{T}}}.$$

So

$$\begin{aligned}\frac{\partial f}{\partial \tau} &= \frac{\hbar^2}{2m}\nabla\cdot\left(\nabla - 2\nabla\log\Psi_{\mathrm{T}}\right)f + \frac{f\hbar^2}{2m}\frac{\nabla^2\Psi_{\mathrm{T}}}{\Psi_{\mathrm{T}}} - \frac{f\phi\Psi_{\mathrm{T}}}{\Psi_{\mathrm{T}}}\\ &= \frac{\hbar^2}{2m}\nabla\cdot\left(\nabla - 2\nabla\log\Psi_{\mathrm{T}}\right)f - \frac{f\hat{H}\Psi_{\mathrm{T}}}{\Psi_{\mathrm{T}}}.\end{aligned}$$

Now add the normalisation constant $\bar{E}$:

$$\begin{aligned}\frac{\partial f}{\partial \tau} &= \frac{\hbar^2}{2m}\nabla\cdot(\nabla - 2\nabla \log \Psi_{\mathrm{T}})\, f - \left(\frac{\hat{H}\Psi_{\mathrm{T}}}{\Psi_{\mathrm{T}}} - \bar{E}\right) f \\ &= -\hat{A}f - \hat{B}f,\end{aligned}$$

$$\begin{aligned}\text{where}\quad \hat{A} &= \frac{\hbar^2}{2m}\nabla\cdot(\nabla - 2\nabla \log \Psi_{\mathrm{T}})\,, \\ \hat{B} &= \frac{\hat{H}\Psi_{\mathrm{T}}}{\Psi_{\mathrm{T}}} - \bar{E}.\end{aligned}$$

The solution is

$$\begin{aligned}f(\tau) &= e^{\tau(\hat{A}+\hat{B})}\, f(0), \qquad \text{which we approximate as} \\ &= e^{\frac{\tau\hat{B}}{2}} e^{\hat{A}} e^{\frac{\tau\hat{B}}{2}}\, f(0) + O(\tau^3).\end{aligned}$$

Properties of the Operators

The DMC operators preserve the ground-state solution.

$$e^{\tau(\hat{A}+\hat{B})}\, \Psi_0\Psi_{\mathrm{T}} = \Psi_0\Psi_{\mathrm{T}}.$$

They are closely related, but not identical, to a density matrix (see later chapter).

$$\begin{aligned}&\langle r|e^{\tau(\hat{A}+\hat{B})}|r'\rangle \\ &\qquad = \frac{\Psi_{\mathrm{T}}(r)}{\Psi_{\mathrm{T}}(r')}\langle r|e^{-\tau\hat{H}}|r'\rangle \\ &\qquad = \frac{\Psi_{\mathrm{T}}(r)}{\Psi_{\mathrm{T}}(r')} \times \rho(r, r'; \tau).\end{aligned}$$

Finally, if the trial wavefunction is exactly correct, that is $\Psi_{\mathrm{T}} = \Psi_0$,

$$\begin{aligned}\hat{B} = \frac{\hat{H}\Psi_{\mathrm{T}}}{\Psi_{\mathrm{T}}} - E_0 = 0, \\ e^{\tau\hat{A}}\, \Psi_{\mathrm{T}}\Psi_{\mathrm{T}} = \Psi_{\mathrm{T}}\Psi_{\mathrm{T}}.\end{aligned}$$

Diffuse/Flow Operator

$$\begin{aligned}\frac{\partial f}{\partial \tau} &= -\hat{A}f \\ &= \frac{\hbar^2}{2m}\nabla\cdot(\nabla - 2\nabla\log\Psi_{\mathrm{T}})\,f \\ &= \frac{\hbar^2}{2m}\nabla^2 f - \nabla\cdot(2\nabla\log\Psi_{\mathrm{T}} f) \\ &\equiv D\nabla^2 f - \nabla\,(f\vec{v})\,.\end{aligned}$$

The term on the left is the same diffusion term that we seen in the introduction, with $D = \frac{\hbar^2}{2m}$. Now, however, we have another term, which is a 'material derivative' with respect to time. It tells us how the local density of a fluid changes when it is flowing according to the velocity field $\vec{v}(x) = \frac{\hbar^2}{m}\nabla\log\Psi_{\mathrm{T}}$. Thus to model this operator, our population of samplers must also be 'flowing' according to velocity field $\vec{v}(x)$.

If we make a further simplifying assumption that $\nabla\log\Psi_{\mathrm{T}}$ (a position operator) is locally a constant, then it commutes with the momentum operator $\hat{p}$ and so we can use Gaussian integration to calculate how the operator $\hat{A}$ propagates the wavefunction from a point r to point r'.

$$\begin{aligned}P(x\to x') &= \langle x|e^{-\tau\hat{A}}|x'\rangle \\ &= \int_p \langle x|e^{-\tau\hat{A}}|p\rangle\langle p|x'\rangle \\ &= \int_p \langle x|p\rangle e^{-\tau\hat{A}}\langle p|x'\rangle \quad \text{as } |p\rangle \text{ is an eigenstate of } \hat{A} \\ &= \int_p dp' e^{\frac{i}{\hbar}p.x'} e^{-\tau\hat{A}} e^{-\frac{i}{\hbar}p.x}.\end{aligned}$$

Note that

$$\begin{aligned}-\hat{A}e^{\frac{-i}{\hbar}p.x'} &= \left(\frac{\hbar^2}{2m}\nabla^2 - v\cdot\nabla - \nabla\cdot v\right)e^{-\frac{i}{\hbar}p.x'} \\ &= -e^{-\frac{i}{\hbar}p.x'}\left(\frac{p^2}{2m} - \frac{ip}{\hbar}v(x') + \nabla\cdot v(x')\right).\end{aligned}$$

For a quadratic wavefunctions, the term $(\nabla\cdot v)$ is a constant that comes out in the wash (normalisation), and we drop it anyway due to our approximation that $\vec{v}$ is locally

constant. Now the propagator becomes

$$\begin{aligned} P(x \to x') &= \int_p e^{\frac{i}{\hbar}p.(x-x')} e^{-\frac{\tau}{2m}\left(p^2 - i\frac{2m}{\hbar}p.v\right)} \\ &= \int_p e^{-\frac{\tau}{2m}\left(p^2 + i\frac{2m}{\hbar\tau}p.(x'-x-\tau v)\right)}. \end{aligned}$$

To evaluate the integral, we need to complete the square in the exponent

$$\begin{aligned} & - \frac{\tau}{2m}\left(p^2 + i\frac{2m}{\hbar\tau}p.(x'-x-\tau v)\right) \\ = & - \frac{\tau}{2m}\left[\left(p + i\frac{m}{\hbar\tau}(x'-x-\tau v)\right)^2 - \left(i\frac{m}{\hbar\tau}(x'-x-\tau v)\right)^2\right]. \end{aligned}$$

Now the p terms can be integrated out and dropped as the integral is only a prefactor

$$P(x \to x') = \left(\frac{2m\pi}{\tau}\right)^{\frac{1}{4}} e^{-\frac{\tau}{2m}\left(\frac{m}{\hbar\tau}(x'-x-\tau v)\right)^2},$$

$$P(x \to x') \propto \exp\left\{-\frac{m}{2\hbar^2\tau}\left[x' - (x+\tau v)\right]^2\right\}.$$

This propagator convolves the distribution $f(x;0)$ to a distribution $f(x';\tau)$.

$$\begin{aligned} f(x';\tau) &= \int_x P(x \to x')\, f(x;0) \\ & \int_x \left(\frac{2m\pi}{\tau}\right)^{\frac{1}{4}} \exp\left\{-\frac{m}{2\hbar^2\tau}\left[x' - (x+\tau v)\right]^2\right\} f(x;0). \end{aligned}$$

Differential Reproduction (Population) Operator

Having accounted for the effect of operator $\hat{A}$ on the p.d.f. $f(r)$, we now look at the operator $\hat{B}$:

$$\hat{B} = \left(\frac{1}{\Psi_{\mathrm{T}}}\hat{H}\Psi_{\mathrm{T}} - E\right).$$

Note that $\hat{H}\Psi_{\mathrm{T}}(r)$ can be evaluated in terms of position only; $\hat{B}$ is a pure position operator which modifies the p.d.f. $f(r)$ at a position. This means that it must modify the population of walkers at a point. As the population at a point considered will be 1, and only integer populations are possible, we interpret the value of $\hat{B}$ as the mean

number of walkers that will represent this point after the operator is applied:

$$f_{\text{new}} = e^{-\tau\left(\frac{1}{\Psi_{\text{T}}}\hat{H}\Psi_{\text{T}} - E\right)} f_{\text{old}}.$$

5.1.2 Simulation by Sampling using Walkers

In DMC, the function $f(x,\tau)$ is represented by a set of random-walkers distributed with probability density $f(x,\tau)$. Thus it is our sampling distribution. We have a time-evolution operator, that was mentioned above, which involves a double convolution and a position operator. Let the shorthand $\int_x$ denote an integral over all space.

$$f(\tau) = e^{\frac{\tau\hat{B}}{2}} e^{\hat{A}} e^{\frac{\tau\hat{B}}{2}} f(0) + O(\tau^3),$$

$$\begin{aligned} f(x'';\tau) &= \int_{x'} \left(\frac{2m\pi}{(\tau/2)}\right)^{\frac{1}{4}} \exp\left\{-\frac{m}{2\hbar^2(\tau/2)}\left[x'' - (x' + (\tau/2)v)\right]^2\right\} \\ &\quad \times \exp\left\{-\tau\left(\frac{1}{\Psi_{\text{T}}}\hat{H}\Psi_{\text{T}}(x') - \bar{E}\right)\right\} \\ &\quad \times \int_x \left(\frac{2m\pi}{(\tau/2)}\right)^{\frac{1}{4}} \exp\left\{-\frac{m}{2\hbar^2(\tau/2)}\left[x' - (x + (\tau/2)v)\right]^2\right\}. \end{aligned}$$

Diffuse/Flow of Walkers

The Diffuse/Flow operator is as follows:

$$P(x \to x';\tau) \propto \exp\left\{-\frac{m}{2\hbar^2\tau}\left[x' - (x + \tau v)\right]^2\right\}.$$

This means that for a walker situated at position x, the probability distribution for its location after timestep τ, is $P(x';x,\tau)$. The distribution is a gaussian (normal distribution) with $\sigma = \hbar\sqrt{\tau/m}$ and $\mu = x + \tau v$; the centre is shifted by τv. We assumed that $\vec{v}$ is constant over the width of the distribution, that is $\vec{v}(x+\sigma) \approx \vec{v}(x-\sigma)$. τv is the 'drift' and is taken out of the gaussian and added separately. If Gaussian() is a unitless gaussian-distributed random number with mean $\mu = 0$ and width $\sigma = 1$, then the position after time τ is:

$$x(\tau) = x(0) + \tau v + \Delta x, \qquad \text{where} \quad \Delta x = \hbar\sqrt{\tfrac{\tau}{m}} \times \text{Gaussian}().$$

For example, using the simple harmonic oscillator as a trial wavefunction gives:

$$\Psi_{\mathrm{T}}(r) = e^{-\frac{m\omega}{2\hbar}r^2},$$
$$v(r) = \frac{\hbar^2}{m}\nabla \log \Psi_{\mathrm{T}} = -\hbar\omega x.$$

Differential Reproduction of Walkers

The population operator multiplies the number of walkers at a point by $e^{-\tau\hat{B}}$, where $B = \left(\frac{\hat{H}\Psi_{\mathrm{T}}}{\Psi_{\mathrm{T}}} - \bar{E}\right)$, in a probabilistic sense. This is like the accept/reject step of traditional Monte Carlo techniques.

- For each walker, we calculate the number $e^{-\tau B}$.
- If $e^{-\tau B} > 1$ the walker is copied $\mathrm{int}(e^{-\tau B})$ times into the next ensemble, and an additional one is added with probability $\mathrm{frac}(e^{-\tau B})$.
- If $e^{-\tau B} < 1$ the walker is copied into the next ensemble with probability $e^{-\tau B}$. If it is not copied into the next ensemble, it is deleted.

Unfortunately, every time a new walker is introduced, it sits on top of its parent so that they are correlated until they diffuse apart. Therefore if there is too much fluctuation in the number of walkers, the correlations will reduce the randomness of the sampling. A good trial-wavefunction should minimise this fluctuation.

Population Control

More significantly, the walkers are multiplied in a stochastic manner. This leads to fluctuations in the overall population, even for a large population. It would be convenient to enforce a constant population, but this is tricky because that becomes a way for walkers to influence each other, and that can introduce a measurable bias to the results.

The population operator depends on a normalising parameter $\bar{E}$, which is typically adjusted [AMK00] in the following way.

$$\bar{E}_{(p+1)} = \bar{E}_p + \frac{K}{\tau}\log\left[N_p/N_{(p+1)}\right],$$

where p is the number of projection steps taken. At long time however, the adjustment falls to zero (as $1/p$), at which point it is no longer able to correct for stochastic drift in the population. With this method, there is nothing to stop the population from exploding, or dying out completely. We mention this in constrast to the norm-conserving method we present below.

Ground State Energy: The Estimator

$\frac{\hat{H}\Psi_\mathbf{T}}{\Psi_\mathbf{T}}$ We already showed that $f \to \Psi_\mathrm{T}\Psi_0$ as $\tau \to \infty$. Thus the normalised f after equilibration is

$$f = \frac{\Psi_\mathrm{T}\Psi_0}{\langle\Psi_\mathrm{T}\Psi_0\rangle}.$$

We know that $\hat{H}$ is a Hermitian operator. This means,

$$\begin{aligned}
\langle\Psi_\mathrm{T}|\hat{H}|\Psi_0\rangle &= \langle\Psi_0|\hat{H}|\Psi_\mathrm{T}\rangle, \\
\langle\Psi_\mathrm{T}|\Psi_0\rangle E_0 &= \int_r \Psi_0\hat{H}\Psi_\mathrm{T} \\
&= \int_r \Psi_0\Psi_\mathrm{T}\left[\frac{1}{\Psi_\mathrm{T}}\hat{H}\Psi_\mathrm{T}\right],
\end{aligned}$$

$$E_0 = \int_r \left[\frac{\Psi_0\Psi_\mathrm{T}}{\langle\Psi_\mathrm{T}|\Psi_0\rangle}\right]\left[\frac{\hat{H}\Psi_\mathrm{T}}{\Psi_\mathrm{T}}\right],$$

$$E_0 = \int_r f\ \tfrac{\hat{H}\Psi_\mathrm{T}}{\Psi_\mathrm{T}}.$$

In words, E_0 is equal to the average of $\frac{\hat{H}\Psi_\mathrm{T}}{\Psi_\mathrm{T}}$, over the distribution f.

Therefore we can approximate E_0 by taking a long time average of $\frac{\hat{H}\Psi_\mathrm{T}}{\Psi_\mathrm{T}}$ over all the walkers, after they have had time to equilibrate. $\frac{\hat{H}\Psi_\mathrm{T}}{\Psi_\mathrm{T}}$ is thus an **estimator** for E_0.

$E_\mathbf{pop}$ Another estimator is the value of E used for normalising the population operator, after applying a small correction that comes from estimating how much it favours increasing the population vs decreasing. We found that it has a larger uncertainty than the $\frac{\hat{H}\Psi_\mathrm{T}}{\Psi_\mathrm{T}}$, but we used it as a test of the quality of the simulations.

5.2 Evaluating trial wavefunctions for use in DMC

5.2.1 First Order Perturbation Theory

First order perturbation theory calculations use a suitable trial wavefunction Ψ_{T} (with its corresponding trial Hamiltonian $\hat{H}_{\mathrm{T}}$) to produce an upper bound on the true energy of the ground state energy of the system. It is an upper bound because any trial wavefunction mixes the ground state wavefunction with those of excited states, unless it is exact.

The trial wavefunction will contain a mixture of the true eigenstates of the real system, so the quality of the estimate of E_0 will depend on how close Ψ_{T} is to Ψ_0:

$$\begin{aligned}
\Psi_{\mathrm{T}} &= \sum_i c_i \Psi_i, \qquad \text{where} \quad \sum_i c_i^2 = 1, \\
E_{\text{estimate}} &= \langle \Psi_{\mathrm{T}} | \hat{H} | \Psi_{\mathrm{T}} \rangle \\
&= \sum_i c_i^2 E_i \\
&= E_0 + \Delta E \qquad \geqslant E_0, \qquad \because E_i \geqslant E_0.
\end{aligned}$$

We have two different ways of performing these calculations: DMC minus the population operator, and Metropolis wavefunction sampling.

5.2.2 DMC with no Population Operator (VMC)

Diffusion Monte Carlo is exact in theory, but Variational Monte Carlo is easier to do, and is a way of doing first order perturbation theory:
We find a suitable trial wavefunction Ψ_{T} and a trial Hamiltonian $\hat{H}_{\mathrm{T}}$:

$$\begin{aligned}
\hat{H} &= \hat{H}_{\mathrm{T}} + \hat{H}', \\
E_{\text{estimate}} &= \langle \Psi_{\mathrm{T}} | \hat{H} | \Psi_{\mathrm{T}} \rangle \\
&= \langle \Psi_{\mathrm{T}} | \hat{H}_{\mathrm{T}} + \hat{H}' | \Psi_{\mathrm{T}} \rangle \\
&= \langle \Psi_{\mathrm{T}} | \hat{H}_{\mathrm{T}} | \Psi_{\mathrm{T}} \rangle + \langle \Psi_{\mathrm{T}} | \hat{H}' | \Psi_{\mathrm{T}} \rangle \\
&= E_{\mathrm{T}} + \langle \Psi_{\mathrm{T}} | \hat{H}' | \Psi_{\mathrm{T}} \rangle \\
&= E_{\mathrm{T}} + E', \\
\text{where} \quad E' &\equiv \langle \Psi_{\mathrm{T}} | \hat{H}' | \Psi_{\mathrm{T}} \rangle.
\end{aligned}$$

This is first order perturbation theory. If $\hat{H}'$ is a pure position operator, then

$$\begin{aligned} E' &= \int dx\, \Psi_{\mathrm{T}}(x)\Psi_{\mathrm{T}}(x)H'(x) \\ &= \int dx\, f(x)H'(x), \\ &\quad \text{where} \quad f(x) \equiv \Psi_{\mathrm{T}}(x)\Psi_{\mathrm{T}}(x). \end{aligned}$$

This can be treated in almost exactly the same way as DMC. The same derivation can be followed, as from the start of this chapter. The only differences are:

$$\begin{array}{ccc} \text{DMC} & \rightarrow & \text{VMC}\,, \\ f = \Psi_0\Psi_{\mathrm{T}} & \rightarrow & f = \Psi_{\mathrm{T}}\Psi_{\mathrm{T}}\,, \\ \frac{\partial\Psi}{\partial\tau} = -\hat{H}\Psi & \rightarrow & \frac{\partial\Psi_{\mathrm{T}}}{\partial\tau} = -\hat{H}_{\mathrm{T}}\Psi_{\mathrm{T}}\,, \\ \left[\frac{\hat{H}\Psi_{\mathrm{T}}}{\Psi_{\mathrm{T}}} - E_0\right] \not\equiv 0 & \rightarrow & \left[\frac{\hat{H}_{\mathrm{T}}\Psi_{\mathrm{T}}}{\Psi_{\mathrm{T}}} - E_{\mathrm{T}}\right] \equiv 0\,. \end{array}$$

The implication of this is that we may use exactly the same diffuse/drift step, but there is no need for multiply/contract step. Simulation of VMC is equivalent to that for DMC but without a multiply/contract step.

At every step τ, we use the full Hamiltonian as an energy-estimator,

$$\frac{\hat{H}\Psi_{\mathrm{T}}}{\Psi_{\mathrm{T}}} = E_{\mathrm{T}} + H'(x(\tau)),$$

but to move the walkers, we need only the trial Hamiltonian:

$$\frac{\partial\Psi_{\mathrm{T}}}{\partial\tau} = -\hat{H}_{\mathrm{T}}\Psi_{\mathrm{T}}.$$

Another way of looking at DMC is that the diffuse/drift step gives you a first order perturbation energy estimate, and the (more troublesome) multiply/contract step accurately closes the gap to the true energy (for the ground state).

5.2.3 Metropolis Wavefunction Sampling of $\Psi_{\mathbf{T}}{}^2$

Another option, is to integrate the wavefunction by Monte Carlo sampling of the probability density function.

For a harmonic wavefunction, this is very simple. Because both the wavefunction and probability density function are Gaussian distributions, a new position-state sample can be generated by calling a Gaussian random number generator, and accepted with 100% probability.

$$\begin{aligned} \phi_{\mathrm{T}} &\propto \exp\left\{-\tfrac{m\omega}{2\hbar}x^2\right\} \quad , \qquad \rho \propto \exp\left\{-\tfrac{2m\omega}{2\hbar}x^2\right\} \quad ; \\ x_{\mathrm{new}} &= \sqrt{\tfrac{\hbar}{2m\omega}} \times \mathrm{Gaussian}() \quad , \qquad \text{accept with } P = 1 \quad . \end{aligned}$$

All wavefunctions suitable for Quantum-Drudes will be based on the simple harmonic wavefunction, but for more complicated trial wavefunctions, there is a perturbation $F(x)$. This requires the accept/reject step.

$$\begin{aligned} \phi_{\mathrm{T}} &\propto \exp\left\{-\tfrac{m\omega}{2\hbar}x^2 + F(x)\right\} \quad , \qquad \rho \propto \exp\left\{-\tfrac{2m\omega}{2\hbar}x^2 + 2F(x)\right\} \quad ; \\ x_{\mathrm{new}} &= \sqrt{\tfrac{\hbar}{2m\omega}} \times \mathrm{Gaussian}() \quad , \\ & \quad \text{accept with } P = \min\left[e^{2F(x_{\mathrm{new}})-2F(x_{\mathrm{old}})}, 1\right] \quad . \end{aligned}$$

5.3 Propagators for Large Timesteps in DMC

Earlier we derived the following result for diffusion with flow on the locally-constant gradient of the trial wavefunction:

$$P(r_0 \to r;\tau) \quad \propto \quad \exp\left\{-\frac{m}{2\hbar^2\tau}\left[r - r_0 - \tau v(r_0)\right]^2\right\}, \qquad \text{where} \quad v = \tfrac{\hbar^2}{m}\nabla \log \Psi_{\rm T}.$$

That derivation assumed that $\vec{v}$ is a constant, but with a little bit more work, we can get much more accuracy. If $\log \Psi_{\rm T}$ is quadratic, we can expand out $v(r)$ without approximation

$$\begin{aligned} v(r) &= v(r_0) + (r - r_0)\cdot\nabla v(r_0) + \cdots \\ &= \tfrac{\hbar^2}{m}\nabla \log \Psi_{\rm T} + \tfrac{\hbar^2}{m}(r - r_0)\cdot\nabla\nabla \log \Psi_{\rm T}. \end{aligned}$$

This second derivative term changes both the diffusion and the drift term, but it is a position operator that does not commute with the momentum operator, so we need a different approach to derive the exact propagator function (exact for quadratic wavefunctions and an improvement for others). We do this by following Feynman's method of deriving the density matrix for a harmonic oscillator (a subtly different problem!).
First write down the propagation equations.

$$\begin{aligned} \tfrac{\partial}{\partial\tau} f_{\rm DMC} &= \tfrac{\hbar^2}{2m}\nabla^2 f_{\rm DMC} - \tfrac{\hbar^2}{m}\nabla\cdot\left(\nabla \log \Psi f_{\rm DMC}\right), \\ f_{\rm DMC}(r,\tau) &= \int_r P(r_0 \to r;\tau) f(r_0, 0), \\ \tfrac{\partial}{\partial\tau} P &= \tfrac{\hbar^2}{2m}\nabla^2 P - \tfrac{\hbar^2}{m}\nabla\cdot\left(\nabla \log \Psi P\right). \end{aligned}$$

Then, following Feynmann, we transform to dimensionless variables:
$y = \sqrt{\frac{m\omega}{\hbar}}x$, $\nabla_y = \sqrt{\frac{\hbar}{m\omega}}\nabla_x$ and $f = \tau\hbar\omega$.

$$\begin{aligned} \tfrac{\partial P}{\partial f} &= \tfrac{1}{2}\nabla^2 P - \nabla\cdot\left(\nabla \log \Psi P\right) \\ &= \tfrac{1}{2}\nabla^2 P - \nabla\cdot\left(v(y)P(y)\right), \qquad (5.1) \\ &\quad \text{where} \qquad v(y) \equiv \nabla_y \log \Psi. \end{aligned}$$

We expect P to have a gaussian form:

$$\begin{aligned} P &= \exp\left\{-a(f)y^2 - b(f)y - c(f)\right\}, \\ \nabla P &= \left[-2ay - b\right]\exp\left\{\quad\right\}, \\ \nabla^2 P &= \left[4a^2y^2 + 4aby + \left(b^2 - 2a\right)\right]\exp\left\{\quad\right\}. \end{aligned}$$

5.3.1 Local Quadratic form

Then can write down an approximation to $v(y)$ that is locally correct to quadratic order in $\log \Psi$.

$$\begin{aligned} v(y) &= -yM(y_0) + v_\Delta(y_0), \\ &\text{where} \quad M(y_0) = \nabla_y \nabla_y \log \Psi(y_0), \\ &\qquad\quad v_\Delta(y_0) = \nabla \log \Psi(y_0) - y_0 \nabla^2 \log \Psi(y_0). \end{aligned}$$

For quadratic wavefunctions, this form is exact, and for the simple harmonic oscillator, $v_\Delta = 0$. Again P is gaussian

$$\begin{aligned} P &= \exp\left\{-a(f)y^2 - b(f)y - c(f)\right\}, \\ \dot{P} &= \left[-\dot{a}y^2 - \dot{b}y - \dot{c}\right] P, \\ -v(y)\nabla P &= \left[-2May^2 - Mby + 2v_\Delta ay + v_\Delta b\right] P, \\ \tfrac{1}{2}\nabla^2 P &= \left[4a^2y^2 + 4aby + \left(b^2 - 2a\right)\right] P. \end{aligned}$$

Using eqn. 5.1, and collecting terms in y^2, y and 1 respectively,

$$\begin{aligned} -\dot{a} &= 2a^2 - 2Ma, \\ -\dot{b} &= 2ab - Mb + 2Mav_\Delta, \\ -\dot{c} &= \tfrac{1}{2}b^2 + 2Mbv_\Delta - a + M. \end{aligned}$$

We can solve these using the boundary conditions $a(f) \to -1/2f$ and $b(f) \to -2y_0 a(f)$ as $f \to 0$:

$$\begin{aligned} a(f) &= \frac{M}{1 - e^{-2Mf}}, \\ b(f) &= \frac{-2My_0 e^{-Mf} - 2v_\Delta\left(1 - e^{-Mf}\right)}{1 - e^{-2Mf}}. \end{aligned}$$

When reinserted into P, this gives,

$$\begin{aligned}
P &\propto \sqrt{\tfrac{1}{2\pi\sigma^2}}\exp\left\{-\tfrac{1}{2\sigma^2}(y-\mu)^2\right\},\\
\sigma(f) &= \sqrt{\frac{1-e^{-2Mf}}{2M}} \qquad \text{(as before)},\\
\mu(f) &= y_0\, e^{-Mf} + v_\Delta M^{-1}\left(1-e^{-Mf}\right),\\
\Delta y &= \mu - y_0 = y_0\left(e^{-Mf}-1\right) + v_\Delta M^{-1}\left(1-e^{-Mf}\right)\\
&= v(y_0)\left[\frac{1-e^{-Mf}}{M}\right]\\
&\approx f v(y_0) - \tfrac{1}{2}f^2 M(y_0)v(y_0).
\end{aligned}$$

Equivalently, we can write these as

$$\begin{aligned}
\sigma(f) &= \sqrt{f}\times\sqrt{\tfrac{1-e^{-2Mf}}{2Mf}} &&= \sqrt{f}\times\sqrt{e^{-Mf}\tfrac{\sinh(Mf)}{Mf}},\\
\Delta y &= v(y_0)\times f\times\left[\tfrac{1-e^{-Mf}}{Mf}\right] &&= v(y_0)\times f\times\left[2e^{-Mf/2}\tfrac{\sinh(Mf)}{Mf}\right].
\end{aligned}$$

Finally we transform coordinates back ($f\to\tau$, $y\to x$).

$$\begin{aligned}
M &= \tfrac{\hbar}{m\omega}\nabla_x^2\log\Psi,\\
f &= \tau\hbar\omega,\\
\epsilon &\equiv -Mf = -\tau\tfrac{\hbar^2}{m}\nabla_x^2\log\Psi,
\end{aligned}$$

$$\begin{aligned}
\sigma_x &= \sqrt{\tfrac{S(\epsilon)\tau\hbar^2}{m}},\\
\tau v_x &= S\left(\tfrac{\epsilon}{2}\right)\tau\tfrac{\hbar^2}{m}\nabla_x\log\Psi,\\
&\qquad \text{where}\quad S(\chi)\equiv\exp(\chi)\frac{\sinh(\chi)}{\chi}.
\end{aligned}$$

We published these expressions in Appendix 1 of [JMC+09]. See Umrigar [UF00] for a discussion of special cases where there are boundary-condition violations (for example when using DMC methods to study fermions, where nodal surfaces are used to divide the space, producing a discontinuities in the first derivative). The formula is valid wherever the locally quadratic approximation to the wavefunction is valid. This can include periodic systems provided that $\tau\hbar^2/mL^2 \ll 1$. For periodic systems, Barker generates a more accurate propagator using Poisson summation[Bar79]. Another solvable case is for a hard-sphere (infinite) potential, where there are discontinuities in the derivative of the wavefunction. This case has been treated by Cao and Berne[CB92b].

5.3.2 Testing via VMC and Metropolis Sampling

For convenience, we can refer to a DMC simulation minus branching, as 'VMC'. It is useful to note that performing VMC is equivalent to sampling ψ_T^2, provided that τ is taken sufficiently small. Although the drift-diffusion $\exp[-\tau(\hat{T}+\hat{D})]$ preserves ψ_T^2, approximations are made in its application, so that the energy estimate will have some timestep dependent error. Since it is possible to sample ψ_T^2 using standard Metropolis Monte Carlo procedures, the efficacy of the approximations made in applying the drift-diffusion evolution operator can be investigated by performing an imaginary time step convergence study of VMC/DMC-sans-branching (The walker number dependence of the DMC results is induced by the branch step only). In this way, measures can be taken to ensure that the imaginary time step of the full DMC computations are limited only by the commutator between the branch and diffusion-drift terms. Additionally as described in Ref. [UNR93], the errors in the drift-diffusion propagator can be reduced using a Metroplis Monte Carlo rejection sampling algorithm that imposes detailed balance by employing the approximate propagator as the *a priori* transition probability and ψ_T^2 as the limiting distribution. This approach has not been implemented herein but it can be employed with the improved drift-diffusion operator above without loss of generality. A minor drawback of Ref. [UNR93] is that persistent or trapped configurations can occur (e.g. diffusion-drift moves of trapped configurations are accepted with very low weight) which requires careful attention to details in the method's implementation.

In Fig. 5.1(left), the results of an imaginary time step study for a 1D oscillator model under DMC-with-the-branching-step-turned-ff are given. When $\epsilon(x)$ is taken to be non-zero, the computations yield correct results; $\langle\psi_T \mid (H-E_0) \mid \psi_T\rangle = \hbar\omega(1-\lambda)^2/(4\lambda) = E_T$; independent of imaginary time step as is to be expected for a Gaussian model. With $\epsilon(x) \equiv 0$, the computations converge at a finite imaginary time step to the correct result.

In Fig. 5.1(right), a similar imaginary time step study is presented for the QDO xenon dimer with ϵ taken to be zero and taken equal to the on-site interaction value. Again, taking $\epsilon \neq 0$ increases significantly the imaginary time step at which convergence is obtained. However, both choices of ϵ converge to the result obtained from a standard Metropolis sampling study.

In order to demonstrate the utility of the local harmonic drift-diffusion propagator for systems in which the trial function is not predominately Gaussian in character, we show the results for the hydrogen atom ($\psi_T = \exp(-r/a_0)$) within the isotropic diagonal 2nd derivative approximation ($\epsilon = -2\tau\hbar^2/(3mra_0)$) simulated in Cartesian coordinates in Fig. 5.2.

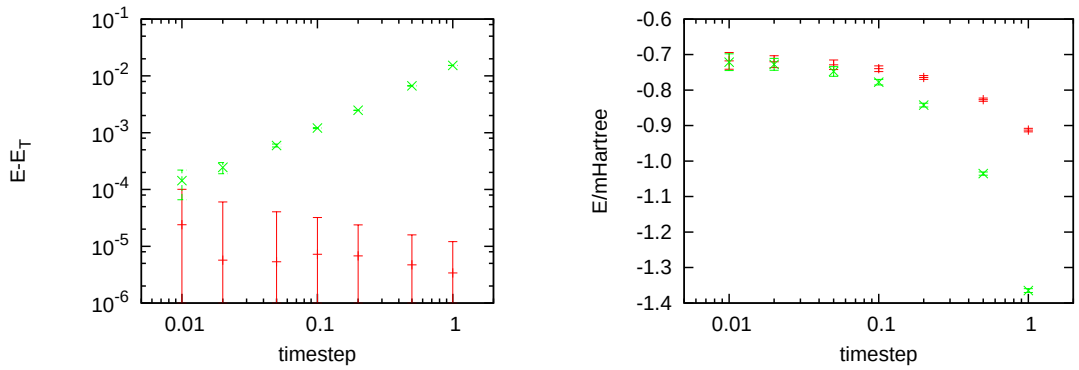


Figure 5.1: (left)The convergence of the variational energy with imaginary time step both with and without the exact drift ($\epsilon \neq 0$, dash, and $\epsilon \equiv 0$, cross), for the one-dimensional harmonic oscillator computing using the DMC method without branching and $\psi_T = \psi_0(\lambda\omega); \lambda = 0.9$. (right) The convergence of the variational energy with imaginary time step both with and without the exact drift ($\epsilon \neq 0$, dash, and $\epsilon \equiv 0$, cross), for the full QDO xenon dimer computing using the DMC method without branching and the on-site-plus-all-pair-multipole trial function.

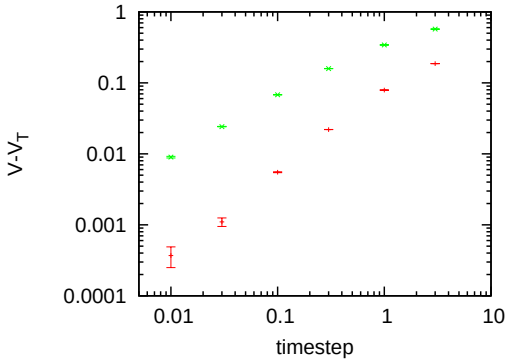


Figure 5.2: (left)The convergence of a non-harmonic system, $V_{\mathrm{T}} = -\langle e^2/r\rangle_{\Psi_{\mathrm{T}}{}^2} = (-e^2/a_0)$ for the hydrogen atom, $\Psi_{\mathrm{T}} = \exp(-r/a_0)$, with imaginary time step both with and without the isotropic diagonal approximation to the trial function 2nd derivative matrix ($\epsilon = -2\hbar^2/(3mra_0)$, dash, and $\epsilon \equiv 0$, cross) computed using the DMC method without branching in Cartesian coordinates.

5.4 Norm conserving DMC method that upholds detailed balance

We have recently proposed a method [JMC+09] to conserve the sampling population with branching, by ensuring that the number of new walkers created by copying is always exactly balanced by the number of walkers that are deleted. The method satisfies detailed balance and therefore can be corrected, and it is stable against the population blowing up or vanishing.

The other operator, the drift-diffusion evolution operator $\exp(-\tau\hat{O})$ preserves the norm

$$\exp(-\tau\hat{O})\psi_T^2 = \psi_T^2 \, .$$

It also preserves the norm of an arbitrary f given that f and ψ_T are well behaved at the boundaries. The given drift-diffusion evolution preserves the norm of the steady state solution; if the overall approximate evolution is to preserve the norm of the steady state solution, then the branch evolution should also preserve the norm.

That is, for the ground state, the expected number of "births" per walker, $\langle n^{(+)}\rangle$, should match the expected number of "deaths" per walker, $\langle n^{(-)}\rangle$, as is defined in the following:

$$\begin{aligned}
\int d\mathbf{r}\, e^{-\tau\hat{V}} f(\mathbf{r}) &= \int d\mathbf{r}\, f(\mathbf{r}), \\
\int d\mathbf{r} \left(e^{-\tau\hat{V}} - 1\right) f(\mathbf{r}) &= 0, \\
\int d\mathbf{r} \left(1 - e^{-\tau\hat{V}}\right) \theta^{(+)} f(\mathbf{r}), &= \int d\mathbf{r} \left(e^{-\tau\hat{V}} - 1\right) \theta^{(-)} f(\mathbf{r}) \, , \\
\langle n^{(-)}\rangle_f &= \langle n^{(+)}\rangle_f \, .
\end{aligned} \tag{5.2}$$

Here, $\hat{V}(r) = \frac{\hat{H}\Psi(r)}{\Psi(r)} - \bar{E}$, $\theta^{(+)} = \theta(V)$, and $\theta^{(-)} = \theta(-V)$, where θ is the switch-on function [$\theta(x) = 1$ for $x \geqslant 0$, and $\theta(x) = 0$ otherwise]. The desired condition simply states that the flux into $\exp(-\tau\hat{V})f$, $\langle n^{(+)}\rangle_f$, must be balanced by the flux out of $\exp(-\tau\hat{V})f$, $\langle n^{(-)}\rangle_f$. Given that the flux matching condition need only hold for the steady state solution, it is clear that a unique $\bar{E}$ that satisfies Eq. (5.2) can be found which supports the existence of the corresponding stationary solution $f(\mathbf{r})$.

Finally, we note that the flux in is equal to the flux out only on average. Therefore, for a finite walker population, the number of walkers in a DMC simulation will fluctuate in time (e.g. unless $\psi_T \equiv \psi_0$). Thus the branching operator does not preserve the norm

for arbitraty f.

5.4.1 The Idea (Flux-balance)

The foregoing analysis of the stationary solution of the split operator DMC method gives insight into construction of a norm-conserving or constant-N, DMC or NC-DMC method. Consider an ensemble of N walkers ($\{r_1 \dots r_N\} = \mathbf{r}$) (each one for the sake of clarity consisting of a single particle moving in three spatial dimensions) selected with probability $\prod_i f(r_i; 0)\, dr_i$. Application of the norm-conserving drift-diffusion term to each ensemble member is straightforward and standard above, as the N-walker system is separable. In light of the discussion in the previous paragraph, it is natural to consider the entire ensemble in developing a new method to apply the branch evolution, $\exp(-\tau\hat{V})$. As described above, flux matching $\langle n^{(-)}\rangle_{\bar{f}} = \langle n^{(+)}\rangle_{\bar{f}}$, is true on average but not instantaneously $n^{(-)}(\mathbf{r}) \neq n^{(+)}(\mathbf{r})$. It is natural to replace the average by a sum over walkers, however, $(1/N)\sum_i n^{(+)}(r_i) \neq (1/N)\sum_i n^{(-)}(r_i)$ unless $N \to \infty$.

It is therefore proposed to enforce norm conservation at each branching step, by enforcing flux balance for the instantaneous N-walker configuration, through a modification to the acceptance rule following earlier work by others[AMK00, BS98, Het84]. This is accomplished by defining

$$\begin{aligned} \bar{n}^{(\pm)}(\mathbf{r}) &= \frac{1}{N}\sum_i n^{(\pm)}(r_i), \\ w^{(\pm)}(\mathbf{r}) &= \frac{\bar{n}^{(+)}(\mathbf{r}) + \bar{n}^{(-)}(\mathbf{r})}{2\bar{n}^{(\pm)}(\mathbf{r})}, \end{aligned} \tag{5.3}$$

and taking

$$P^{(\pm)}(i;\mathbf{r}) = w^{(\pm)}(\mathbf{r}) n^{(\pm)}(r_i), \tag{5.4}$$

where $P^{(\pm)}(i;\mathbf{r})$ is probability walker i contributes to the flux into/out of the new ensemble (e.g. the action of the branch operator). The modification creates the "flux-matching branch operator" for each walker

$$\langle \mathbf{r} \mid \exp[-\tau\hat{o}_2] \mid \mathbf{r}\rangle \equiv \left[w^{(+)}(\mathbf{r}) n^{(+)}(r_i) - \mathrm{Min}\left(w^{(-)}(\mathbf{r}) n^{(-)}(r_i), 1\right) + 1\right],$$

defined for $N > 1$ which when summed over $i = 1, N$ yields unity by definition. In

more detail, a self-consistent computation is required

$$w^{(-)}(\mathbf{r})\bar{n}^{(-)}(\mathbf{r}) = \frac{1}{N}\sum_i \mathrm{Min}\left(w^{(-)}(\mathbf{r})n^{(-)}(r_i), 1\right), \tag{5.5}$$

$$w^{(-)}(\mathbf{r}) = \frac{w^{(-)}(\mathbf{r})\bar{n}^{(-)}(\mathbf{r})}{4w^{(-)}(\mathbf{r})\bar{n}^{(-)}(\mathbf{r}) - 2\bar{n}^{(+)}(\mathbf{r})}, > 0$$

to ensure the ensemble branch operator is positive definite $w^{(-)}(\mathbf{r})n^{(-)}(r_i) \leq 1$; in practice, self-consistent cycles can be avoided as will be discussed later. It is assumed that all $n^{(-)}(r_i)$ are strictly less than unity or equivalently there are no boundary condition violations. The parameter $\bar{E}$ is still selected such that $\langle n^{(-)}\rangle_{\bar{f}} = \langle n^{(+)}\rangle_{\bar{f}}$. Note, $w^{(\pm)}(\mathbf{r})$ approaches unity as $N \to \infty$ and the branch operator properly reduces to the original form. This approach is new, and is not equivalent to that of the earlier work we cited [BS98, Het84, AMK00].

The new NC-DMC method strictly conserves the number of walkers in the ensemble and hence ensures the stability of the simulation for any reasonable choice of $\bar{E}$. Thus, neither a rare fluctuation far away from flux matching nor a deviation, $\bar{E} = \tilde{E}_0 + \delta$, can cause the walker population to grow or shrink by an unacceptably large amount as in the original method. Note, $\bar{E}$ is only known within some error and first order perturbation theory in δ is invoked to analyze the simulation results here (as in any DMC simulation). The lack of weights ensures that statistical averages can be taken safely to the long time limit but with the $\mathcal{O}(1/N)$ bias as in standard DMC. Unlike standard DMC, the acceptance rule does not modulate in time and the method strictly satisfies detailed balance allowing for corrections given below. Below, in (5.4.3), details of the new branching process are given along with prescriptions (*i*) to choose $\bar{E}$, to estimate E_0, (*ii*) to check for the correctness of the simulation results, (*iii*) to treat configurations of measure zero wherein the ensemble branch operator becomes undefined or nearly so (e.g. requiring self-consistent cycles), and (*iv*) to treat systems where boundary condition violations ($n^{(-)}(r_i) \equiv 1$ for any i) cannot be avoided.

5.4.2 Reweighting for the elimination of $\frac{1}{N}$ bias

The NC-DMC method has a bias that leads to $\mathcal{O}(1/N)$ error as in standard DMC with its time varying acceptance rule. However, under NC-DMC it is possible to correct for the bias by defining appropriate weights in the usual way. We begin by defining the

exact weight by rewriting the ensemble branch operator

$$\begin{aligned}\langle \mathbf{r} \mid \exp[-\tau \hat{O}] \mid \mathbf{r} \rangle &= \exp[-\tau V_{\rm eff}(r_i)] \\ &\times \exp\Big\{\log\Big[\,1+(w^{(+)}(\mathbf{r})-1)\,n^{(+,\dagger)}(r_i) \\ &\qquad -(w^{(-)}(\mathbf{r})-1)\,n^{(-,\dagger)}(r_i)\,\Big]\Big\}\,,\end{aligned}$$

where the minimum condition was removed for simplicity and

$$n^{(\pm,\dagger)}(r_i) = e^{\tau V_{\rm eff}(r_i)} n^{(\pm)}(r_i)\,.$$

The weighting factor for each walker is then

$$\begin{aligned}W^{(\rm corr)}(r_i,\mathbf{r};\tau) &= \exp\Big\{-\log\Big[\,1+(w^{(+)}(\mathbf{r})-1)\,n^{(+,\dagger)}(r_i) \\ &\qquad -(w^{(-)}(\mathbf{r})-1)\,n^{(-,\dagger)}(r_i)\,\Big]\Big\}\,.\end{aligned}$$

Accumulating the weights for projection time, $P\tau$, requires NP extra storage and communication (in parallel computations) and removes all population bias. It is possible to define a mean field correction that provides a weight for the entire ensemble and introduces extra storage of size P and no additional communication (1 double),

$$\begin{aligned}W^{(\rm corr,mf)}(\mathbf{r};\tau) &= \exp\left\{-\frac{1}{N}\sum_i \log\Big[\,1+(w^{(+)}(\mathbf{r})-1)\,n^{(+,\dagger)}(r_i) \right. \\ &\qquad \left. -(w^{(-)}(\mathbf{r})-1)\,n^{(-,\dagger)}(r_i)\,\Big]\right\}\end{aligned}$$

The mean-field weight distinguishes 'good collections' of size N from 'better collections' of size N thereby correcting approximately for the effect of the strict norm constraint.

Under NC-DMC, the parameter, $\bar{E}$, can naturally be determined so the average of the correction in mean field is zero. Assuming the simulation is performed using a good estimate $\bar{E} = \bar{E}_{\rm true} + \delta$.

$$\bar{E}_{\rm true} = \bar{E} - \frac{\left\langle \log\left[W^{(\rm corr,mf)}(\mathbf{r};\tau)\right]\right\rangle}{\left\langle \frac{\partial \log\left[W^{(\rm corr,mf)}(\mathbf{r};\tau)\right]}{\partial \bar{E}}\right\rangle}\,,$$

where the average is not corrected. This result reduces to the standard condition

$$\begin{aligned} \bar{E}_{\text{true}} &= \bar{E} + \frac{\left\langle \bar{n}^{(-)} - \bar{n}^{(+)} \right\rangle}{\tau \left\langle \bar{n}^{(-)} - \bar{n}^{(+)} + 1 \right\rangle} + \mathcal{O}\left(\delta^2\right) \\ &= \bar{E} + \frac{\left\langle \bar{n}^{(-)} - \bar{n}^{(+)} \right\rangle}{\tau} + \mathcal{O}\left(\delta^2\right), \qquad \text{as} \quad N \to \infty. \end{aligned}$$

In the following, we shall not apply the weights but the improvement wrought by the mean field weights is given in Fig. 5.6(b) below (the number of walkers required to obtain a given accuracy is reduced by a factor of ≈ 4).

It should be noted in passing that even re-weighting does not lead to exact converged results in practice. This is because of the way the sampling distibution actually overlaps with the correct distribution. This is a common problem where importance sampling is involved, not just our particular method. The problem arises whereever the sampled distribution is small, but the correct distribution is large. This requires a large weighting to be put on any of those samples. But as the sampling is stochastic, this produces the occasional sample with very large weight, which dominates over most of the other samples, and prevents good averaging.

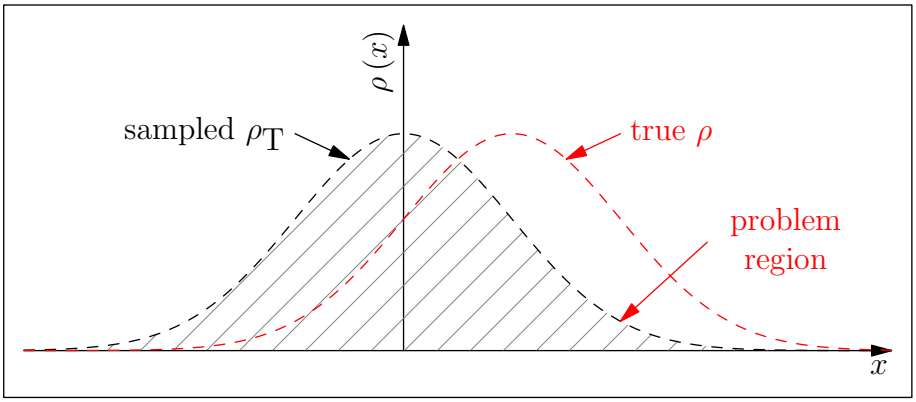


Figure 5.3: The true distribution (red) vs the sampling distribution (black): the left hand side is sampled very well, but the right hand side is sampled very poorly. To make up for this, the right hand side receives bigger weights (ρ/ρ_{T}) to pump it up to the true distribution. The further to the right the sample is, the bigger the weight, so an occasional sample can have a very big weight and make the results very stochastic.

5.4.3 Estimating the energy in NC-DMC

The ground state energy can be estimated in two ways from a DMC simulation[KW86, HJR94, Cep95]. First, using the properties of a Hermitian operator

$$\int d\mathbf{r} f(\mathbf{r}) \frac{\hat{H}\psi_T(\mathbf{r})}{\psi_T(\mathbf{r})} \quad = \quad \int d\mathbf{r} \psi_T(\mathbf{r}) \hat{H}\psi_0(\mathbf{r}) = E_0.$$

Hence, an estimator for E_0 is simply

$$E_0^{(\mathrm{est,T})}(\mathbf{r};\tau) = \frac{1}{N}\sum_i \frac{\hat{H}\psi_T(r_i)}{\psi_T(r_i)},$$

which is then averaged over the stochastic DMC process or DMC "trajectory" to yield a variational estimate for E_0 (assuming the steady state is of the form $f(\tau) = \psi_T\psi_0$).

The second estimator is constructed using the fact that $\bar{E}$ should be taken equal to $\tilde{E}_0(\tau)$. If $\bar{E}$ is not taken exactly equal to E_0, then the average of $\bar{n}^{(+)}(\mathbf{r}) - \bar{n}^{(-)}(\mathbf{r})$ over the DMC trajectory will not be zero. Using 1st order perturbation theory, letting $\bar{E} = \tilde{E}_0(\tau) + \delta$, expanding $n^{(\pm)}(r_i)$ to lowest order in τ yields

$$E_0^{(\mathrm{est},\tau)}(\mathbf{r};\tau) \quad = \quad \bar{E} + \frac{\bar{n}^{(-)}(\mathbf{r}) - \bar{n}^{(+)}(\mathbf{r})}{\tau}, \tag{5.6}$$

which must be averaged over the DMC trajectory to predict $\tilde{E}_0(\tau)$.

Typically, $\langle E_0^{(\mathrm{est,T})}\rangle$ has lower variance but if $\langle E_0^{(\mathrm{est},\tau)}\rangle$ does not closely agree with the latter, the simulation should be rerun with a smaller time step, τ given $\langle(\bar{n}^{(-)} - \bar{n}^{(+)})/\tau\rangle$ is sufficiently small. If the latter average is large, a new run with a refined choice of $\bar{E} = \langle E_0^{(\mathrm{est},\tau)}\rangle$ will be required. Note, the above averages over the NC-DMC trajectory are assumed to be invariant to 1st order variations in $\bar{E}$.

5.4.4 Selecting an $\bar{E}$

It is important to select a good value of $\bar{E}$. A useful procedure is to start $\bar{E}$ equal to the variational energy, $\bar{E} = \int d\mathbf{r}\psi_T(\mathbf{r})\hat{H}\psi_T(\mathbf{r})$, run NC-DMC for a few hundred steps, refine $\bar{E}$ using eqn. 5.6, and so on until reasonable convergence is achieved. A long run may then be spawned with $\bar{E}$ fixed. The simulations are stable independent of $\bar{E}$ and the quality of a NC-DMC simulation with fixed $\bar{E}$ can be assessed from the estimators of the ground state energy as described above. Computing 1 dimensional distribution such as $P(\bar{n}^{(+)})$, $P(\bar{n}^{(-)})$ and $P(\bar{n}^{(-)} - \bar{n}^{(+)})$ can also help

judge the quality of the simulations. From the central limit theorem, as $N \to \infty$, $P(\bar{n}^{(\pm)})$ approaches a Gaussian characterized by mean $\sigma\sqrt{2/\pi}$ and standard deviation $\sigma\sqrt{(\pi-2)N_c/(\pi N)}$ and $P(\bar{n}^{(-)} - \bar{n}^{(+)})$ approaches a Gaussian characterized by zero mean and standard deviation $\sigma\sqrt{N_c/N}$ where N_c is a walker correlation number. If the quantity, $1 - \exp[-\tau V_{\text{eff}}]\exp[-\tau V_{\text{eff}}]$, itself exhibits zero mean Gaussian statistics with standard deviation, σ, then $N_c \equiv 1$. The average of the weighting function $\langle w^{(\pm)}(\mathbf{r})\rangle$ over the trajectory should be close to unity or too few walkers have been employed.

5.4.5 Avoiding 'Sign Collapse'

The NC-DMC method divides walkers into two populations at every step, those below $\bar{E}$, and those above and equalizes the flux out of the ensemble (arising from walkers with energy above $\bar{E}$) to the flux into the ensemble (arising walkers with energy below $\bar{E}$). Thus, the NC-DMC method can be become ill-defined if all walkers have $V_{\text{eff}} \geq 0$ or $V_{\text{eff}} < 0$ for a given $N-$walker configuration. Making the reasonable assumption that the probability of any walker having a V_{eff} of sign positive/negative is 0.5 for the correct choice of $\bar{E}$, the probability that N walkers all have the V_{eff} with the same sign is $P^{(\text{all }\pm)}(N) = 2^{1-N}$. For N as small as $N = 50$, $P^{(\text{all }\pm)}(50) \approx 2 \times 10^{-15}$ while for $N = 20$, $P^{(\text{all }\pm)}(20) \approx 2 \times 10^{-6}$. Thus, walker "sign collapse" is a rare event given large enough N. Note, due to walker correlations, the effective number of walkers is reduced, $N \to N/N_c$, where N_c is a walker correlation number. Now, for sign collapsed configurations, we can simply choose to take all present walkers into the next ensemble without prejudice (e.g. no flux in or out). This norm conversing choice is microscopically reversible because a sign collapse state is history independent (depends only on configuration). It effectively takes $\prod_i f(r_i) = \prod_i \psi_T^2(r_i)$ for any sign collapsed configuration which is, in fact, exact for the case $\psi_T = \psi_0$ where $V_{\text{eff}} \equiv 0$, all configurations are sign collapsed (e.g the sign of the number 0 is by convention positive) and there is no flux $\bar{n}^{(\pm)} \equiv 0$. The "accept all" choice for the sign collapsed configurations defines the $N = 1$ limit wherein NC-DMC samples $f(\mathbf{r}) = \psi_T^2(\mathbf{r})$. It also preserves the variational character of the NC-DMC technique and leads to the "ensemble flux matching" branch operator definition for each walker

$$\begin{aligned}
\langle \mathbf{r} \mid \exp[-\tau\hat{o}_2] \mid \mathbf{r}\rangle &= \left[\left\{w^{(+)}(\mathbf{r})n^{(+)}(r_i) - \text{Min}\left(w^{(-)}(\mathbf{r})n^{(-)}(r_i), 1\right)\right\}\chi(\mathbf{r}) + 1\right], \\
\chi(\mathbf{r}) &= 1 - \left[\prod_i \theta^{(-)}(V_{\text{eff}}(r_i))\right] - \left[\prod_i \theta^{(+)}(V_{\text{eff}}(r_i))\right],
\end{aligned}$$

where $\chi(\mathbf{r}) = 0$ for sign collapsed configurations and $\chi(\mathbf{r}) = 1$ otherwise. Similarly, configurations that are not sign collapsed but for which the self-consistent equation for

$\bar{n}^{(-)}(\mathbf{r})$, Eq. (5.5) would be required are also accepted without prejudice in the absence of boundary condition violations.

Walker sign collapse occurs for a vanishingly small fraction of configurations given $N \geq 18$ and $\psi_T \neq \psi_0$. That is, sign collapsed configurations become points of measure zero in the N-walker configuration space (e.g. $\mathbf{r}$) for N remarkably small. The NC-DMC method is not recommended for use with $N < 18$ where sign collapsed configurations contribute more than 1 part in 1×10^5 unless ψ_T is a particularly good estimate of ψ_0. The "accept all" choice for sign collapsed configurations ensures the method preserves the $\psi_T \to \psi_0; V_{\text{eff}} \to 0$ limit where $\bar{n}^{(\pm)} \equiv 0$ leads to ambiguities in the more naive NC-DMC acceptance rule and all configurations are, by definition, sign collapsed. For the case $\psi_T \neq \psi_0$, the "accept all" condition makes the variationally consistent choice $\prod_i f(r_i) = \prod_i \psi_T^2(r_i)$ for a set of configurations which rapidly approaches measure zero as $N \to \infty$. It, also, neatly defines the $N = 1$ limit of the NC-DMC method to simply be a variational computation with trial function, ψ_T. If too many sign-collapsed configurations are identified, the number of walkers should be increased. However, reweighting either in mean field or exactly corrects for sign collapse but error to be corrected scales as $\mathcal{O}(1/N)$.

5.4.6 Special Case Boundary Conditions

For completeness, we consider systems in which the diffusion-drift operator cannot be applied such that the boundary conditions are satisfied, for example where $\Psi_{\text{T}} \to 0$ in some region. Note, the order of operations should be reversed for this case. The QDO Hamiltonian studied herein is sufficiently simple that the drift-diffusion operator can be applied "properly". In general, any walker that violates the boundary conditions must be rejected and its rejection probability, $n^{(-)}(r_i)$ cannot be scaled so that it is different from unity. If $\bar{n}^{(-)} < \bar{n}^{(+)}$ and the self-consistency condition given in Eq. (5.5) has a solution, the NC-DMC method needs no modification. When $\bar{n}^{(-)} < \bar{n}^{(+)}$, the new self-consistency condition

$$\begin{aligned} w^{(-)}(\mathbf{r})\bar{n}^{(-)}(\mathbf{r}) &= \sum_i n^{(-)}(r_i)\theta^{(+)}(n^{(-)}(r_i) - 1) \\ &+ w^{(-)}(\mathbf{r}) \sum_i n^{(-)}(r_i)\theta^{(-)}(1 - n^{(-)}(r_i)), \\ w^{(-)}(\mathbf{r}) &= \frac{w^{(-)}(\mathbf{r})\bar{n}^{(-)}(\mathbf{r})}{4w^{(-)}(\mathbf{r})\bar{n}^{(-)}(\mathbf{r}) - 2\bar{n}^{(+)}(\mathbf{r})} > 0, \end{aligned}$$

is imposed. For those rare N-walker configurations for which self-consistent solutions do not exist, all walkers that violate the boundary conditions are removed. The new ensemble of size $N_{new} < N$ is then grown back to size N by randomly copying the

remaining walkers with equal probability as described in the first paragraph of this subsection. Again, these fixes for configurations of measure zero are not history dependent. Note, the method of Ref. [UNR93] can be implemented so as to reject boundary condition violating moves at the drift/diffusion step thereby obviating the above procedure. That is, we can use Diffuse/Flow propagators that respect the boundary conditions.

5.4.7 Testing NC-DMC

In order to test the new NC-DMC technique and the new drift-diffusion evolution operator, the harmonic oscillator is studied

$$\begin{aligned}
\hat{H} &= \hat{T} + \frac{m\omega^2\hat{x}^2}{2}, \\
E_0 &= \frac{\hbar\omega}{2}, \\
\psi_0 &= \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} \exp\left(\frac{m\omega x^2}{2\hbar}\right), \\
\psi_T(\lambda) &= \left(\frac{m\omega\lambda}{\pi\hbar}\right)^{1/4} \exp\left(\frac{m\omega\lambda x^2}{2\hbar}\right), \\
E_T &= \langle\psi_T \mid \hat{H} - E_0 \mid \psi_T\rangle = \frac{\hbar\omega(1-\lambda)^2}{4\lambda}.
\end{aligned}$$

Tests are performed using $\lambda = 0$ and $\lambda = 0.9$.

In Fig. 5.4, a NC-DMC study of the convergence of the ground state energy with imaginary time step, Fig. 5.4(top), and walker number, Fig. 5.4(bot), are presented for the $\lambda = 0$ trial function (no importance sampling). The results converge as τ^2 to the correct value ($E_0(\tau) - E_0$ is presented in the figure). Approximately 200 walkers are required to generate converged results.

In Fig. 5.5, the same study is performed using the $\lambda = 0.9$ trial function both with and without the improved drift-diffusion evolution operator. Again, appropriate convergence behavior is observed with the improved drift-diffusion evolution operator exhibiting faster convergence with τ. Approximately 200 walkers seems to be required to generate accurate results.

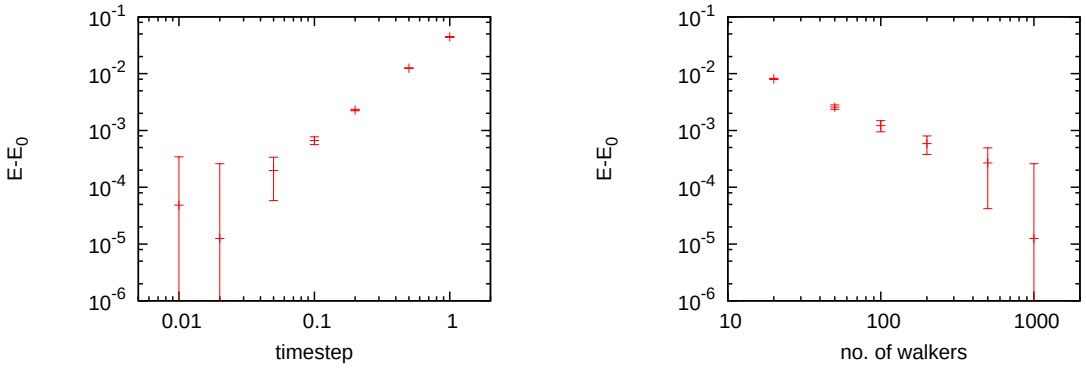


Figure 5.4: (left) The convergence of the ground state energy with imaginary time step for the one-dimension harmonic oscillator computed using the NC-DMC method with $\psi_T = 1$ and $N = 1000$ walkers. (right) The corresponding convergence of the ground state energy with walker number computed using an imaginary time step of $\tau\hbar\omega = 0.02$.

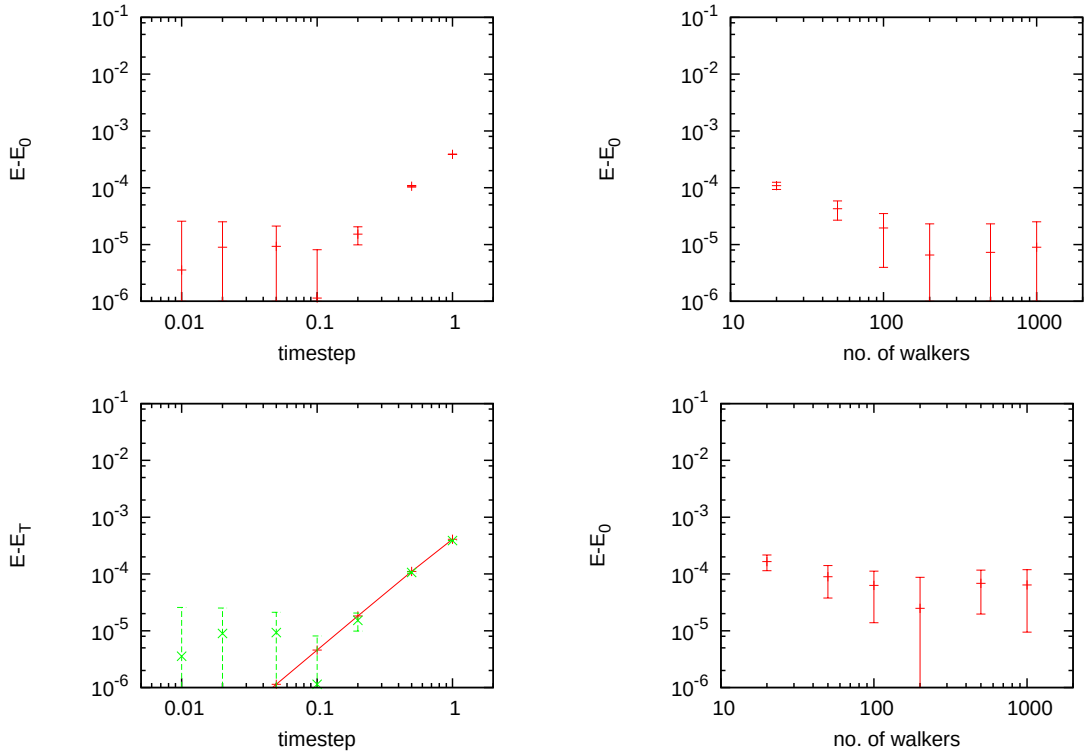


Figure 5.5: (top-left)The convergence of the ground state energy with imaginary time step for the one-dimensional harmonic oscillator computed using the NC-DMC method with $\psi_T = \psi_0(\lambda\omega); \lambda = 0.9; \epsilon \neq 0$ and $N = 1000$ walkers. (top-right)The corresponding convergence of the ground state energy with walker number computed using an imaginary time step of $\tau\hbar\omega = 0.1$. (bot-left)The convergence of the ground state energy with imaginary time step for the one-dimensional harmonic oscillator computed using the NC-DMC method with $\psi_T = \psi_0(\lambda\omega); \lambda = 0.9; \epsilon = 0$ and $N = 1000$ walkers. (bot-right)The corresponding convergence of the ground state energy with walker number computed using an imaginary time step of $\tau\hbar\omega = 0.002$.

5.4.8 Analytical studies of Diffuse-Drift Error

Exact results for harmonic systems can be obtained and used to aid in accessing the accuracy of the new approach. The general expression

$$\langle x \mid \exp(-\tau\hat{O}) \mid x' \rangle \equiv \langle x \mid \Psi_{\rm T} \exp(-\tau\hat{H}) {\Psi_{\rm T}}^{-1} \mid x' \rangle,$$

can be evaluated analytically for the special case $\hat{H} = \hat{T} + m\omega^2 x^2/2$ and $\hat{H}_{\rm T} = \hat{T} + m\omega^2\lambda^2 x^2/2$ as described in Ref. [Fey72],

$$\begin{aligned}\langle x \mid \exp(-\tau\hat{O}) \mid x' \rangle = \\ \left[\frac{m\omega}{2\hbar\sinh(\tau\hbar\omega)}\right]^{1/2} \exp\left\{-\frac{m\omega}{2\hbar}\left[\frac{1}{\sinh(\tau\hbar\omega)}(x-x')^2\right.\right. \\ \left.\left. + \tanh\left(\frac{\tau\hbar\omega}{2}\right)(x^2+x'^2) + \lambda(x^2-x'^2)\right]\right\}.\end{aligned}$$

The approximate propapagor of the text can also be generated analytically for harmonic systems,

$$\begin{aligned}\langle x \mid \exp\left(-\tau\hat{\tilde{O}}(\tau)\right) \mid x' \rangle &= \langle x \mid e^{-\tau\hat{V}/2} e^{-\tau\hat{O}} e^{-\tau\hat{V}/2} \mid x' \rangle \\ &= \left[\frac{m\omega a(\tau,\omega,\lambda)}{2}\right]^{1/2} \exp\left\{-\frac{m\omega}{2\hbar}\left[a(\tau,\omega,\lambda)(x-x')^2\right.\right. \\ &\quad \left.\left. + b(\tau,\omega,\lambda)(x^2+x'^2) + \lambda(x^2-x'^2)\right]\right\},\end{aligned}$$

where

$$\begin{aligned} a(\tau,\omega,\lambda) &= \left[\frac{\lambda}{\sinh(\tau\hbar\omega\lambda)}\right]\left[1+\left(\frac{\tau\hbar\omega(1-\lambda^2)}{2\lambda}\right)\tanh\left(\frac{\tau\hbar\omega\lambda}{2}\right)\right]^{-1}, \\ b(\tau,\omega,\lambda) &= \lambda\tanh\left(\frac{\tau\hbar\omega\lambda}{2}\right) - 2a(\tau,\omega,\lambda) + \frac{2\lambda}{\sinh(\tau\hbar\omega\lambda)}. \end{aligned}$$

Since the approximate expression has same the form as the exact result, the finite τ propagator can be mapped on to its exact solution at arbitrary $\beta = P\tau$ through the definition of three effective parameters

$$\begin{aligned} \tilde{\omega}(\tau) &= \left(\frac{2}{\tau\hbar}\right)\sinh^{-1}\left(\left[\frac{a(\tau,\omega,\lambda)}{2b(\tau,\omega,\lambda)}\right]^{1/2}\right), \\ \tilde{m}(\tau) &= \left(\frac{m\omega a(\tau,\omega,\lambda,m)}{\tilde{\omega}(\tau)}\right)\sinh(\tau\hbar\tilde{\omega}(\tau)), \\ \tilde{\lambda}(\tau) &= \left(\frac{m\omega\lambda}{\tilde{m}(\tau)\tilde{\omega}(\tau)}\right). \end{aligned}$$

Here, only the τ dependence of the effective parameters is referenced explicitly. Finite imaginary time step dependent expressions for the DMC energy estimators follow

$$\begin{aligned}
\tilde{E}_0(\tau) &= \frac{\hbar\tilde{\omega}(\tau)}{2}, \\
\tilde{E}_H(\tau) &= \left\langle \frac{\hat{H}\psi_T(\lambda)}{\psi_T(\lambda)} \right\rangle_{\tilde{f}(\tau)} = \frac{\hbar\omega\lambda}{2} + \left\langle \frac{m\omega^2(1-\lambda^2)}{2}x^2 \right\rangle_{\tilde{f}(\tau)} \\
&= \frac{\hbar\omega}{2}\left[\lambda + \left(\frac{[1-\lambda^2]\tilde{\lambda}(\tau)}{[1+\tilde{\lambda}(\tau)]\lambda}\right)\right].
\end{aligned}$$

where

$$\tilde{f}(x;\tau) = \left(\frac{\tilde{m}(\tau)\tilde{\omega}(\tau)[1+\tilde{\lambda}(\tau)]}{2\pi}\right)^{1/2} \exp\left(-\frac{\tilde{m}(\tau)\tilde{\omega}(\tau)[1+\tilde{\lambda}(\tau)]}{2}x^2\right).$$

The $\tilde{E}_0(\tau)$ estimator exists when $\lambda \to 0$ because the limits of $a(\tau,\omega,0) = 1/\tau\hbar\omega$ and $b(\tau,\omega,0) = \tau\hbar\omega/2$ are well defined. In the limit $\lambda \to 1$ both estimators properly yield the ground state energy $E_0 = \hbar\omega/2$. In Fig. 5.6(l), the convergence of the NC-DMC method to the exact finite imaginary time step results is shown for both a small and a large imaginary time step. The faster convergence of the mean field correction to NC-DMC is demonstrated at finite τ in Fig. 5.6(r).

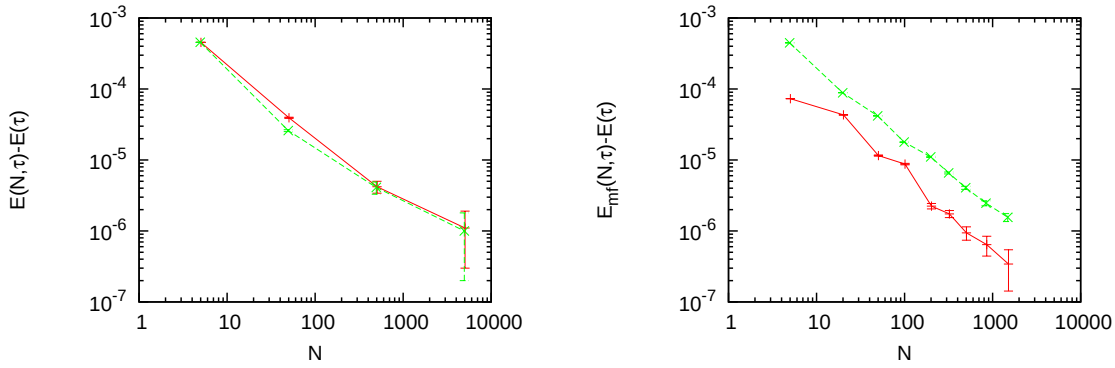


Figure 5.6: (a) The convergence of the ground-state energy as a function of walker number for the one-dimensional harmonic oscillator computed using the NC-DMC method with $\Psi_T = \Psi_0(\lambda\omega)$; $\lambda = 0.9$; $\epsilon \neq 0$ at two imaginary-time steps: $\tau\hbar\omega = 0.6$(dash) and $\tau\hbar\omega = 0.06$(cross). Convergence is presented relative the exact finite imaginary-time step results $\tilde{E}_H(0.6/\hbar\omega) - \hbar\omega/2 = 1.576 \times 10^{-4}$ and $\tilde{E}_H(0.06/\hbar\omega) - \hbar\omega/2 = 1.642 \times 10^{-6}$, respectively. (b) The improvement wrought by mean-field trajectory weighting procedure at imaginary-time step $\tau\hbar\omega = 0.06$. The crosses are generated using the mean-field trajectory weighting correction while the xs are generated using NC-DMC.

Chapter 6

New Trial Wavefunctions for the QDO

We explore various new trial wavefunctions for the QDO, suitable for use in DMC. We start with quadratic dipole-dipole wavefunctions inspired by chapter 3, then incorporate higher-multipole interaction terms, and then describe an iterative method for generating many-body multipole interaction terms, the method that inspired the diagrammatic expansion in chapter 4. We finally settle on a trial wavefunction useful for Quantum Drudes which very accurately includes the effect of the off-site Coulomb interaction (Drude - charge) with the on-site harmonic interaction (Drude - nucleus), except there is damping. Thus the trial wavefunction has some 3-body character (pseudo-electron/pseudo-nucleus/external-charge) which is specific to the Quantum Drude (it is similar to but not the same as a Hydrogen atom interacting with a charge). This trial wavefunction allows us to calculate energies for large systems to high accuracy, as will be demonstrated in the following chapter. We also sketch current work towards a version of this wavefunction that has improved close-range damping.

6.1 Trial wavefunctions for the Dipole limit

6.1.1 General Approach for a Quadratic Perturbation

We can extend the harmonic oscillator trial wavefunction by adding a quadratic matrix term $x\mathbf{M}x$ that couples the oscillation coordinates x:

$$\Psi_{\mathrm{T}} \;=\; e^{-\frac{m\omega}{2\hbar}\left(x^2+\alpha x\mathbf{M}x\right)} \,.$$

Then we plug it into the formulae for DMC. For QDO's only the coupling is interesting, so we drop the constant zero energy of the oscillation of the dipoles ($\frac{3N}{2}\hbar\omega$). Recall $\alpha = q^2(m\omega^2)^{-1}$. This gives us:

$$\begin{aligned} \hat{H} &= -\tfrac{\hbar^2}{2m}\nabla^2 + \tfrac{1}{2}m\omega^2 x^2 - \tfrac{3N}{2}\hbar\omega + \phi \;, \\ v_{\text{drift}} &\equiv \tfrac{\hbar}{m}\nabla \log \Psi_{\text{T}} = -\omega x - \omega\alpha\mathbf{M}x \;. \end{aligned}$$

Now we introduce our energy estimator and multiply/contract energy $\texttt{hpsi}(x) = \frac{\hat{H}\Psi_{\text{T}}}{\Psi_{\text{T}}}$, as used in our DMC code.

$$\begin{aligned} \texttt{hpsi}(x) &= -\tfrac{\hbar^2}{2m}\left[\nabla \log(\Psi_{\text{T}})\right]^2 - \tfrac{\hbar^2}{2m}\left[\nabla^2 \log(\Psi_{\text{T}})\right] + \tfrac{1}{2}m\omega^2 x^2 - \tfrac{3N}{2}\hbar\omega + \phi \\ -\tfrac{\hbar^2}{2m}\left[\nabla \log(\Psi_{\text{T}})\right]^2 &= -\tfrac{\hbar^2}{2m}\left[-\tfrac{m\omega}{\hbar}\left(x + \alpha\mathbf{M}x\right)\right]^2 \\ &= -\tfrac{1}{2}m\omega^2 x^2 - m\omega^2\alpha x\mathbf{M}x - m\omega^2\tfrac{\alpha^2}{2}x\mathbf{M}\mathbf{M}x \\ -\tfrac{\hbar^2}{2m}\left[\nabla^2 \log(\Psi_{\text{T}})\right] &= -\tfrac{\hbar^2}{2m}\left[-\tfrac{m\omega}{\hbar}\left(3N + \alpha\mathbf{Tr}(\mathbf{M})\right)\right] \\ &= +\tfrac{3N}{2}\hbar\omega + \tfrac{1}{2}\hbar\omega\alpha\mathbf{Tr}(\mathbf{M}) \end{aligned} \;, \\ \begin{aligned} \texttt{hpsi}(x) &= \phi(x) - m\omega^2\alpha x\mathbf{M}x - m\omega^2\tfrac{\alpha^2}{2}x\mathbf{M}\mathbf{M}x + \tfrac{1}{2}\hbar\omega\alpha\mathbf{Tr}(\mathbf{M}) \\ &= \phi(x) - q^2 x\mathbf{M}x - q^2\tfrac{\alpha}{2}x\mathbf{M}\mathbf{M}x + \tfrac{1}{2}\hbar\omega\alpha\mathbf{Tr}(\mathbf{M}) \;. \end{aligned}$$

We want to reduce the variance of $\texttt{hpsi}(x)$ so we can choose $\mathbf{M}$ accordingly. The quantity $\frac{1}{2}\hbar\omega\alpha\mathbf{Tr}(\mathbf{M})$ would then be part of the zero-point energy.

One way to choose $\mathbf{M}$ is to use a variational approach; we would try various forms of $\mathbf{M}$ with different parameters to get the smallest variance of $\texttt{hpsi}(x)$. However that would be time consuming and dependent on the system/configuration under study.

A more general approach is to use these remaining quadratic terms to cancel quadratic terms in the potential:
In this context, that would mean using both terms to subtract the dipole-dipole parts of the total interaction: $(-q^2 x\mathbf{M}x - q^2\frac{\alpha}{2}x\mathbf{M}\mathbf{M}x = -\frac{1}{2}q^2 x\mathbf{T}x)$, but that involves taking the square root of a matrix which would be too expensive.
Instead, we will use one term $(-q^2 x\mathbf{M}x)$ as the second term is usually small.

Symmetrise the Perturbation

Expresssions are simpler when the matrix $\mathbf{M}$ is symmetric, so any potential terms that we wish to cancel should also be expressed in a symmetric form.

For example, the sum over all two-body dipole-dipole interactions:

$$\begin{aligned} \phi_{\text{dipole-dipole}} = & \quad \textstyle\sum_{i>j} q^2 x_i \mathbf{T}_{ij} x_j \ , & \text{triangular form} \\ \rightarrow & \quad \textstyle\sum_{i\neq j} \frac{1}{2} q^2 x_i \mathbf{T} i j x_j \ . & \text{symmetric form} \end{aligned}$$

This factor of $\frac{1}{2}$ will appear in every case when we sum up two-body interactions.

Note on Implementation

For implementation in a computational algorithm, we made the following rearrangement, which is possible if $\mathbf{M}$ is symmetric (for example, $\mathbf{T}$ is symmetric).

$$\begin{aligned} x\mathbf{M}^2 x \quad \text{means} \quad & x^T \mathbf{M}\mathbf{M} x \\ = \quad & (x^T \mathbf{M}^T)(\mathbf{M} x) \\ \equiv \quad & y^2 = \sum_i y_i^2 \ , \\ \text{where} \quad y_i \quad = \quad & M_{ij} x_j \ . \end{aligned}$$

As the vector $\mathbf{M}x$ gets used twice for `hpsi`(x) and once in the drift v, it is more efficient to store it once it is calculated. This kind of saving comes up with PIMD as well.

6.1.2 Simple Dipole Approximation

Using the form just described, we want an $\texttt{hpsi}(x)$ in which the real-space dipole-dipole part of the electrostatic potential is cancelled by a term from the wavefunction.

$$\begin{aligned} \phi_{\text{dpl}} &= -\tfrac{1}{2}q^2 x\mathbf{T}x \ ; \\ \text{we want} \quad \phi &\to \phi - \phi_{\text{dpl}} = \phi + \tfrac{1}{2}q^2 x\mathbf{T}x \ . \end{aligned}$$

So choose $\mathbf{M} = -\frac{1}{2}\mathbf{T}$:

$$\begin{aligned} \texttt{hpsi}(x) &= \phi(x) - q^2 x\mathbf{M}x - q^2\tfrac{\alpha}{2}x\mathbf{M}\mathbf{M}x + \tfrac{1}{2}\hbar\omega\alpha\mathbf{Tr}(\mathbf{M}) \\ &= \phi(x) + \tfrac{q^2}{2}x\mathbf{T}x - \tfrac{q^2\alpha}{8}x\mathbf{T}\mathbf{T}x - \tfrac{\alpha}{4}\hbar\omega\mathbf{Tr}(\mathbf{T}) \\ &= \phi(x) + \tfrac{q^2}{2}x\mathbf{T}x - \tfrac{q^2\alpha}{8}x\mathbf{T}\mathbf{T}x \\ &\qquad \because \mathbf{Tr}(\mathbf{T}) = 0 \quad \text{(in real space)}. \end{aligned}$$

The implied wavefunction turns out to be the first-order perturbed wavefunction that was derived in the dipole-limit chapter.

$$\Psi_{\text{T}} = e^{-\frac{m\omega_0}{2\hbar}\left(x^2 - \frac{\alpha}{2}\mu\mathbf{T}\mu\right)} \ .$$

Plug everything we need into our formulae:

$$v_{\text{drift}} \equiv \tfrac{\hbar}{m}\nabla \log \Psi_{\text{T}} = -\omega x + \tfrac{\alpha}{2}\omega\mathbf{T}x \ .$$

$$\texttt{hpsi}(x) = \phi(x) + \tfrac{q^2}{2}x\mathbf{T}x - \tfrac{q^2\alpha}{8}x\mathbf{T}\mathbf{T}x \ .$$

This leaves only one object that is expensive to calculate $\left[\mathcal{O}\left(N^2\right)\right]$:

$$\mathbf{M}x = -\tfrac{1}{2}\mathbf{T}x \qquad \Longrightarrow \qquad (\mathbf{M}x)_i = -\tfrac{1}{2}T_{ij}x_j \ .$$

6.1.3 Reaction Field Approximation

We can also adapt the trial wavefunction to cancel out the effective reaction electric-field that is used in the simulation (see Leeuw, Perram and Smith for the derivation[dLPS80]). In practice we found that this was too small a term to be bothered about, but we show the notes because they are interesting, and it is an illustration of how both wavefunction terms can be used.

This time the term we wish to cancel is the effective field that reacts to the total

dipole moment of the periodic simulation box:

$$\begin{aligned}\phi_{\text{react}} &\equiv +\tfrac{1}{2}\frac{4\pi}{\text{vol}}\frac{1}{2\epsilon_s+1}\left|\sum_i q\vec{x}_i\right|^2 \\ &= +\tfrac{1}{2}q^2 Rx\boldsymbol{\Sigma}x\ ,\end{aligned}$$

$$\text{where}\quad R \equiv \frac{4\pi}{\text{vol}}\frac{1}{2\epsilon_s+1}\ ,\quad \boldsymbol{\Sigma} \equiv \begin{pmatrix}\delta & \delta & \cdots & \delta\\ \delta & \delta & & \delta\\ \vdots & & \ddots & \vdots\\ \delta & \delta & \cdots & \delta\end{pmatrix}\ .$$

This time we can do clever things with the matrix $\boldsymbol{\Sigma}$ that allow us to cancel both terms with no extra (computational) work:

$$\begin{aligned}&\text{we want}\quad \phi \to \phi - \phi_{\text{react}} = \phi - \tfrac{1}{2}q^2 Rx\boldsymbol{\Sigma}x\ ;\\ &\implies -\tfrac{1}{2}q^2 Rx\boldsymbol{\Sigma}x = -q^2 x\mathbf{M}x - q^2\tfrac{\alpha}{2}x\mathbf{M}\mathbf{M}x\\ &\implies R\boldsymbol{\Sigma} = 2\mathbf{M} + \alpha\mathbf{M^2}\ .\end{aligned}$$

The following insight allows us to solve the equation exactly:

$$\boldsymbol{\Sigma}^2 \equiv \begin{pmatrix}\delta & \delta & \cdots & \delta\\ \delta & \delta & \cdots & \delta\\ \vdots & & & \\ \delta & \delta & \cdots & \delta\end{pmatrix}\begin{pmatrix}\delta & \delta & \cdots & \delta\\ \delta & \delta & & \delta\\ \vdots & \vdots & & \vdots\\ \delta & \delta & & \delta\end{pmatrix} = \begin{pmatrix}N\delta & N\delta & \cdots & N\delta\\ N\delta & N\delta & & N\delta\\ \vdots & & \ddots & \vdots\\ N\delta & N\delta & \cdots & N\delta\end{pmatrix} = N\boldsymbol{\Sigma}\ ,$$

$$\begin{aligned}\text{let}\quad \mathbf{M} &= M\boldsymbol{\Sigma}\ ,\\ \text{then}\quad R\boldsymbol{\Sigma} &= 2M\boldsymbol{\Sigma} + \alpha N M^2\boldsymbol{\Sigma}\ ,\\ \implies M &= \tfrac{1}{N\alpha}\left(\pm\sqrt{1+N\alpha R}-1\right)\ ,\\ M &= \tfrac{1}{N\alpha}\left(\sqrt{1+N\alpha R}-1\right)\qquad \text{(choose the smallest magnitude).}\end{aligned}$$

Now we need to know $\mathbf{Tr}(\mathbf{M})$ and $\mathbf{M}x$:

$$\mathbf{Tr}(\mathbf{M}) = M\mathbf{Tr}(\boldsymbol{\Sigma}) = 3NM,$$

$$\mathbf{M}x \equiv M\boldsymbol{\Sigma}x = M\begin{pmatrix}\delta & \delta & \cdots & \delta\\ \delta & \delta & \cdots & \delta\\ \vdots & & & \\ \delta & \delta & \cdots & \delta\end{pmatrix}\begin{pmatrix}\vec{x}_1\\ \vec{x}_2\\ \vdots\\ \vec{x}_N\end{pmatrix} = M\begin{pmatrix}\sum_i \vec{x}_i\\ \sum_i \vec{x}_i\\ \vdots\\ \sum_i \vec{x}_i\end{pmatrix} = M\vec{x}_{\text{tot}}\begin{pmatrix}1\\ 1\\ \vdots\\ 1\end{pmatrix},$$

$$\text{where} \quad \vec{x}_{\text{tot}} \equiv \sum_i \vec{x}_i \,, \quad M \equiv \tfrac{1}{N\alpha}\left(\sqrt{1+N\alpha R}-1\right) \,, \quad R \equiv \frac{4\pi}{\text{vol}}\frac{1}{2\epsilon_s+1} \,.$$

Finally plug them into our formulae:

$$\begin{aligned} \vec{v}_{\text{drift}} &= -\omega\vec{x} - \omega\alpha\mathbf{M}\vec{x} \\ &= -\omega\vec{x} - \tfrac{1}{N}\left(\sqrt{1+N\alpha R}-1\right)\omega\vec{x}_{tot} \,, \end{aligned}$$

$$\begin{aligned} \texttt{hpsi}(x) &= \phi(x) - q^2 x\mathbf{M}x - q^2\tfrac{\alpha}{2}x\mathbf{M}\mathbf{M}x + \tfrac{1}{2}\hbar\omega\alpha\mathbf{Tr}(\mathbf{M}) \\ &= \phi(x) - \phi_{\text{react}} + \tfrac{3}{2}\left(\sqrt{1+N\alpha R}-1\right)\hbar\omega, \end{aligned}$$

$$\text{where} \quad \mathbf{M} \equiv \tfrac{1}{N\alpha}\left(\sqrt{1+N\alpha R}-1\right)\mathbf{\Sigma} \,, \quad R \equiv \frac{4\pi}{\text{vol}}\frac{1}{2\epsilon_s+1} \,.$$

It turns out that the reaction energy is small. It is also interesting to note that it is a function of the number density (N/vol). This means that it will also fall in significance for larger systems, which will have correspondingly larger energies.

6.1.4 Reciprocal Dipole Approximation

We can further extend the dipole trial wavefunction so that it cancels all the dipole terms from both the real-space and reciprocal-space parts of the ewald sum.
This time we want

$$\begin{aligned} \phi &\rightarrow \phi - \phi_{\text{dipole}} \,, \\ \phi_{\text{dipole}} &= \phi_{\text{dipole,short}} + \phi_{\text{dipole,long}} \quad \text{(defined in the chapter on QDO's)} \\ &= x_{i\alpha}x_{j\beta}\nabla_{i\alpha}\nabla_{j\beta}\,\phi_{\text{dipole,short}} + x_{i\alpha}x_{j\beta}\nabla_{i\alpha}\nabla_{j\beta}\,\phi_{\text{dipole,long}} \,. \end{aligned}$$

We have 3 distinct parts, for which the theory is separable from one another, but in practice all 3 need to be applied as they cancel each other to some extent.

$\phi_{\text{dipole,short}}$ is straightforward, as we use the same tensor as for a pure dipole potential. We need to be a little careful, as this tensor has non-zero trace, and there will be effective self-interactions due to images outside the first box (The same is also true of

the long-range part).

$$\mathbf{Tr}(\mathbf{M}_{\rm short}) \;=\; \sum_{i,\ell>0} \tfrac{1}{2}q^2\,[1-{\rm damp}(|R|)]\,\tfrac{4\alpha_{\rm ew}^3}{\sqrt{\pi}}e^{-\alpha_{\rm ew}^2R^2}\;.$$

$\phi_{\rm dipole,long}$ is more complicated, as it is not immediately obvious that we can deconstruct it like this. It can be done, however: see the appendix.

$$\begin{aligned}
(\mathbf{M}x)_{i\alpha} &= \tfrac{1}{q^2}\tfrac{1}{2}\nabla_{i\alpha}\,\phi_{\rm dipole,long} \\
&= \tfrac{4\pi}{\det(\mathbf{h})}\sum_{\forall k\in\frac{1}{2}\mathbb{Z}^3}\left[\tfrac{1}{\kappa^2}e^{-\frac{\kappa^2}{4\alpha_{\rm ew}^2}}\,\kappa_\alpha(a_{\kappa,i}A_\kappa+b_{\kappa,i}B_\kappa)\right]-\tfrac{2\alpha_{\rm ew}^3}{3\sqrt{\pi}}x_{i\alpha}\;, \\
\mathbf{Tr}(\mathbf{M}_{\rm long}) &= \tfrac{1}{q^2}\tfrac{1}{2}\nabla^2\phi_{\rm dipole,long} \\
&= \tfrac{4\pi}{\det(\mathbf{h})}\sum_{\forall k\in\frac{1}{2}\mathbb{Z}^3}\left[\tfrac{1}{\kappa^2}e^{-\frac{\kappa^2}{4\alpha_{\rm ew}^2}}\,N\kappa^2\right]-\tfrac{2(3N)\alpha_{\rm ew}^3}{3\sqrt{\pi}}\;,
\end{aligned}$$

$$\begin{aligned}
\text{with}\quad \kappa &= 2\pi k\mathbf{h}^{-1}\;, \\
a_{\kappa,i} &= \cos(\kappa\cdot R_i)\;, \qquad A_\kappa=\sum_i(x_i\cdot\kappa)\cos(\kappa\cdot R_i)\;, \\
b_{\kappa,i} &= \sin(\kappa\cdot R_i)\;, \qquad B_\kappa=\sum_i(x_i\cdot\kappa)\sin(\kappa\cdot R_i)\;.
\end{aligned}$$

$\phi_{\rm react}$ was done in a precise way above but we used a simpler version (first order in $\mathbf{M}$) that avoids cross-terms with the dipole matrix.

$$\begin{aligned}
(\mathbf{M}x)_{i\alpha} &= \tfrac{1}{2}Rx_{{\rm tot},\alpha}\;, \quad \text{where}\quad \vec{x}_{\rm tot}=\sum_j\vec{x}_j \\
\mathbf{Tr}(\mathbf{R}) &= 3NR\;, \\
\text{where}\quad R &\equiv \frac{4\pi}{\rm vol}\frac{1}{2\epsilon_s+1}\;.
\end{aligned}$$

6.2 Trial Wavefunctions for the Full Model

The full model contains polarisability not just in the dipole limit but beyond. This means that the true wavefunction will be non-quadratic, but the efficiency of DMC would be improved if we could include some non-quadratic perturbations. There are 3

ways of attacking this problem:

Fitted Jastrow terms which suffer from the fact any simple functional form will not be very like the true wavefunction.

Multipole expansion which suffers from the fact that the series is divergent, and each term needs to be damped by a function of interatomic separation.

Direct Coulomb-perturbation which shows better promise, and can be damped with one function.

6.2.1 Fitted Jastrow Terms

One major flaw of the above wavefunctions is that they are very good approximations at long distance, but are very poor at short range, where the magnitude of the coulombic interactions are actually the greatest. One of the major sources of error is when a drude particle gets close to another drude particle or another nucleus. There is a damping term on the potential that prevents nasty singularities at $r = 0$, but at intermediate-small r, the trial wavefunction contains nothing that reflects the distribution of particles due to close-range repulsion or attraction. This means that the distribution is shaped almost entirely by the stochastic multiply/contract step of DMC, which results in a higher variance of hpsi, and ultimately requires much more sampling.

Here we discuss an attempt shape the wavefunction with a Jastrow function to reduce the reliance on the multiply/contract or population operator. In the end, it did not perform well, but we include it for interest.

Form and Implementation

But we can add terms to the potential in an empirical manner that helps to correct the distribution and reduce the variance of the sampling.

We tried the following form for the wavefunction:

$$\begin{aligned} \Psi_{\mathrm{T,jastrow}} &= \exp\left\{(ar-b)\left[1+\left(\tfrac{r}{c}\right)^8\right]^{-1}\right\} \\ &= \exp\left\{J(r)\right\} , \qquad \text{which defines } J(r(x)) . \end{aligned}$$

One set of parameters $a_{\mathrm{DD}}, b_{\mathrm{DD}}, c_{\mathrm{DD}}$ were chosen to try to match an empirical radial distribution of the drudes with other drudes, and another set $a_{\mathrm{DN}}, b_{\mathrm{DN}}, c_{\mathrm{DN}}$ to match that of drudes around other nuclei. If we combine this wavefunction by multiplying it

with our existing ones, this gives us:

$$\begin{aligned}\Psi_{\mathrm{T}} &= \exp\left\{-\tfrac{m\omega}{\hbar}\left(\tfrac{1}{2}x^2+\alpha F(x)\right)+J(x)\right\}\\ &= \exp\left\{-\tfrac{m\omega}{\hbar}\left(\tfrac{1}{2}x^2+\alpha F(x)-\tfrac{\hbar}{m\omega}J(x)\right)\right\}\ .\end{aligned}$$

But because we ignore the zero point energy of the harmonic oscillator, we can split up the wavefunction into two parts:

$$\begin{aligned}\Psi_{\mathrm{T}} &= \Psi_{\mathrm{SHO}}\Psi_{\Delta}\ ,\\ \Psi_{\mathrm{SHO}} &= \exp\left\{-\tfrac{m\omega}{2\hbar}x^2\right\}\ ,\\ \Psi_{\Delta} &= \exp\left\{-\tfrac{m\omega}{\hbar}\alpha F(x)+J(x)\right\}\ .\end{aligned}$$

By analogy from the multipole wavefunction, we can simply do:

$$\hat{H} = -\tfrac{\hbar^2}{2m}\nabla^2+\tfrac{1}{2}m\omega^2x^2-\tfrac{3N}{2}\hbar\omega+\phi \qquad \text{(as always)},$$

$$v_{\mathrm{drift}} \equiv \tfrac{\hbar}{m}\nabla\log\Psi_{\mathrm{T}} = -\omega x+\tfrac{\hbar}{m}\nabla\log\left(\Psi_{\Delta}\right)\ ,$$

$$\begin{aligned}\texttt{hpsi}(x) &\equiv \frac{1}{\Psi_{\mathrm{T}}}\hat{H}\Psi_{\mathrm{T}}\\ &= -\tfrac{\hbar^2}{2m}\left[\nabla\log\left(\Psi_{\mathrm{T}}\right)\right]^2-\tfrac{\hbar^2}{2m}\left[\nabla^2\log\left(\Psi_{\mathrm{T}}\right)\right]+\tfrac{1}{2}m\omega^2x^2-\tfrac{3N}{2}\hbar\omega+\phi\end{aligned}$$

$$\begin{aligned}-\tfrac{\hbar^2}{2m}\left[\nabla\log\left(\Psi_{\mathrm{T}}\right)\right]^2 &= -\tfrac{\hbar^2}{2m}\left[-\tfrac{m\omega}{\hbar}x+\nabla\log\left(\Psi_{\Delta}\right)\right]^2\\ &= -\tfrac{1}{2}m\omega^2x^2+\hbar\omega\ x\cdot\nabla\log\left(\Psi_{\Delta}\right)-\tfrac{\hbar^2}{2m}\left[\nabla\log\left(\Psi_{\Delta}\right)\right]^2\\ -\tfrac{\hbar^2}{2m}\left[\nabla^2\log\left(\Psi_{\mathrm{T}}\right)\right] &= -\tfrac{\hbar^2}{2m}\left\{-\tfrac{m\omega}{\hbar}(3N)+\nabla^2\log\left(\Psi_{\Delta}\right)\right\}\\ &= +\tfrac{3N}{2}\hbar\omega-\tfrac{\hbar^2}{2m}\nabla^2\log\left(\Psi_{\Delta}\right)\end{aligned}\ ,$$

$$\texttt{hpsi}(x) = \phi(x)+\hbar\omega\ x\cdot\nabla\log\left(\Psi_{\Delta}\right)-\tfrac{\hbar^2}{2m}\left[\nabla\log\left(\Psi_{\Delta}\right)\right]^2-\tfrac{\hbar^2}{2m}\nabla^2\log\left(\Psi_{\Delta}\right)\ .$$

We need to calculate $\nabla\log\left(\Psi_{\Delta}\right)$ and $\nabla^2\log\left(\Psi_{\Delta}\right)$:

$$\begin{aligned}\nabla\log\left(\Psi_{\Delta}\right) &= -\tfrac{m\omega}{\hbar}\alpha\nabla F(x)+\nabla J(x)\ ,\\ \nabla^2\log\left(\Psi_{\Delta}\right) &= -\tfrac{m\omega}{\hbar}\alpha\nabla^2F(x)+\nabla^2J(x)\ .\end{aligned}$$

The expressions for $\nabla J(x)$ and $\nabla^2 J(x)$ can be calculated using the chain-rule for differentiation. They are not complicated to implement, but they are complicated to write down and unenlightening, so I have omitted them.

Evaluation of this approach

We realised at a very early stage that the Jastrow corrections were contributing very little to the accuracy of calculations, suggesting we needed a more complex form than this. Initial results showed that dipole-moments and higher multipole corrections (see next section) actually made this form of the Jastrow function less useful, so we did not pursue it in detail.

6.2.2 Trial wavefunction with Multipole-Expansion

We can expand any two body potential as follows, using $\vec{R}$ as the internuclear coordinates, and $\vec{x}$ as the drude coordinates. As our Quantum Drudes are centred on pseudo-nuclei of equal and opposite charge, the $\mathcal{O}\left(x^0\right)$ term of the potential is cancelled. Then, at $\mathcal{O}\left(x^n\right)$, we get n terms:

$$\phi_{ij} = \sum_{n=1}^{\infty} \phi_{(n)ij} \ , \qquad \text{where} \quad \phi_{(n)ij} = q^2 \sum_{m=1}^{n-1} \tfrac{1}{m!}\tfrac{1}{(n-m)!} x_i^{(m)} \mathbf{T}_{(n)ij}(\vec{R}_i, \vec{R}_j) x_j^{(n-m)} \ ,$$

$$\begin{aligned} \text{where} \quad & \mathbf{T}_{(n)} && = \text{the } n^{th} \text{ rank tensor derivative of } \left(\tfrac{1}{R}\right) \ , \\ \text{so that} \quad & x_i^{(m)} && \equiv x_{i\alpha_1} x_{i\alpha_2} \cdots x_{i\alpha_m} \ , \\ \text{and} \quad & x_j^{(n-m)} && \equiv x_{j\alpha_{m+1}} \cdots x_{j\alpha_n} \ . \end{aligned}$$

Previously we had:

$$\Psi_{\mathrm{T}} \quad = \quad e^{-\frac{m\omega}{2\hbar}\left(x^2 + \alpha x \mathbf{M} x\right)} \ .$$

We can generalise this again:

$$\Psi_{\mathrm{T}} \quad = \quad \exp\left\{-\tfrac{m\omega}{\hbar}\left(\tfrac{1}{2}x^2 + \alpha F(x)\right)\right\} \ .$$

Which gives us a general expression for the energy estimator, (more below)

$$\frac{1}{\Psi_{\mathrm{T}}}\hat{H}\Psi_{\mathrm{T}} \quad = \quad \phi(x) - q^2 x \cdot \nabla F(x) - q^2 \tfrac{\alpha}{2}\left[\nabla F(x)\right]^2 + \tfrac{1}{2}\hbar\omega\alpha\nabla^2 F(x) \ .$$

It turns out that for QDO's in a Coulomb field, $\nabla^2 F$ is approximately zero, whereas $\left[\nabla F(x)\right]^2$ does not vanish, and is an order more complicated than F itself. As the variance of the energy estimator and the error of the wavefunction are related, we can improve the quality of the wavefunction by seeking to reduce the variance of the

estimator. We do this by choosing F such that strongly varying terms are cancelled, in particular the external potential:

$$\phi(x) = q^2 x \cdot \nabla F(x) \ . \tag{6.1}$$

In the dipole limit, this would return to the simple dipole approximation we made above, but now we have a result that is generally true.

Constructing the Perturbation

Now we choose a particular form of $F(x)$ to cancel terms in the potential $\phi(x)$:

$$\begin{aligned} \phi(x) = & \quad \textstyle\sum_{i>j} \phi_{ij}(x) \quad && \text{triangular form (as in the code)} \\ = & \quad \tfrac{1}{2} \textstyle\sum_{i \neq j} \phi_{ij}(x) && \text{symmetric form (clearer maths),} \\ & \text{where} \quad \phi_{ij}(x) && = \displaystyle\sum_{n=2}^{4} \phi_{(n)ij}(x) + \mathcal{O}\left(x^5\right) \ . \end{aligned}$$

Likewise

$$\begin{aligned} F(x) &= \sum_{i>j} F_{ij}(x) = \tfrac{1}{2} \sum_{i \neq j} F_{ij}(x) \ , \\ F_{ij} &\equiv \sum_{n=2}^{4} \tfrac{1}{n} f_{(n)ij}(x) \ , \qquad \text{where} \quad f_{(n)ij} \equiv \tfrac{1}{q^2} \phi_{(n)ij} \ . \end{aligned}$$

Notice that each component function $f_{(n)}$ is the n^{th} order term of the Taylor expansion of the potential. The extra factor of $\frac{1}{n}$ is added because

$$x \cdot \nabla f_{(n)ij} = \left[x_i \cdot \nabla_i + x_j \cdot \nabla_j \right] f_{(n)ij} = n f_{(n)ij} \ ,$$

$$\begin{aligned} \implies \quad q^2 x \cdot \nabla \left(\tfrac{1}{n} f_{(n)ij} \right) &= q^2 f_{(n)ij} = \phi_{(n)ij} \ , \\ \implies \quad q^2 x \cdot \nabla F_{ij} &= \phi_{ij} - \mathcal{O}\left(x^5\right) \ , \\ \implies \quad q^2 x \cdot \nabla F(x) &= \phi(x) - \mathcal{O}\left(x^5\right) \ . \end{aligned}$$

So $F(x)$ is derived from $\phi(x)$ but is different.

Plugging into DMC

We plug the new wavefunction into our standard formulae,
remembering to drop the zero-point energy $\left(\frac{3N}{2}\hbar\omega\right)$, and that $\alpha = q^2(m\omega^2)^{-1}$.
This gives us:

$$\hat{H} = -\tfrac{\hbar^2}{2m}\nabla^2 + \tfrac{1}{2}m\omega^2x^2 - \tfrac{3N}{2}\hbar\omega + \phi \,,$$

$$v_{\text{drift}} \equiv \tfrac{\hbar}{m}\nabla \log \Psi_{\text{T}} = -\omega x - \omega\alpha\nabla F(x) \,,$$

$$\begin{aligned}
\texttt{hpsi}(x) &\equiv \frac{1}{\Psi_{\text{T}}}\hat{H}\Psi_{\text{T}} \\
&= -\tfrac{\hbar^2}{2m}\left[\nabla \log\left(\Psi_{\text{T}}\right)\right]^2 - \tfrac{\hbar^2}{2m}\left[\nabla^2 \log\left(\Psi_{\text{T}}\right)\right] + \tfrac{1}{2}m\omega^2x^2 - \tfrac{3N}{2}\hbar\omega + \phi \,, \\
&\quad -\tfrac{\hbar^2}{2m}\left[\nabla \log\left(\Psi_{\text{T}}\right)\right]^2 = -\tfrac{\hbar^2}{2m}\left[-\tfrac{m\omega}{\hbar}\left(x + \alpha\nabla F(x)\right)\right]^2 \\
&\qquad\qquad = -\tfrac{1}{2}m\omega^2x^2 - m\omega^2\alpha x\cdot\nabla F(x) - m\omega^2\tfrac{\alpha^2}{2}\left(\nabla F(x)\right)^2 \,, \\
&\quad -\tfrac{\hbar^2}{2m}\left[\nabla^2 \log\left(\Psi_{\text{T}}\right)\right] = -\tfrac{\hbar^2}{2m}\left\{-\tfrac{m\omega}{\hbar}\left(3N + \alpha\nabla^2F(x)\right)\right\} \\
&\qquad\qquad = +\tfrac{3N}{2}\hbar\omega + \tfrac{1}{2}\hbar\omega\alpha\nabla^2F(x) \,,
\end{aligned}$$

$$\begin{aligned}
\texttt{hpsi}(x) &= \phi(x) - m\omega^2\alpha x\cdot\nabla F(x) - m\omega^2\tfrac{\alpha^2}{2}\left[\nabla F(x)\right]^2 + \tfrac{1}{2}\hbar\omega\alpha\nabla^2F(x) \\
&= \phi(x) - q^2x\cdot\nabla F(x) - q^2\tfrac{\alpha}{2}\left[\nabla F(x)\right]^2 + \tfrac{1}{2}\hbar\omega\alpha\nabla^2F(x) \,.
\end{aligned}$$

Energy Terms

We have a term that cancels the potential up to $\mathcal{O}\left(R^{-6}\right)$:

$$\begin{aligned}
x\cdot\nabla F(x) &= f_{(2)}(x) + f_{(3)}(x) + f_{(4)}(x) \\
&= \tfrac{1}{q^2}\phi(x) \quad - \mathcal{O}\left(x^5R^{-6}\right) \,.
\end{aligned}$$

And we have a remainder term, which is also $\mathcal{O}\left(R^{-6}\right)$:

$$\begin{aligned}
\nabla F(x) &\approx f_{(2)}(x) \approx \mathcal{O}\left(xR^{-3}\right) \,, \\
\implies \left(\nabla F(x)\right)^2 &\approx \mathcal{O}\left(x^2R^{-6}\right) \,.
\end{aligned}$$

Beyond R^{-6}, 3-body terms (for example triple-dipole interactions) appear in this term, and would need similar terms to appear in the wavefunction to cancel them.

Fortunately this choice of $F(x)$ allows us to omit $\nabla^2 F(x)$ because $\nabla^2 f_{(n)}(x)$ is identically zero for all n (see appendix), which derives from the fact that $\nabla^2\phi(x)$ equals zero, and therefore all its tensor derivatives are also zero. Note, this is only true of this two-body multipole expansion form, which is a special case. It is not true of the following wavefunctions, for which $\nabla^2 F(x)$ may be non-zero. That leaves

$$\texttt{hpsi}(x) \;=\; \phi(x) - q^2 x\cdot\nabla F(x) - q^2\tfrac{\alpha}{2}\left[\nabla F(x)\right]^2 \ .$$

All that remains now is to evaluate the vector $\nabla F(x)$, for each term in the expansion.

Extension to Reciprocal Space?

It would be possible to extend the multipole expansion into reciprocal space if necessary, in a manner analogous to the dipole limit, but we have not examined this in great detail yet, as these are not the major contributors to the energy.

6.2.3 Trial wavefunction with Many-Body terms

The kind of wavefunction we used above is perfectly general:

$$\Psi_{\rm T} \;=\; \exp\left\{-\tfrac{m\omega}{\hbar}\left(\tfrac{1}{2}x^2 + \alpha F(x)\right)\right\} \ .$$

We generalise it again slightly by $\alpha F(x) \to G(x)$:

$$\Psi_{\rm T} \;=\; \exp\left\{-\tfrac{m\omega}{\hbar}\left(\tfrac{1}{2}x^2 + G(x)\right)\right\} \ .$$

Again, it gives us:

$$\texttt{hpsi}(x) \;=\; \phi(x) - \tfrac{q^2}{\alpha}x\cdot\nabla G(x) - \tfrac{q^2}{2\alpha}\left(\nabla G(x)\right)^2 + \tfrac{1}{2}\hbar\omega\nabla^2 G(x) \ .$$

Our approach in eqn. 6.1 above was to use $\left[-\frac{q^2}{\alpha}x\cdot\nabla G\right]$ to cancel simply $\phi(x)$.
We can go further by using $\left[-\frac{q^2}{\alpha}x\cdot\nabla G\right]$ to cancel not only $\phi(x)$ but also successive orders of $\left[\frac{q^2}{2\alpha}\left(\nabla G\right)^2\right]$ and $\left[\frac{1}{2}\hbar\omega\nabla^2 G\right]$, except for terms that come out as constants, which we will put into successive terms of the energy E.

In algebra, we break G into sets $G_{n,m}$, where n is the number of times a coulomb interaction (yellow box) appears in the term, and m is the number of extra contractions

(more than the minimum required to connect all the instances). These can be visualised in the diagrams in section 4.4 (figs. 4.2, 4.3 and 4.4). Thus n refers to the column of the diagram, counting from $n = 1$, while m refers to the row, counting from $m = 0$.

$$G(x) \quad = \quad G_{1,0} \; + \; [G_{2,0} + G_{2,1} + \cdots] \; + \; [G_{3,0} + G_{3,1} + \cdots] \; + \; \cdots \; .$$

They are related as follows:

$$\begin{aligned} x\cdot\nabla G_{1,0}(x) &\equiv \tfrac{\alpha}{q^2}\phi(x) \; , \\ x\cdot\nabla G_{n,m}(x) &\equiv \tfrac{1}{2}\sum_{\nu=1}^{n-1}\sum_{\mu=0}^{m}[\nabla G_{\nu,\mu}(x)]\cdot[\nabla G_{n-\nu,m-\mu}(x)] \; , \\ &\quad + \tfrac{\alpha}{2q^2}\hbar\omega\nabla^2 G_{n,m-1}(x) \; . \end{aligned}$$

Each term in G is trivial to integrate if each is a polynomial of degree integer ℓ; the $x\cdot\nabla$ operation simply multiplies by ℓ. This series can be 'seeded' by writing down the potential ϕ as a multipole expansion.

$$\phi \quad = \quad \sum_{\ell=0}^{\infty}\phi_\ell, \qquad \text{where} \quad \phi_\ell \equiv \sum_k \tfrac{1}{(k)!(\ell-k)!}x^k\mathbf{T}^{(\ell)}x^{\ell-k} \; .$$

Terms of the interaction energy can be read off by summing up terms that are constant with respect to x. This is how the the expansion in fig. 4.2 on pg. 52 was generated.

$$E_{n,m} = \tfrac{q^2}{2\alpha}G_{n,m}(\not{x}) = \tfrac{q^2}{2\alpha}G_{n,m}(x=0) \; .$$

The terms which are of order $1 = (\hbar\omega)^0$, (the set $E_{n,0}$) would all exist in the classical limit. That means that they exist for classical drude oscillators. If you look at the diagram in the QDO chapter, these are all the terms in the top row. Note that there are rather a lot of them, and they get quickly very complex. The Quantum Drude has even more, and one of the great advantages of the Quantum Drude model, is that these energies can be sampled directly without calculating individual terms.

The above derivation shows how many-body multipole terms could be added iteratively to a DMC wavefunction, so give more than just an energy expansion. Unfortunately this approach does not transfer well to PIMD beyond quadratic terms (that is, dipole-dipole). Additionally, the multipole expansion in general is not a good description when charges get too close to one another (the series diverges, and would require damping on each term). The following solution should be better able to cope with this.

6.3 Trial wavefunction for Coulomb Interaction

The dipole-limit expansion allows us to project out about half the (ground-state Born-Oppenheimer) interaction energy for a Quantum Drude model Xenon dimer near its minimum (see fig. 7.1 in the next chapter), and this is because the dipole-dipole interaction is only about half the interaction energy at that point.

Thus we tried a multipole expansion. The great advantage of a multipole expansion is that it generates polynomial expressions that are easy to integrate term by term, like the dipole-limit. The major disadvantage is that the series diverges when a particle gets too close to the other charge. A multipole expansion could be used, but it would require damping on every term, requiring a very involved parameter-fitting procedure.

It would be far better to find a single function $F(x)$ that cancels the entire coulomb potential $\phi(x)$ in one. Following the pattern established above, in eqn. 6.1, this means finding a function that satisfies

$$x\cdot\nabla F(x) \quad = \quad \tfrac{1}{q^2}\,\phi(x)\ .$$

Perhaps surprisingly, it turns out that there is a simple function that does this even for the off-centre coulomb potentials that perturb Quantum Drudes in the full model.

$$x\cdot\nabla F(x) \quad = \quad \frac{1}{|x-R|}\ . \tag{6.2}$$

There is a small complication because we want ∇F to be well-behaved at $x=0$. This means that we must have $(x\cdot\nabla F)|_{x=0}=0$, so we need to subtract a $\frac{1}{|R|}$ from the right hand side and add it to the Hamiltonian as a constant. This problem did not arise for the dipole-limit, or multipole expansion because we did not consider the 0^{th} multipole moment (charge) as we do now. Thus we have:

$$x\cdot\nabla F(x) \quad = \quad \frac{1}{|\vec{x}-\vec{R}|}-\frac{1}{|\vec{R}|}\ . \tag{6.3}$$

$$\tag{6.4}$$

6.3.1 Solution of F

Transform the equation into spherical polars, where it simplifies:

$$\begin{aligned} r\tfrac{\partial}{\partial r}F(r;\theta,\phi) \quad &= \quad \frac{1}{|\vec{r}-\vec{R}|}-\frac{1}{|\vec{R}|}\ , \\ F(r;\theta,\phi) \quad &= \quad \int_0^r \frac{1}{r}\left[\tfrac{1}{|\vec{r}-\vec{R}|}-\frac{1}{|\vec{R}|}\right]\ . \end{aligned}$$

Now let $\vec{R}$ be along the polar axis:

$$F(r;\theta,\phi) \;=\; \int_0^r \frac{1}{r\sqrt{r^2+R^2-2rR\cos\theta}} - \frac{1}{rR}\,.$$

Then integrate (we used Wolfram online integrator):

$$F(r;\theta,\phi) \;=\; -\tfrac{1}{R}\ln\left[R\sqrt{r^2+R^2-2rR\cos\theta}+R\left(R-r\cos\theta\right)\right]\,.$$

And transform back to cartesian coordinates.

$$\begin{aligned} F(x) &= -\tfrac{1}{|R|}\ln\left[|R|\,|R-x|+R\cdot(R-x)\right]\,, \\ \nabla F(x) &= \frac{1}{|R|\,|R-x|}\,\frac{|R|\,(\vec{R}-\vec{x})+|R-x|\,\vec{R}}{\left[|R|\,|R-x|+R\cdot(R-x)\right]}\,, \\ \nabla^2 F(x) &= 0\,. \end{aligned}$$

Sum F contributions

For a particular set of molecules, each coulomb interaction as to be cancelled by an F-like term. For an example, we can look at the Xenon model (described later). All we need to know right now is that each Quantum-Drude is centred on a nucleus of opposite charge. This allows some simplification.

$$\begin{aligned} \phi_{\text{total}} &= \sum_{j>k}\frac{q_{\text{D}}q_{\text{D}}}{|R_{jk}+x_k-x_j|}+\frac{q_{\text{D}}q_{\text{n}}}{|R_{jk}-x_j|}+\frac{q_{\text{n}}q_{\text{D}}}{|R_{jk}+x_k|}+\frac{q_{\text{n}}q_{\text{n}}}{|R_{jk}|} \\ &= \sum_{j>k}\left[\frac{q_{\text{D}}q_{\text{D}}}{|R_{jk}+x_k-x_j|}-\frac{q_{\text{D}}q_{\text{D}}}{|R_{jk}|}\right]+\left[\frac{q_{\text{D}}q_{\text{n}}}{|R_{jk}-x_j|}-\frac{q_{\text{D}}q_{\text{n}}}{|R_{jk}|}\right] \\ &\quad+\left[\frac{q_{\text{n}}q_{\text{D}}}{|R_{jk}+x_k|}-\frac{q_{\text{n}}q_{\text{D}}}{|R_{jk}|}\right]+\frac{q_{\text{D}}q_{\text{D}}}{|R_{jk}|}+\frac{q_{\text{n}}q_{\text{D}}}{|R_{jk}|}+\frac{q_{\text{D}}q_{\text{n}}}{|R_{jk}|}+\frac{q_{\text{n}}q_{\text{n}}}{|R_{jk}|}\,. \end{aligned}$$

Thus, our total F perturbation is,

$$F_{\text{total}} \;=\; \sum_{j>k} q_{\text{D}}q_{\text{D}}\,F\left(x_j-x_k;R_{jk}\right)+\sum_{j\neq k} q_{\text{D}}q_{\text{n}}\,F\left(x_j;R_{jk}\right)\,,$$

and the overall energy contains a constant, which disappears for a neutral atom, where $q_{\rm D} + q_{\rm n} = 0$.

$$\phi(0) \quad = \quad \frac{(q_{\rm D} + q_{\rm n})^2}{|R_{jk}|} \; .$$

6.3.2 Wavefunction and Corrections

The above derivation assumed a wavefunction of the following form

$$\Psi_{\rm T} \quad = \quad \exp\left\{-\tfrac{m\omega}{\hbar}\left(\tfrac{1}{2}x^2 + \alpha F(x)\right)\right\} \; .$$

And the formulae we derived above (for multipoles):

$$\hat{H} \quad = \quad -\tfrac{\hbar^2}{2m}\nabla^2 + \tfrac{1}{2}m\omega^2 x^2 - \tfrac{3N}{2}\hbar\omega + \phi \; ,$$

$$v_{\rm drift} \quad \equiv \quad \tfrac{\hbar}{m}\nabla \log \Psi_{\rm T} = -\omega x - \omega\alpha\nabla F(x) \; ,$$

$$\texttt{hpsi} \quad = \quad \cancel{\phi(x)} - \cancel{q^2 x\cdot\nabla F(x)} - q^2\tfrac{\alpha}{2}\left[\nabla F(x)\right]^2 + \tfrac{1}{2}\hbar\omega\alpha\nabla^2 F(x) \; .$$

The remaining error involves $\left[\nabla F(x)\right]^2$, and is related to 3-body correlations.

6.3.3 Simple Damping

The F function causes problems because it has a singularity in the potential at $|x - R| = 0$, which is due to the singularity in the potential that it is designed to cancel.

The potential we actually use is damped, but it is difficult to integrate, so we use an approximate method for damping, with a new parameter σ.

$$\begin{aligned} x\cdot\nabla F(x) \quad &\approx \quad \left(1 - e^{-|x-R|^4/\lambda^4}\right)\tfrac{1}{|x-R|} - \left(1 - e^{-|R|^4/\lambda^4}\right)\tfrac{1}{|R|} \; , \\ F(x) \quad &\equiv \quad -\tfrac{1}{|R|}\ln\left\{\mathrm{damp}_\sigma\left[|R|\,|R-x| + R\cdot(R-x)\right]\right\} \; . \end{aligned}$$

Unfortunately this simple form leads to very complex expressions, but these can be

kept manageable using the chain-rule for differentiation, both here and in the code.

$$\begin{aligned}
\nabla F &= -\tfrac{1}{|R|}\frac{1}{\mathrm{damp}_\sigma}\frac{\partial\,\mathrm{damp}_\sigma}{\partial y}\,\nabla y\;, \\
\nabla^2 F &= \tfrac{1}{R^2}\left(\tfrac{1}{\mathrm{damp}_\sigma}\tfrac{\partial\,\mathrm{damp}_\sigma}{\partial y}\nabla^2 y + \tfrac{1}{\mathrm{damp}_\sigma}\tfrac{\partial^2\,\mathrm{damp}_\sigma^2}{\partial y}\,[\nabla y]^2 - \tfrac{1}{\mathrm{damp}_\sigma^2}\left[\tfrac{\partial\,\mathrm{damp}_\sigma}{\partial y}\nabla y\right]^2\right)\;, \\
&\quad \text{where } y \text{ stands for } \left[|R|\,|R-x| + R\cdot(R-x)\right]\;.
\end{aligned}$$

The simple damping method is the one we used for most of our computational work, as it worked well enough for our requirements. It turned out that, for each kind of damped interaction, Drude-Drude, nucleus-Drude, we could use exactly the same damping parameter, e.g. $\sigma_{\mathrm{DD}} = \gamma\mathrm{DD}$, $\sigma_{\mathrm{nD}} = \gamma\mathrm{nD}$

Unfortunately, this simple approach also has some potential problems. The damped F-function implies a potential, via the expression $\phi = \phi_0 + x\cdot\nabla F$. By damping the argument of the log function, we introduce anisotropies around the charge centre. The shape of the well is also the wrong shape. We tried replacing the existing damped potential with this implicit potential but found that it did not perform well (no close-range repulsion emerged). If we could create an F with an implicit potential that matched the existing fitted potential reasonably well, we could drop it altogether, given that they are intended to cancel each other.

6.3.4 Intelligent Damping

The function F was constructed as a function of the polar coordinates for x; r and θ. Because the function is constructed via an integration in r, there is a θ-dependent constant of integration, but it can be set to zero if the function is written in the following form:

$$\begin{aligned} F(r,\theta) &= -\tfrac{1}{|R|}\log\left[|R||R-x| + R\cdot(R-x)\right] + \cancel{c(\theta)} \\ &= -\tfrac{1}{|R|}\log\left[|R||R-x|(1-\cos\alpha)\right] \ , \qquad \text{where} \quad \cos\alpha = \tfrac{R(x-R)}{|R||R-x|} \ . \end{aligned}$$

The problem with this expression is that it is infinite-valued where $\cos\alpha = 1$, caused by integrating through the singularity at $|R-x| = 0$. This infinite value arises when $(1-\cos\alpha) \to 0$, so we would like to rearrange the expression to remove this singularity:

$$(1-\cos\alpha) = \frac{\sin^2\alpha}{(1+\cos\alpha)} = \frac{|x|^2}{|x-R|^2}\frac{\sin^2\theta}{(1+\cos\alpha)} \ .$$

Plugging this back in, we can find a solution that is finite for $\cos\alpha = 1$ (though now infinite for $\cos\alpha = -1$), and the singularity is pulled out into an expression purely dependent on θ, which can be absorbed into the integration constant $c(\theta)$.

$$\begin{aligned} F(r,\theta) &= -\tfrac{1}{|R|}\log\left[|R||R-x|(1-\cos\alpha)\right] \\ &= -\tfrac{1}{|R|}\log\left[|R||R-x|\frac{|x|^2}{|x-R|^2}\frac{\sin^2\theta}{(1+\cos\alpha)}\right] \\ &= -\tfrac{1}{|R|}\log\left[\frac{|x|^2|R|^2\sin^2\theta}{|R||x-R|(1+\cos\alpha)}\right] \ , \end{aligned}$$

$$\begin{aligned} F(r,\theta) &= -\tfrac{1}{|R|}\log\left[\frac{|x|^2|R|^2\sin^2\theta}{|R||x-R| + R\cdot(x-R)}\right] + c(\theta) \\ &= +\tfrac{1}{|R|}\log\left[|R||x-R| + R\cdot(x-R)\right] \\ &\qquad -\tfrac{1}{R}\log\left[|x|^2|R|^2\right] - \tfrac{1}{R}\log\left[\sin^2\theta\right] + c(\theta) \ . \end{aligned}$$

This function tells us what F looks like on the far side of a damped charge (its asymptotic form). What we need to do now is to develop a damped function to stitch them together correctly in the vicinity of the charge. For example, if we want a constant, flat potential in the vicinity of the charge:

$$F(r,\theta) = \int \frac{V_0}{r} = V_0\log r + c(\theta) \ .$$

Thus there are 3 parts to this function. The integration constants have been pulled within the log bracket.

$$\begin{aligned} F(r,\theta) \approx & \quad -\tfrac{1}{|R|}\log\left[|R||R-x| + R\cdot(R-x)\right] && 0 < r < R_-(\theta) \\ \approx & \quad +\tfrac{1}{|R|}\log\left[c_1(\theta)\, r^{V_c}\right] && R_-(\theta) < r < R_+(\theta) \\ \approx & \quad +\tfrac{1}{|R|}\log\left[c_2(\theta)\, \tfrac{|R||x-R|+R(x-R)}{|x|^2|R|^2}\right] && R_+(\theta) < r\ . \end{aligned}$$

These could all be put under a single log bracket, but unfortunately, the outer parts have opposite sign, forcing one or other to have a singularity at $|x-R| = 0$. On the other hand, we want the damping to be inside the brackets because $\log(0)$ is singular. This is a nuisance. We need to define suitable switching functions.

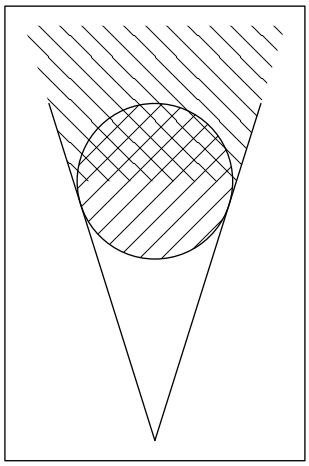


Figure 6.1: 3 regions requiring 3 forms for the F-function. The region left unshaded is where the existing (undamped) $F(r;\theta)$ is accurate. The circular area (// shading) is the region in which the coulomb potential is damped, but where both the existing F and the transformed F would be diverge at $\theta = 0$, so a smooth F is needed. The region beyond this (\\ shading) is where the transformed $F(r;\theta)$ is accurate, up to a θ-dependent integration factor.

Fitting Procedure for the damped F

Produce an numerical function $F(r)$ by integrating $\phi(r-R)/r$ numerically, well into the far region where the potential damping has disappeared, and then fit the tail (transformed) function, which means fitting the the θ-dependent integration constant $c_2(\theta)$ above.

To allow both forms to be damped, we need to damp out the branch point at $|r-R|=0$, from both log-functions, before we attempt to use a switching function. This involves damping the argument of the log-function as for the simple damping case; we can use the same damping function with a tighter radial parameter.

Then the front and back forms can be 'stitched' using a suitable switching function, e.g. $\frac{1}{2}\left(\tanh\left[r-r_0(\theta)\right]-1\right)$ and finally the central form can be 'stitched' in using an appropriate short-range damping form.

This work is in progress. The branch points were a major headache, but using the prescription just described, we anticipate that it will now be straightforward to fit F-functions to arbitrary damped-coulomb potentials.

Chapter 7

Xenon - a QDO Model

The ability of a full QDO model to treat Xenon solid will be stringently tested using the accurate DMC methods described in the previous chapter. We begin by describing the Quantum Drude model of Xenon that was originally constructed and fitted to gas-phase measurements by Martyna and Whitfield[WM06], showing in particular how the basic parameters of the Quantum Drude were chosen.

Next, using the DMC methods described in chapter 4, we perform DMC simulations, including convergence tests, for the Xenon dimer, and Xenon FCC-crystal in the dipole-limit and the full model. We graph the Born-Oppenheimer energy surface that exists between two QD Xenon atoms, and demonstrate how well each trial wavefunction contributes to the results, including the 'Coulomb' trial wavefunction that performs particularly well for the full Quantum Drude model. For the dipole-limit FCC-crystal, we have analytic results from matrix diagonalisation, which they reproduce perfectly. For the full model, we calculate an accurate lattice constant plus accurate bulk modulus for a QD Xenon FCC-crystal which both come out very close to experimental values, showing that the way Quantum Drudes capture many-body dispersion makes them truly transferable between gas and condensed phases, demonstrating the potential of Quantum Drudes in general to produce highly-transferable, widely useful molecular models.

7.1 The Model

The Quantum-Drude is a harmonic oscillator having two parameters m and ω that describe its quantum distribution,

$$\hat{H}_0 = -\tfrac{\hbar^2}{2m}\nabla^2 + \tfrac{1}{2}m\omega^2 x^2.$$

plus q its charge, which defines the strength of its interaction with other charges (including other Quantum-Drudes).

It interacts with other Quantum Drudes and other nuclei through damped coulomb potentials. The nuclei interact via coulomb potentials plus a repulsive correction which attempts to take account of short range forces, especially Pauli exclusion which resists overlap between atoms.

$$\begin{aligned} V(x,r) &= \tfrac{1}{2}\sum_{j\neq k} V_{\mathrm{DD}}\left(|x_j + r_j - x_k - r_k|\right) \\ &\quad + \sum_{jk} V_{\mathrm{nD}}\left(|x_j + r_j - r_k|\right) \\ &\quad + \tfrac{1}{2}\sum_{j\neq k} V_{\mathrm{nn}}\left(|r_j - r_k|\right) \\ &\quad + \tfrac{1}{2}\sum_{j\neq k} V_{\mathrm{rep}}\left(|r_j - r_k|\right) \ , \end{aligned}$$

where the potentials have a standard form,

$$\begin{aligned} V_{\mathrm{DD}}(r) &= \frac{q_{\mathrm{D}} q_{\mathrm{D}} \left(1 - e^{-(r/\gamma_{\mathrm{DD}})^4}\right)}{r}, \\ V_{\mathrm{nD}}(r) &= \frac{q_{\mathrm{n}} q_{\mathrm{D}} \left(1 - e^{-(r/\gamma_{\mathrm{nD}})^4}\right)}{r}, \\ V_{\mathrm{nn}}(r) &= \frac{q_{\mathrm{n}} q_{\mathrm{n}} \left(1 - e^{-(r/\gamma_{\mathrm{nn}})^4}\right)}{r}, \\ V_{\mathrm{rep}}(r) &= \kappa_1\, e^{-\lambda_1 r} + \kappa_2\, e^{-\lambda_2 r} + \kappa_3\, e^{-\lambda_3 r}. \end{aligned}$$

A xenon atom is charge-neutral, so the charge on the Quantum Drude and that on the nucleus must cancel: $q_{\mathrm{D}} = -q_{\mathrm{n}}$. The mass of a Xenon atom is also known: $m_{\mathrm{Xe}} = 131.29$ a.m.u. The damping parameter between two nuclei should not matter greatly as the nuclei should never get too close, and in any case, its effect can be balanced by modifications to the repulsive correction. For aesthetic reasons only, Martyna and Whitfield [WM06] chose $\gamma_{\mathrm{nn}} = \gamma_{\mathrm{DD}}$, although there is no real relationship between these two parameters.

That leaves the following parameters to be fitted:

$$\begin{array}{ccc} q_{\mathrm{D}} & m & \omega \\ \gamma_{\mathrm{DD}} & \gamma_{\mathrm{nD}} & \end{array}$$

$$\begin{array}{cccc} \kappa_1 & \kappa_2 & \kappa_3 & \\ \lambda_1 & \lambda_2 & \lambda_3 & . \end{array}$$

7.1.1 Fitting Polarisabilities

In a previous chapter (section 4.3), we derived the following relations,

$$\begin{aligned} \alpha_1 &= \frac{q^2}{\mu\omega^2} , & C_6 &= \tfrac{3}{4}\alpha_1\alpha_1\hbar\omega , \\ \alpha_2 &= \tfrac{3}{4}\left(\tfrac{\hbar}{\mu\omega}\right)\alpha_1 , & C_8 &= 5\left(\tfrac{\hbar}{\mu\omega}\right)\times C_6 , \\ \alpha_3 &= \tfrac{5}{4}\left(\tfrac{\hbar}{\mu\omega}\right)^2\alpha_1 , & C_{10} &= \tfrac{245}{8}\left(\tfrac{\hbar}{\mu\omega}\right)^2\times C_6 , \\ & & C_9 &= \tfrac{1}{4}\alpha_1\times C_6 , \end{aligned}$$

even though they hold perfectly only for Quantum-Drudes, they suggested the following fitting procedure,

$$\begin{aligned} &(1) \quad \omega = \frac{1}{\hbar}\frac{4C_6}{3\alpha_1^2} , \\ &(2) \quad m = \frac{\hbar}{\omega}\frac{3\alpha_1}{4\alpha_2} \qquad \text{or} \qquad m = \frac{\hbar}{\omega}\frac{5C_6}{C_8} , \\ &(3) \quad q = \pm\sqrt{m\omega^2\alpha_1} . \end{aligned}$$

Martyna and Whitfield [WM06] generated Quantum-Drude parameters by this procedure, using the following values for α_1, C_6, C_8, and neglecting α_2, as well as α_3 and C_{10}.

$$\begin{array}{lcl} \alpha_1 = 27.3 & & \omega = 0.5152 \\ C_6 = 288 & \Longrightarrow & m = 0.2541 \\ C_8 = 11000 & & q_{\mathrm{D}} = -1.357 . \end{array}$$

Then we can plug these back in to the polarisation / dispersion relations, and compare

the results with Certain's newer paper [SC85]. Note that this means the values for C_6 and C_8 no longer match perfectly, even though they were just used for fitting. Nevertheless, this gives a good idea for what level of error still exists.

		Property	Value	Target	Deviation
		α_1	27.30	27.30±0.03	0.0%±0.0%
q	−1.357	α_2	156.4	212.6±0.2	−26.4%±0.0%
m	0.2541	α_3	1991	3602±4	−44.7%±0.1%
ω	0.5152	C_6	288.0	298.5±26.5	−3.52%±8.88%
		C_8	11000	11400±2500	−3.51%±21.93%
		C_{10}	514700	551500±123500	−6.67%±22.39%

Table 7.1: Polarisation and Dispersion coefficients for the Xenon model

7.1.2 Fitting Damping

As the damping contributes only part of the repulsive potential, and the rest of the potential is added as an explicit force between the nuclei, there is no simple way to determine exactly what the damping parameters should be. However, Martyna and Whitfield discovered that the behaviour is fortunately not overly too sensitive to the damping parameters [WM06]. In the end, the damping parameters were chosen by eye, by their effect on the shape of the Born-Oppenheimer surface they generated at close range. the set that best reproduced the experimental gas-phase potential[BWL$^+$74] at intermediate- and long-range, thereby limiting the explicit repulsion correction to the short-range.

7.2 DMC Simulations and Results

Here, the ability the trial wavefunctions developed for use with the QDO model to treat model and realistic systems, is explored. The QDO trial wavefunctions are tested through NC-DMC and variational Monte Carlo (VMC) studies of the xenon dimer and the fcc-xenon crystal. Where possible comparisons are made to analytical or experimental results as applicable. The notation N_{Xe} will be used to denote the number of xenon Drude oscillator atoms and N the number of NC-DMC walkers, below. Diffusion Monte Carlo with rejection[UNR93] appropriate for use with the more complex trial functions has not been implemented.

7.2.1 Xenon Dimer, full QDO model, at T=0

It is useful to evaluate the quality of the QDO trial functions presented in the DMC chapter through tests on the most basic system to which they can be applied, the xenon dimer. In fig. 7.1, the variational dimer energy as a function of nuclear separation, $E_{\mathrm{T}}(R)$, is given for some of the trial wavefunctions described in the text, the on-site-only trial function, the on-site-plus-dipole-dipole trial function, some related multipole trial functions, and the on-site-plus-pair-coulomb model .

Results are compared to a high quality NC-DMC simulation estimate of $E_0(R)$ (the correct results). The dipole-dipole trial function is a multipole expansion correct to order $\mathcal{O}\left(\mathbf{x}^2\right)$, whereas the dipole-quadrupole trial function and dipole-octopole/quadrupole-quadrupole trial function are the expansions correct to $\mathcal{O}\left(\mathbf{x}^3\right)$ and $\mathcal{O}\left(\mathbf{x}^4\right)$ respectively (that is the dipole-quadrupole trial function builds on the dipole-dipole trial function and so on). These results illustrate the fact that a multipole expansion (in this case for a trial function) will converge at long range, but not at close range.

The on-site-plus-pair-coulomb wavefunction performs particularly well, even at very close range, predicting the well depth to within 10% percent.

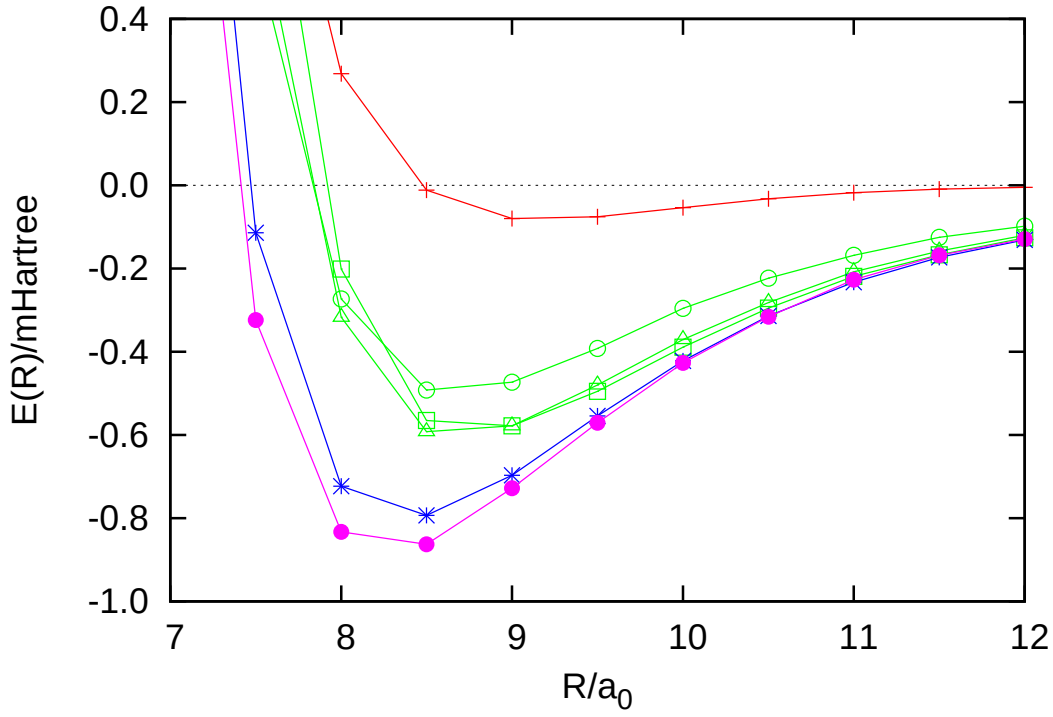
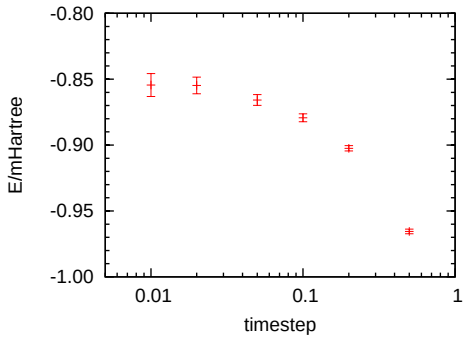
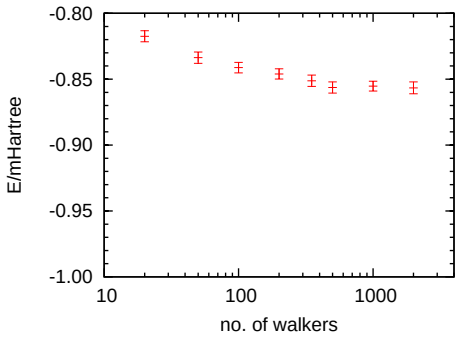


Figure 7.1: Variational energy $E_{\mathrm{T}}(R)$ vs nuclear separation, for the full QDO model xenon dimer is presented for 3 trial wavefunctions described in the text, from top to bottom: the on-site-only trial function (red), the on-site-plus-dipole-dipole trial function (green circle), the on-site-plus-dipole-quadrupole trial function (green triangle), the on-site-plus-dipole-octopole/quadrupole-quadrupole trial function (green square) and the on-site-plus-pair-coulomb trial function (blue). Results are compared to $E_0(R)$, calculated from a converged NC-DMC study (purple).

Next, the convergence of the ground state energy of the xenon dimer at its minimum, $E_0(R_{\mathrm{min}})$, with imaginary time step and walker number is presented for the on-site-plus-pair-coulomb trial function under NC-DMC (fig. 7.2). The observed behavior is in accord with expectations (e.g. the convergence with N and τ is uniform and has the appropriate power law dependence given the method of [UNR93] has not been implemented.).



(a) Convergence of E_0 vs imaginary-time step for the full QDO model xenon dimer computed using the NC-DMC method with the on-site-plus-pair-coulomb trial function.



(b) Convergence of E_0 vs walker number computed using a time step of $\hbar\tau = 0.02$au. Note, the energy scale is on order 10μHartree, stringently testing the method.

Figure 7.2: Convergence tests for DMC, Xenon dimer

7.2.2 Xenon FCC-crystal, in the dipole-limit, at T=0

Having demonstrated the stability and accuracy of the new techniques on small systems, it is natural to examine larger systems for which high quality "exact" results can still be obtained. The dipole limit QDO model for the perfect fcc-xenon solid can be solved quasi-analytically in reciprocal space (a $6n_{\text{site}} \times 6n_{\text{site}}$ matrix constructed by appropriate G-vector sums is diagonalized at each of the n_{cell}^3 k-points in Brillouin zone where $n_{\text{site}} = 4$ for the fcc lattice and n_{cell} is the number of fcc unit cells in the crystal of interest, as opposed to diagonalizing a single $3N_{\text{Xe}} \times 3N_{\text{Xe}}$ matrix). The results of a NC-DMC study of the dipole limit QDO model of the perfect fcc-xenon solid at the experimental lattice constant performed using $\hbar\tau = 0.01$ $N = 1000$ and the on-site-plus-dipole-dipole trial function are compared to the analytical results as a function of system size, N_{Xe}, in fig. 7.3. Although the NC-DMC imaginary time step must be decreased as $\sqrt{N_{\text{Xe}}}$ with increasing system size, the NC-DMC simulation estimates of the ground state energies match the analytical answers within the error bars.

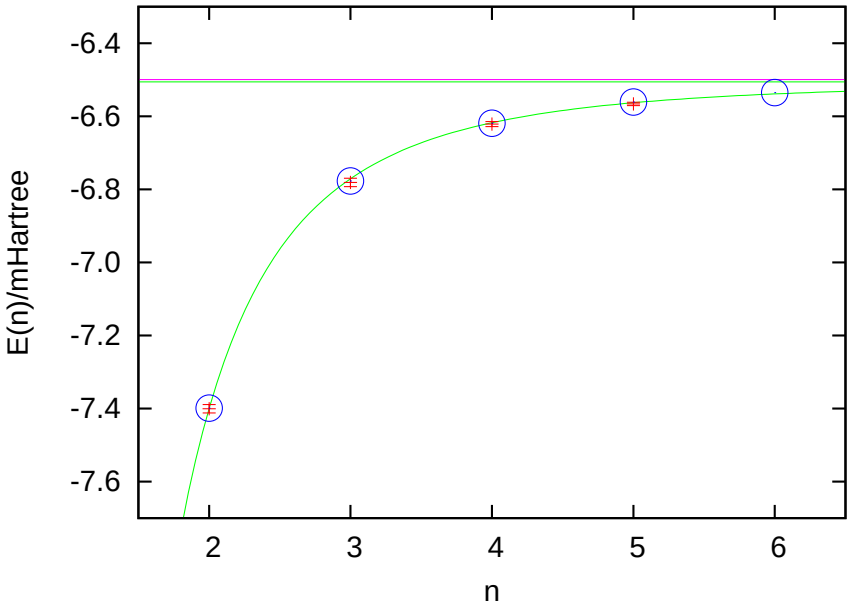


Figure 7.3: Convergence of E_0 of the dipole-limit of QDO model xenon fcc-solid at the equilibrium lattice constant vs system size ($n = n_{\text{cell}}$), computed using NC-DMC with the on-site-plus-dipole-dipole trial function, $N = 1000$ walkers and a time step of $\hbar\tau = 0.01$. Simulation results (red) are compared to the analytical results (blue). A fit to the data is included: $f(n) = E_0 - \Delta/n^3$, along with the extrapolated value of $E_0 = -6.5055$mHartree/atom (green), which is close to the large-N analytic result of $E_0 = -6.4996$mHartree/atom, calculated for $N = 32000$ (purple).

7.2.3 Xenon FCC-crystal, full QDO model, at T=0

The results of the previous subsection suggest that the new techniques are capable of generating high quality results also for the non-trivial full QDO model of the fcc-xenon crystal (where there are no analytical answers with which to compare).

Proceeding carefully, therefore, in fig. 7.4, the convergence of the ground state energy with imaginary time step for the full QDO model N_{Xe} =32 atom xenon solid computed using the NC-DMC method with the on-site-plus-pair-coulomb trial function and $N = 1000$ walkers is given. Further computations on the QDO model solid given below were performed using $\hbar\tau = 0.02$au. and $N = 1000$ and an appropriate reduction of τ as $\sqrt{N_{\text{Xe}}}$ with increasing system size.

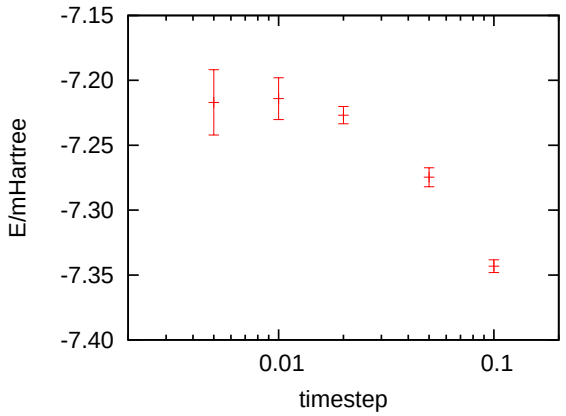


Figure 7.4: Convergence of the E_0 of the full xenon QDO model fcc-solid at the equilibrium lattice constant, vs imaginary time step, computed using NC-DMC with the on-site-plus-pair-coulomb trial function, with 32 atoms and $N = 1000$ walkers.

Fig. 7.5 shows the ground state energy of the fcc-xenon solid as a function of system size, N_{Xe} at the experimentally determined lattice constant, $a_{eq} = 6.12$Å under the full model QDO, along with the extrapolation of the results to $N_{\mathrm{Xe}} \to \infty$ limit. The experimental $T = 0$ binding energy is $E_0(a_{eq}) = -6.05$mHartree/atom[BWL$^+$74] while the present full QDO predicts $E_0(a_{eq}) = -6.27$mHartree/atom. The zero point energy is estimated to be 0.2mHartree/atom[BB83] and hence the agreement is good. It is important to note that the full model QDO is fit to reproduce the BWLSL pair potential in the gas phase. The high accuracy gas phase pair potential predicts a $T = 0$ crystal binding energy of $E_0(a_{eq}) = -6.81$mHartree/atom. Thus, the full QDO model introduces substantial many-body corrections in the condensed phase.

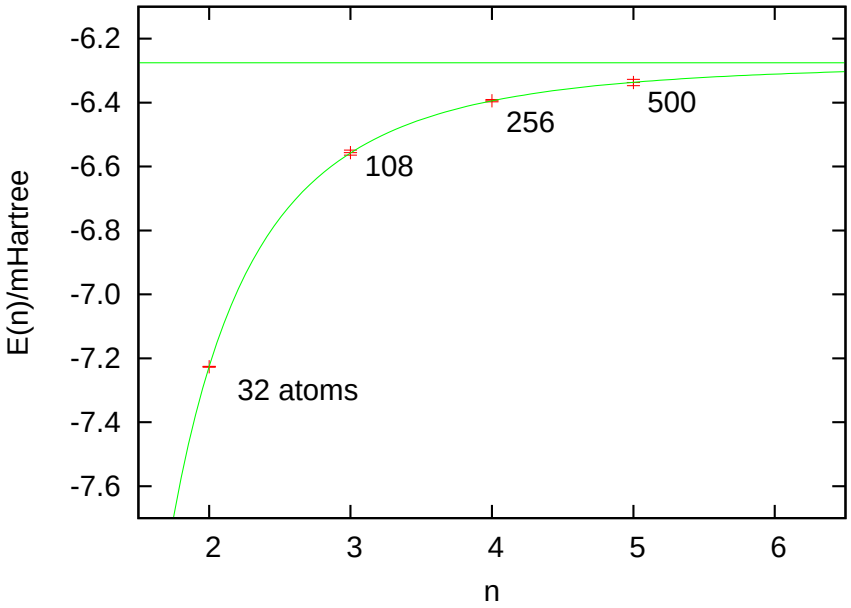
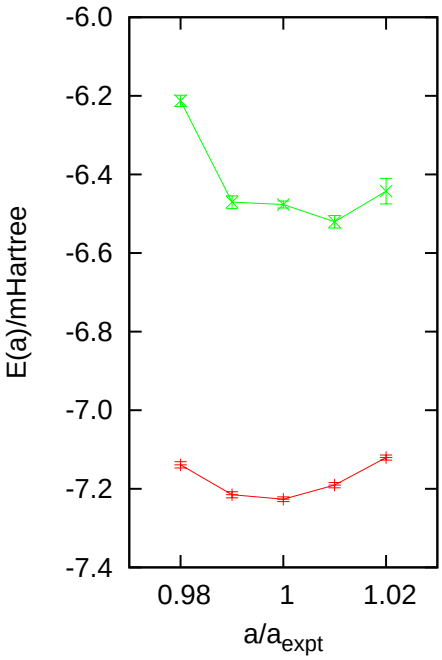


Figure 7.5: Convergence of the E_0 of the full xenon QDO model fcc-solid at the equilibrium lattice constant vs system size ($n = n_{cell}$), computed using NC-DMC with the on-site-plus-pair-coulomb trial function (red), $N = 1000$ walkers and an imaginary time step of $\hbar\tau = 0.02$au. A fit to the data is included $f(n) = E_0 - \Delta/n^3$ (green curve), along with the extrapolated value of $E_0 = -6.275$mHartree/atom (green horizontal).

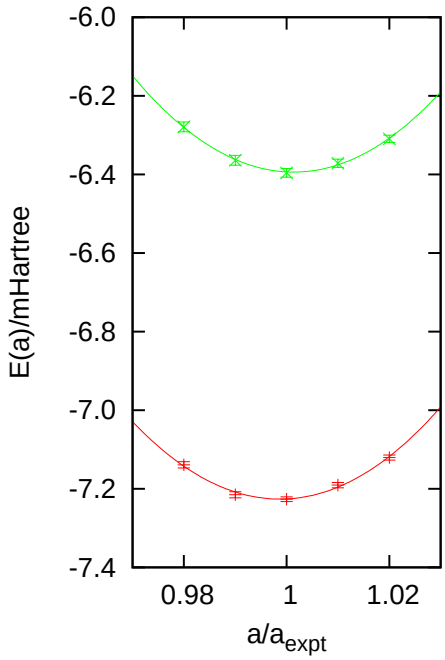
In fig. 7.6(a), the ground state energy of the N_{Xe}=32 atom fcc-xenon solid as a function of lattice constant, $E_0(a)$, is presented along with the variational energy, $E_{\mathrm{T}}(a)$, of the on-site-plus-pair-coulomb trial function. It is clear that $E_{\mathrm{T}}(a)$ is not accurate enough to describe the solid well.

Finally, fig. 7.6(b) shows the ground state energy of the N_{Xe}=32 and N_{Xe}=256 atom fcc-xenon solids as a function of lattice constant is presented . The lattice constant predicted by the QDO model is in very good agreement with experiment ($a_{eq} = 6.12$Å). In contrast, the BWLSL gas-phase pair potential, which is reproduced by the QDO model in the two-body, gas-phase limit, predicts $a_{eq} = 6.04$Å. Nuclear quantum effects (see 7.2.3) are estimated to increase the lattice constant by 0.03Å. The QDO prediction of the bulk modulus for two system sizes, $\kappa = 4.0 \pm 0.1$GPa for N_{Xe}=32 atoms and $\kappa = 4.2 \pm 0.2$GPa for N_{Xe}=256 atoms are in good agreement with each other and nuclear quantum effects are estimated to decrease these values by 0.3GPa. Thus, the model prediction is within 3 to 10 percent of the experimental value, $\kappa = 3.64$GPa. Note, the BWLSL gas phase pair potential predicts a rather large modulus, $\kappa = 4.55$GPa. Thus, the QDO does capture the many-body terms that arise in condensed phase

reasonably well. These new accurate simulation results show the present QDO xenon model somewhat "stiff" in the solid phase although the lattice constant and the binding energy are predicted reasonably. In general, the QDO parameters need to be tuned to reflect the higher quality solid state simulation data that can now be generated.



(a) E_0 vs lattice constant, of the N_{Xe}=32 atom fcc-xenon solid, under the full QDO model, computed using NC-DMC with the on-site-plus-pair-coulomb trial function (bottom curve), $N = 1000$ walkers and an imaginary time step of $\hbar\tau = 0.02$au, compared to the variational result of the on-site-plus-pair-coulomb trial function (top curve).



(b) E_0 vs lattice constant of fcc-xenon for N_{Xe}=32 atom (bottom curve) and N_{Xe}=256 atom (top curve) cells, under the full QDO model, computed using NC-DMC with the on-site-plus-pair-coulomb trial function, $N = 1000$ walkers and an imaginary time step of $\hbar\tau = 0.02$au.

Figure 7.6: Xenon FCC-crystal bulk modulus curves

Nuclear Quantum Effects

The relative importance of quantum effects due to the nuclear motion of Lennard Jones (LJ) atoms can be quantified with the de Boer parameter, $\Lambda = h/\sigma\sqrt{m\epsilon}$, where m is the mass of the LJ atom, and σ and ϵ are its LJ interaction radius and energy, respectively[dB48]. Properties of Lennard Jonesium in reduced units depend linearly on Λ, for example its reduced density can be fitted to the equation $\rho^* = -1.0789 - 0.845 \cdot \Lambda$, as ascertained from Table II in Ref. [Cha02]. Using results for solid argon on the quantum effects in the lattice constant and bulk modulus from the literature [SM01], we can use this linear dependence on Λ to estimate the quantum effects in xenon. The LJ parameters for argon used in Ref.[SM01] were $\sigma = 3.405$ Å, and $\epsilon = 120\ k_B K$, while reasonable values for the condensed phase of xenon are $\sigma = 4.055$ Å, and $\epsilon = 228\ k_B K$, thus the ratio $\Lambda_{\mathrm{Xe}}/\Lambda_{\mathrm{Ar}} = 0.42$. The relative difference in lattice constant and elastic constants of quantum versus classical for argon were reported to be 1.2% and -17%, respectively. Given the ratio for Λ stated above, a 0.5% increase in the classical limit lattice constant of xenon and a 7% decrease in the classical limit bulk modulus of xenon due to quantum effects can be expected.

Chapter 8

Path Integral Molecular Dynamics

We use Path Integral Molecular Dynamics [PR84, TBJK93] as it allows us to calculate accurate nuclear forces, is order $N \log N$ (N is the number of nuclei), and can be systematically improved by increasing discretization. In this way, it provides us with a powerful tool for applying QDOs to realistic systems.

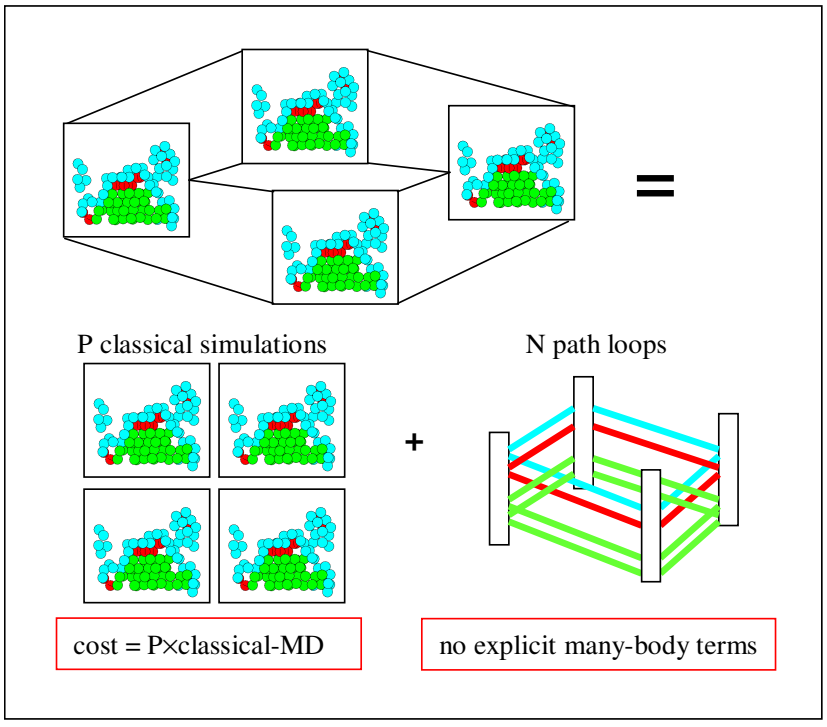


Figure 8.1: Pictorial summary of Path Integral Molecular Dynamics

We outline the theory of PIMD, starting with the density operator, density matrix,

and Feynman's path integral formalism, describing some density matrices and their properties along the way, especially the concept of a 'high-temperature' approximate density matrix. Then we make the conceptual leap to sampling the partition function by PIMD, using Newton's equations of motion on a fictitious classical potential to sample the microcanonical ensemble, then adding thermostats to properly sample the canonical ensemble. We also outline the techniques required to efficiently sample the paths, including bead-staging, and low-variance estimators for the energy, that come from integrating out the large on-site energy from the harmonic oscillator. We then extend PIMD to the motion of nuclei on the implied Born-Oppenheimer surface. We describe the efficiency gains from increasing the Drude temperature, how to properly calculate forces and pressures, and how to increase efficiency by increasing the faux-mass of the Quantum Drude, while maintaining adiabatic separation between the quantum and classical coordinates, (which each have a different temperature).

8.1 From the Boltzmann Distribution to the Density Matrix

The operator $\hat{\rho} = e^{-\beta\hat{H}}$ [Fey72], which is the quantum operator equivalent of the classical Boltzmann distribution, is known as the density operator. It is diagonal in the basis of the $\hat{H}$-eigenstates, and it will tend to increase the weight of the lowest-energy states in any given superposition.

When applied between two position states, $\hat{\rho}$ becomes a non-diagonal density matrix, $\rho(x, x'; \beta)$. When the temperature is low, there is more structure due to the potential, and correlation distances are long. When the temperature is high, there is less structure (it behaves more like a free particle), and correlation lengths fall, but we can integrate over chains of approximate (simplified) high-temperature density matrices to build accurate low-temperature ones; a path integration.

The diagonal-elements of the density matrix, $\rho(x, x; \beta)$, a kind of position-state 'self-correlation' function, turn out to be the temperature-dependent probability density function, which becomes the classical Boltzmann distribution, in the classical limit, when $\beta = 1/kT \to 0$ or $\hbar \to 0$.

$$\rho(x, x; \beta) \to e^{-\beta V(x)} \qquad \text{(classical limit)}.$$

8.1.1 Density Operator

The quantum Boltzmann distribution leads to the definition of the partition function Z, in terms of the density operator.

$$\begin{aligned} \hat{\rho} &\equiv e^{-\beta \hat{H}} = \sum_i |i\rangle e^{-\beta E_i} \langle i|, \\ Z &= \sum_i e^{-\beta E_i} \langle i|i\rangle = \mathrm{Tr}\left(\hat{\rho}\right). \end{aligned}$$

and from Z when can compute the thermodynamic quantities,

$$\begin{aligned} F &= -kT \log Z \ , \\ E &= -\tfrac{\partial}{\partial \beta} \log Z \ , \\ P &= -\tfrac{\partial}{\partial V} F_{NVT} = kT \tfrac{\partial}{\partial V} \log Z_{NVT} \ . \end{aligned}$$

It is also possible to calculate observables using the density operator.

$$\begin{aligned} \langle A \rangle &= \sum A_i \rho_i = \sum_i \langle i|\hat{A}|i\rangle \langle i|\hat{\rho}|i\rangle \\ &= \sum_{ii'} \langle i|\hat{A}|i'\rangle \langle i'|\hat{\rho}|i\rangle \ , \qquad \because \langle i'|\hat{\rho}|i\rangle = \langle i|\hat{\rho}|i\rangle \delta_{ii'} \ , \\ &= \sum_i \langle i|\hat{A}\hat{\rho}|i\rangle = \mathrm{Tr}\left(\hat{A}\hat{\rho}\right) \ . \end{aligned}$$

The density operator is a mathematical operator that can be expressed in different ways or in different bases. $\hat{\rho}$ is an energy eigenoperator; that is, the energy basis is its diagonal basis, but it can be applied in different bases.

$$\begin{aligned} \rho(x, x'; \beta) &= \langle x|\hat{\rho}(\beta)|x'\rangle, \qquad \text{the position-state density matrix,} \\ \rho(p, p'; \beta) &= \langle p|\hat{\rho}(\beta)|p'\rangle, \qquad \text{the momentum-state density matrix.} \end{aligned}$$

The traces can also be calculated in different bases. It is possible to split the operator by factorising the exponential, allowing us to use off-diagonal elements of the relevant

density matrix.

$$\begin{aligned}
\mathrm{Tr}\left(\hat{\rho}(\beta)\right) &= \int_x \langle x|\hat{\rho}(\beta)|x\rangle = \int_x \rho(x,x;\beta) \\
&= \int_x \langle x|\hat{\rho}(\beta-\tau)\hat{\rho}(\tau)|x\rangle \\
&= \int_x \int_{x'} \langle x|\hat{\rho}(\beta-\tau)|x'\rangle\langle x'|\hat{\rho}(\tau)|x\rangle \\
&= \int_x \int_{x'} \rho(x,x';\beta-\tau)\rho(x',x;\tau) \qquad \text{etc ...}
\end{aligned}$$

It is also possible to calculate expectation values for operators. If $\hat{A}$ is a position-operator, then it can be expressed as a function $A(x)$.

$$\begin{aligned}
\mathrm{Tr}\left(\hat{A}\hat{\rho}(\beta)\right) &= \int_x \langle x|\hat{A}\hat{\rho}(\beta)|x\rangle = \int_x A(x)\rho(x,x;\beta) \\
&= \int_x \int_{x'} \langle x|\hat{A}\hat{\rho}(\beta-\tau)|x'\rangle\langle x'|\hat{\rho}(\tau)|x\rangle \\
&= \int_x \int_{x'} A(x)\rho(x,x';\beta-\tau)\rho(x',x;\tau) \qquad \text{etc ...}
\end{aligned}$$

For an operator $\hat{B}$ that is not a position operator, then it could be expressed as a non-diagonal matrix $B(x,x')$ and the diagonal terms of $\rho(x,x';\beta)$ could be used to complete the trace:

$$\begin{aligned}
\mathrm{Tr}\left(\hat{B}\hat{\rho}(\beta)\right) &= \int_x \int_{x'} \langle x|\hat{A}|x'\rangle\langle x'|\hat{\rho}(\beta)|x\rangle \\
&= \int_x \int_{x'} B(x,x')\rho(x',x;\beta) \ .
\end{aligned}$$

8.1.2 Some Example Density Matrices

Free Particle Density Matrix

For an example, consider a free particle

$$\hat{\rho}(\beta) = e^{-\beta\hat{H}_{\mathrm{free}}} = e^{-\beta\frac{\hat{p}^2}{2m}} = e^{+\beta\frac{\hbar^2}{2m}\nabla^2} \ .$$

So the free particle density matrix is (See [Fey72, pg. 48]) :

$$\begin{aligned}\rho(x,x';\beta) &= \langle x|e^{-\frac{\beta\hat{p}^2}{2m}}|x'\rangle \\ &= \int_p \langle x|e^{-\frac{\beta\hat{p}^2}{2m}}|p\rangle\langle p|x'\rangle \\ &= \sqrt{\frac{m}{2\pi\beta\hbar^2}} \times e^{-\frac{m}{2\beta\hbar^2}(x'-x)^2} .\end{aligned}$$

Simple Harmonic Density Matrix

Also consider the harmonic oscillator

$$\hat{\rho}(\beta) = e^{-\beta\left(\frac{\hat{p}^2}{2m}+\frac{m}{2}\omega^2x^2\right)} .$$

The corresponding density matrix is (See [Fey72, pg. 48]).

$$\begin{aligned}\rho(x,x';\beta) &= \sqrt{\frac{m\omega}{2\pi\hbar\sinh f}} \times \exp\left\{\frac{-m\omega}{2\hbar}\left[\frac{(x-x')^2}{\sinh f}+\tanh\tfrac{f}{2}(x^2+x'^2)\right]\right\}, \\ &\text{where} \qquad f=\beta\hbar\omega.\end{aligned}$$

Structure at Low Temperature (of Harmonic Oscillator)

At low temperature (β, f are large), we get

$$\rho(x,x';\beta) \rightarrow \sqrt{\frac{m\omega}{\pi\hbar e^f}}\exp\left\{-\tfrac{m\omega}{2\hbar}(x^2+x'^2)\right\} .$$

Except for the normalisation, this looks like a product of the groundstate wavefunction at two points $\phi_0(x)\phi_0(x')$.
In fact that is always true as $f, \beta \to \infty$.

We can show this by expanding the density operator in terms of energy basis states:

$$\begin{aligned}\rho(x,x';\beta) &= \langle x|e^{-\beta H}|x'\rangle \\ &= \sum_i \langle x|\phi_i\rangle e^{-\beta E_i}\langle\phi_i|x'\rangle, \\ &\to \langle x|\phi_0\rangle e^{-\beta E_0}\langle\phi_0|x'\rangle \\ &= e^{-\beta E_0}\phi_0(x)\phi_0(x').\end{aligned}$$

Simplicity at High Temperature

At high temperature (β, f are small), the density matrix shrinks to a delta function:

$$\begin{aligned}\rho(x,x';\beta) &\rightarrow \sqrt{\tfrac{m\omega}{2\pi\hbar f}}\exp\left\{\tfrac{m\omega}{2\hbar f}(x-x')^2\right\} \\ &\rightarrow \delta(x-x') \qquad \text{as} \quad f\rightarrow 0\end{aligned}$$

Notice that as temperature increases, the density matrix behaves more like that of a free-particle. As one might expect, as temperature increases, the structure of the groundstate wavefunction is lost, and correlation times (actually imaginary time $\hbar\tau$) and correlation lengths become shorter, ending up as a delta function.

8.2 Approximate Density Matrices

8.2.1 Operator Splitting

We need to be able to do create density matrices that are guaranteed to be correct to a given order in β. In many cases we do not have the exact expression that we need. For example, the exponential of two operators that do not commute, we can do [Tro59, Suz76, Sch81]

$$\begin{aligned}e^{\tau(\hat{A}+\hat{B})} &= 1+\tau\left(\hat{A}+\hat{B}\right)+\tfrac{\tau^2}{2}\left(\hat{A}^2+\hat{A}\hat{B}+\hat{B}\hat{A}+\hat{B}^2\right)+\mathcal{O}\left(\tau^3\right), \\ e^{\tau\hat{A}}e^{\tau\hat{B}} &= \left[1+\tau\hat{A}+\tfrac{\tau^2}{2}\hat{A}^2\right]\left[1+\tau\hat{B}+\tfrac{\tau^2}{2}\hat{B}^2\right]+\mathcal{O}\left(\tau^3\right) \\ &= 1+\tau\left(\hat{A}+\hat{B}\right)+\tfrac{\tau^2}{2}\left(\hat{A}^2+\hat{B}^2+2\hat{A}\hat{B}\right)+\mathcal{O}\left(\tau^3\right) \\ &= e^{\tau(\hat{A}+\hat{B})}+\tfrac{\tau^2}{2}\left[\hat{A},\hat{B}\right]+\mathcal{O}\left(\tau^3\right), \\ &\quad \text{where} \quad \left[\hat{A},\hat{B}\right]\equiv\hat{A}\hat{B}-\hat{B}\hat{A} \qquad \text{is the commutator.}\end{aligned}$$

Fortunately, this can easily be improved by symmetrising the splitting:

$$\begin{aligned}e^{\frac{\tau}{2}\hat{A}}e^{\frac{\tau}{2}\hat{B}}e^{\frac{\tau}{2}\hat{B}}e^{\frac{\tau}{2}\hat{A}} &= e^{\tau(\hat{A}+\hat{B})}+\tfrac{\tau^2}{2}\left\{\tfrac{1}{4}\cancel{\left[\hat{A},\hat{B}\right]}+\tfrac{1}{4}\cancel{\left[\hat{B},\hat{A}\right]}\right\}+\mathcal{O}\left(\tau^3\right), \\ e^{\frac{\tau}{2}\hat{A}}e^{\tau\hat{B}}e^{\frac{\tau}{2}\hat{A}} &= e^{\tau(\hat{A}+\hat{B})}+\mathcal{O}\left(\tau^3\right).\end{aligned}$$

We call this a 2nd-order factorisation because when we consider P applications, the error becomes $\mathcal{O}\left(P\tau^3\right)=\mathcal{O}\left(\beta\tau^2\right)=\mathcal{O}\left(\tau^2\right)$ where $\beta{=}P\tau{=}1/kT$. This decomposition was originally due to Trotter and Suzuki [Tro59, Suz76, Sch81]. It is mentioned by Martyna and Whitfield[WM07], along with a 4th-order factorisation, which uses another scheme by Suzuki[Suz86, Suz94, Chi97].

The Hamiltonian usually contains a 'kinetic' term containing momentum operators and a 'potential' term which is a position operator. Momentum and position operators do not commute, making it is non-trivial to find a solution to the density matrix. Instead, we build approximations.

The simplest thing to do is to simply separate the kinetic term $\hat{T}$ from the potential term $\hat{V}$. Because the density matrix for the harmonic oscillator is known, an alternative is to split the harmonic oscillator hamiltonian ($\hat{H}_{\text{h.o.}} = \hat{T} + \hat{V}_{\text{h.o.}}$) from the external potential (V_{ext}). Other (new) density matrices are explored later on, towards the end of this chapter.

8.2.2 Free Particle Reference

We have two operators $\hat{T} = \frac{\hat{p}^2}{2m} = -\frac{\hbar^2}{2m}\nabla^2$, and $\hat{V} = V(x)$. Because $\hat{V}$ is a position operator, it makes sense to put it on the outside of the symmetric splitting. The operator $\hat{T}$ generates the free-particle reference propagator, as derived by Feynman [Fey72].

$$
\begin{aligned}
\hat{\rho} &= e^{-\tau\hat{H}}, \\
\hat{\rho} &\rightarrow e^{-\frac{\tau}{2}\hat{V}} e^{-\tau\hat{T}} e^{-\frac{\tau}{2}\hat{V}}, \\
\rho(x,x';\tau) &= \langle x|e^{-\tau\hat{H}}|x'\rangle, \\
\rho(x,x';\tau) &\rightarrow \langle x|e^{-\frac{\tau}{2}\hat{V}} e^{-\tau\hat{T}} e^{-\frac{\tau}{2}\hat{V}}|x'\rangle \\
&= e^{-\frac{\tau}{2}V(x)}\langle x|e^{-\tau\hat{T}}|x'\rangle e^{-\frac{\tau}{2}V(x')} \\
&= e^{-\frac{\tau}{2}V(x)}\ \rho_{\text{free}}(x,x';\tau)\ e^{-\frac{\tau}{2}V(x')} \\
&= e^{-\frac{\tau}{2}V(x)}\ e^{-\frac{m}{2\hbar^2\tau}(x-x')^2}\ e^{-\frac{\tau}{2}V(x')}, \\
&\quad \text{which is correct to } \mathcal{O}\left(\tau^3\right).
\end{aligned}
$$

8.2.3 Harmonic Oscillator Reference

We have two operators $\hat{H}_{\text{h.o.}} = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2x^2$, and $\hat{V}_{\text{ext}} = V_{\text{ext}}(x)$, where $\hat{H}_{\text{h.o.}}$ generates the harmonic-oscillator reference propagator [Fey72].

$$
\begin{aligned}
\hat{\rho} &\rightarrow e^{-\frac{\tau}{2}\hat{V}_{\text{ext}}} e^{-\tau\hat{H}_{\text{h.o.}}} e^{-\frac{\tau}{2}\hat{V}_{\text{ext}}}, \\
\rho(x,x';\tau) &\rightarrow \langle x|e^{-\frac{\tau}{2}\hat{V}_{\text{ext}}} e^{-\tau\hat{H}_{\text{h.o.}}} e^{-\frac{\tau}{2}\hat{V}_{\text{ext}}}|x'\rangle \\
&= e^{-\frac{\tau}{2}V_{\text{ext}}(x)}\langle x|e^{-\tau\hat{H}_{\text{h.o.}}}|x'\rangle e^{-\frac{\tau}{2}V_{\text{ext}}(x')} \\
&= e^{-\frac{\tau}{2}V_{\text{ext}}(x)}\ \rho_{\text{h.o.}}(x,x';\tau)\ e^{-\frac{\tau}{2}V_{\text{ext}}(x')}, \qquad \text{also correct to } \mathcal{O}\left(\tau^3\right),
\end{aligned}
$$

where

$$\begin{aligned}
\rho_{\text{h.o.}}(x,x';\tau) &= \sqrt{\tfrac{m\omega}{2\pi\hbar\sinh f}} \times \exp\left\{ \frac{-m\omega}{2\hbar}\left[\frac{1}{\sinh f}(x-x')^2 + \tanh\tfrac{f}{2}(x^2+x'^2)\right]\right\} \\
&= \exp\left\{ \tfrac{1}{2}\mathcal{N} - \tfrac{1}{2}\mathcal{A}(x-x')^2 - \tfrac{1}{2}\mathcal{B}(x^2+x'^2)\right\}, \\
&\quad \text{where} \quad f = \tau\hbar\omega\,, \\
&\qquad\qquad \mathcal{N} = \log\left[\tfrac{m\omega}{2\pi\hbar\sinh f}\right]\,, \\
&\qquad\qquad \mathcal{A} = \tfrac{m\omega}{\hbar}\tfrac{1}{\sinh f}\,, \\
&\qquad\qquad \mathcal{B} = \tfrac{m\omega}{\hbar}\tanh\tfrac{f}{2}\,.
\end{aligned}$$

8.3 From Density Matrices to Discretised Path Integrals

8.3.1 Density Matrices as an Integral of Paths

Approximate density matrices are most accurate at small-τ or -β (high-temperature: where short τ means there is only short-range spatial correlation; little structure to be lost). We can use this fact to split a large-β density matrix (low-temperature: where long τ allows long-range spatial correlation; and more structure) into an integral over a product of multiple small-τ density matrices; lots of little timesteps [Fey72]. Thus when we refer to 'high-temperature' density matrices, we are referring to simpler, approximate ones, small-τ from which we build the desired 'low-temperature' density matrices.

We do this by splitting/factorising the density operator ρ into several (P of them), and inserting complete sets of states between them. Let $\beta = P\tau$.

$$\begin{aligned}
\rho(x, x'; \beta) &= \langle x|\hat{\rho}(\beta)|x'\rangle \\
&= \langle x|e^{-\beta\hat{H}}|x'\rangle \\
&= \langle x|\left[e^{-\tau\hat{H}}\right]^P|x'\rangle \\
&= \int \mathrm{d}x_1\mathrm{d}x_2...\mathrm{d}x_{\mathrm{P}-1}\langle x|e^{-\tau\hat{H}}|x_1\rangle\langle x_1|e^{-\tau\hat{H}}|x_2\rangle\langle x_2|\cdots e^{-\tau\hat{H}}|x'\rangle \\
&= \int \mathrm{d}x_1\mathrm{d}x_2...\mathrm{d}x_{\mathrm{P}-1}\,\rho(x, x_1; \tau)\rho(x_1, x_2; \tau)\cdots\rho(x_{\mathrm{P}-1}, x'; \tau).
\end{aligned}$$

The high-T density matrices can be thought of as path-segments, and to construct the low-T density matrix requires an integral over every intermediate coordinate, thus integrating over all possible paths made of high-T path-segments. Hence the name "path integral".

We want to know the density of each state, which means taking the trace of the density matrix, and this in turn implies a closed path that finishes where it starts. Each coordinate on the chain is equivalent as far as the integration is concerned, so we can use integrate using measurements all round the chain.

8.3.2 The Partition function and Measurables

The partition function is obtained by closing the path and integrating over that start point, becoming an integral over all closed paths through state-space [Fey72].

$$\begin{aligned} Z(\beta) &= \mathrm{Tr}\,(\hat{\rho}) = \int \mathrm{d}x \rho(x,x;\beta) \\ &= \int \mathrm{d}x_0\,\mathrm{d}x_1 \cdots \mathrm{d}x_{\mathrm{P}-1}\, \rho(x_0,x_1;\tau)\rho(x_1,x_1;\tau)\cdots\rho(x_{\mathrm{P}-1},x_0;\tau) \\ &= \int \mathrm{d}^P x \prod_{i=0}^{\mathrm{P}-1} \rho(x_i,x_{i+1};\tau)\ , \qquad \text{where} \quad x_{i\pm\mathrm{P}} \equiv x_i\ . \end{aligned}$$

And similarly for measurables (of position operators) [Fey72]:

$$\begin{aligned} \langle A \rangle &= \mathrm{Tr}\left(\hat{A}\hat{\rho}\right) = \int_x A(x)\,\rho(x,x;\beta) \\ &= \int \mathrm{d}^P x\, A(x_0) \prod_{i=0}^{\mathrm{P}-1} \rho(x_i,x_{i+1};\tau)\ . \end{aligned}$$

The closed polymer is symmetric under rotation of the indices, meaning that the measurable $A(x)$ can be calculated at any of the indices, and therefore can be averaged over all of them [WM07].

$$\langle A \rangle = \int \mathrm{d}^P x \left[\frac{1}{P}\sum_{i=0}^{\mathrm{P}-1} A(x_i)\right]\left[\prod_{i=0}^{\mathrm{P}-1} \rho(x_i,x_{i+1};\tau)\right]\ .$$

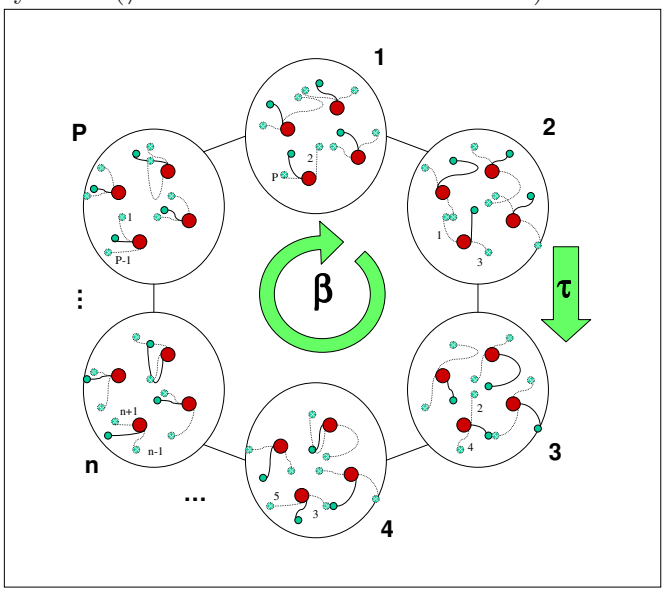
8.3.3 Integration over all Closed Quantum Paths

We are left with something that looks like an integration over all possible states of a cyclic-polymer, with each node being a state in $3N$-space. This is only true for distinguishable particles, including the Quantum Drude, but is not complete for Bosons or Fermions where exchange interactions occur see pg. 142.

$$
\begin{aligned}
x &\rightarrow \mathbf{x} = (x_0, x_1 ..., x_{\mathrm{P}-1}) \ , \\
\rho(x) &\rightarrow \rho(\mathbf{x}) = \prod_{i=0}^{\mathrm{P}-1} \rho(x_i, x_{i+1}; \tau) \ , \\
\int \mathrm{d}x &\rightarrow \int \mathrm{d}^P x \quad , \qquad \text{or equivalently} \qquad \int_x \rightarrow \int_{\mathbf{x}} \quad .
\end{aligned}
$$

The same way as in classical mechanics, each state has a certain weight, in quantum mechanics for distinguishable particles, each possible configuration for the cyclic polymer has a certain weight. The distribution can be sampled by various methods, including Monte Carlo methods (PIMC), or Molecular Dynamics (PIMD).

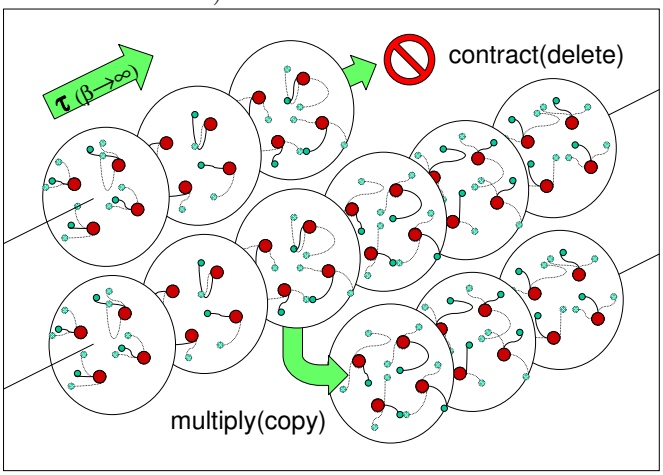
Figure 8.2: Path Integrals: Cyclic polymer has rotational symmetry through 'imaginary time' ($\neq$ to simulation time-coordinate)



8.3.4 DMC as an Integral over Open-Ended Paths

The mechanics and mathematics of DMC are very different from path integrals and PIMD, but there is this homology between them.

Figure 8.3: DMC: Infinite polymer proceeds through 'imaginary time'
(= to simulation time-coordinate)



DMC and Density Matrices

The propagation steps of DMC are like density matrices [PC84], except that they contain some extra weighting due to the trial wavefunction. But because the chain does not close, we do not get the trace, and so we do not know the ground-state density $\rho(x, x; \beta \to \infty, T \to 0) = |\Psi_0(x)|^2$. What we do get is the ground-state wavefunction (weighted by the trial wavefunction). If a DMC simulation could continue indefinitely, the paths would fill all space with the density f.

$$P(x' \to x, \tau) \quad \propto \quad \rho(x, x'; \tau) \times \frac{\Psi_{\mathrm{T}}(x)}{\Psi_{\mathrm{T}}(x')},$$

$$
\begin{aligned}
f(x;0) &= {\Psi_{\rm T}}^2(x) \qquad \text{(usual initialisation)},\\[2ex]
f(x;\tau) &= \int {\rm d}x_0\; P(x_0 \to x;\tau) f(x_0;0),\\
f(x;P\tau) &= \int {\rm d}x_{\rm P-1} \cdots \int {\rm d}x_1 \int {\rm d}x_0\; P(x_{\rm P-1} \to x) \cdots P(x_1 \to x_2;\tau)\\
&\qquad P(x_0 \to x_1;\tau)\; f(x_0;0)\\
&= \int {\rm d}x_{\rm P-1} \cdots {\rm d}x_0\; \Psi_{\rm T}(x)\, \rho(x, x_{\rm P-1};\tau)\, \cancel{\Psi_{\rm T}^{-1}(x_{\rm P-1})}\\
&\qquad\qquad \cancel{\Psi_{\rm T}(x_{\rm P-1})} \cdots \cancel{\Psi_{\rm T}^{-1}(x_1)}\\
&\qquad\qquad \cancel{\Psi_{\rm T}(x_1)}\, \rho(x_1, x_0;\tau)\, \cancel{\Psi_{\rm T}^{-1}(x_0)} {\Psi_{\rm T}}^2(x_0),\\
f(x;P\tau) &= \Psi_{\rm T}(x) \int {\rm d}x_0\; \rho(x, x_0; P\tau)\; \Psi_{\rm T}(x_0),\\[2ex]
f(x;P\tau) &\to c_0 \Psi_{\rm T}(x) \Psi_0(x), \qquad \text{as} \quad \beta = P\tau \to \infty.
\end{aligned}
$$

8.3.5 Path Integrals for Many-particle Statistics

The Quantum Drude Oscillator is a distinguishable particle, as each particle is attached to a unique centre, or, if on the same centre, have unique parameters. Real particles, however, fall into one of two categories; Bosons or Fermions; and each has a different statistical behaviour that comes from the symmetry of their wavefunctions; Boson many-particle wavefunctions are symmetric under permutation of indices, while Fermion many-particle wavefunctions are antisymmetric.

Identical or indistinguishable particles satisfy the following property: the Hamiltonian is invariant to particle exchange. That is, if one labels particle A and particle B as separate entities, and the Hamiltonian is not changed by swapping A and B, performing $\hat{H}$, then swapping them back again. We write this down as $\hat{p}\hat{H}\hat{p} \equiv \hat{H}$ for identical particles, where $\hat{p}$ is the permutation operator. $\hat{p}$ has the property that it is its own inverse; $\hat{p}^2 = 1$.

$$\begin{aligned} \hat{p}\hat{H}\hat{p}\Psi &= \hat{H}|\Psi\rangle = E|\Psi\rangle, \\ \hat{p}^2\hat{H}\hat{p}\Psi &= \hat{p}E|\Psi\rangle, \qquad \text{but } E \text{ is just a constant,} \\ \hat{H}\hat{p}\Psi &= E\hat{p}|\Psi\rangle. \end{aligned}$$

This means that $\hat{p}|\Psi\rangle$ is an eigenstate of $\hat{H}$. If we make one further reasonable assumption, that $\hat{p}$, being a book-keeping/labelling operation with no physical consequences, does not change the state (the permutation of two non-identical particles would be a physical consequence), then it at most adds some constant factor; that means $|\Psi\rangle$ must be an eigenstate of $\hat{p}$. This makes something very weird happen; the wavefunction must be either symmetric or antisymmetric under permutation.

$$\begin{aligned} \hat{p}^2 &= 1, \\ \hat{p}^2|\Psi\rangle &= |\Psi\rangle, \\ \hat{p}|\Psi\rangle &= (\pm 1)\,|\Psi\rangle. \end{aligned}$$

Relativity is grafted onto standard quantum mechanics by allowing Fermions to occupy only antisymmetric eigenstates of $\hat{H}$, and Bosons to occupy only symmetric eigenstates of $\hat{H}$. Wavefunctions of the correct symmetry can be selected using a symmetrisation

/ antisymmetrisation as follows (if we neglect spin degeneracy).

$$\begin{aligned}\Psi_\pm(\hat{p}\mathbf{r}) &= \tfrac{1}{N!}\sum_p (\pm 1)^p\, \Psi(\hat{p}\mathbf{r}) \\ &= \Psi(\mathbf{r}) \quad \text{or} \quad 0 \qquad \text{depending on symmetry}\end{aligned}$$

where '$\pm$' denotes restriction to even (odd) functions for bosons (fermions), $\mathbf{r} \equiv (r_0, r_1, \cdots, r_{N-1})$, $\hat{p}\mathbf{r}$ means a permutation of the elements of $\mathbf{r}$ (note that $\mathbf{r}$ is an vector of length N containing 3d-states, not the $P \times N$ vector $\mathbf{x}$ used elsewhere).

The construction of the density matrix is as before, except that for bosons, only even states exist, and for fermions, only odd states. This forces us to perform a sum restricted to only even / odd states respectively.

$$\rho_\pm(\mathbf{r}, \mathbf{r}'; \beta) = \sum_{k\pm} e^{-\beta E_k}\, \Psi_k(r)\, \Psi_k(r).$$

However, restricted sums are inconvenient for building path integrals, so we would like to convert it into unrestricted sum where the even / odd solutions are predicted out. We do this using the symmetrisation / antisymmetrisation above. Then we can rearrange the sum and notice that it translates to a symmetrised / antisymmetrised density matrix.

$$\begin{aligned}\rho_\pm(\mathbf{r}, \mathbf{r}'; \beta) &= \sum_k e^{-\beta E_k}\, \tfrac{1}{N!} \sum_p (\pm 1)^p\, \Psi_k(\mathbf{r})\, \Psi_k(\hat{p}\mathbf{r}') \\ &= \tfrac{1}{N!} \sum_p (\pm 1)^p \sum_k e^{-\beta E_k}\, \Psi_k(\mathbf{r})\, \Psi_k(\hat{p}\mathbf{r}'), \\ \rho_\pm(\mathbf{r}, \mathbf{r}'; \beta) &= \tfrac{1}{N!} \sum_p (\pm 1)^p\, \rho_{\text{boltz}}(\mathbf{r}, \hat{p}\mathbf{r}'; \beta).\end{aligned}$$

The permutations have the effect of joining the chains between different particles (unlike the distinguishable case where they are separate), so that they 'grow together'. For fermions the permutation operator $\hat{p}$ antisymmetrises the density matrix. This is a problem because an anti-symmetric function must be negative somewhere, and thus the the diagonal elements of the density matrix cannot be interpreted naively as a probability density function, which cannot be negative (diagonal meaning $r_1 = r_1'$, not $r_1 = r_2$).

8.4 From Path Integral to P.I.M.D.

Traditional Molecular Dynamics samples states using the true physical dynamics, using true momenta and forces derived from the form of $\phi(x)$. The momenta can be initialised from a gaussian distribution.

$$\begin{aligned} \dot{x} = p/m \quad , \qquad \dot{p} = F \quad , \qquad F = -\nabla\phi(x) \quad , \\ \rho(p) = e^{-\beta\frac{p^2}{2m}} \qquad \Longrightarrow \qquad p_{\text{init}} = \sqrt{\tfrac{m}{\beta}} \times \text{Gaussian}() \quad . \end{aligned}$$

Unfortunately, the true physical dynamics conserves energy, and therefore traditional MD samples only the microcanonical ensemble (NVE - Number of particles, Energy and Volume are conserved), assuming ergodicity.

$$\begin{aligned} \Gamma &= \int_x \int_p \exp\left\{-\beta H(x,p)\right\}\ \delta(H-E), \\ \langle A\rangle &= \frac{1}{\Gamma}\int_x \int_p A(x)\ \exp\left\{-\beta H(x,p)\right\}\ \delta(H-E), \\ &\quad \text{where} \quad H(x,p) = \tfrac{p^2}{2m} + \phi(x). \end{aligned}$$

But we are interested in sampling a Boltzmann distribution, which is the canonical, or NVT ensemble (constant Temperature but not constant energy due to contact with a heat-bath).

$$\begin{aligned} Z &= \int_x e^{-\beta\phi(x)}, \qquad \text{or} \\ &= \int_x \int_p \exp\left\{-\beta H(x,p)\right\}, \\ \langle A\rangle &= \frac{1}{Z}\int_x \int_p A(x)\ \exp\left\{-\beta H(x,p)\right\}, \end{aligned}$$

which we shall accomplish by introducing a dynamics that connects the system to heat bath in such a way that the equations of motion are ergodic.

Canonical Dynamics

We need some way of introducing the conceptual 'heat-bath' into the molecular dynamics. There are many possible methods, including stochastic methods which scale the momenta, but we prefer deterministic methods if possible, not least because they can be time-reversible. One deterministic method is that of Nosé and Hoover [Hoo85], which adds the following to the dynamics

$$\begin{aligned} \dot{x} = \frac{p}{m} \quad , \qquad \dot{p} &= F - p\frac{p_\theta}{Q} \quad , \qquad \dot{\theta}_n = \frac{p_{\theta n}}{Q} \quad , \\ \dot{p}_\theta &= \left[\frac{p^2}{m} - kT\right] \quad . \end{aligned}$$

where Q is a mass-like parameter that controls the coupling between momenta and thermostats.

However, it has been shown [LLM09] that Nosé-Hoover dynamics is not ergodic for 1d systems such as oscillators. In real systems, heat-energy is conducted away from molecule to molecule, so for any given molecule its immediate 'heat bath' also fluctuates. This effect can itself be modelled, and the momentum sampling distribution improved, by themostatting the chains in turn, leading to the Nosé Hoover Chain method [MKT92], where the n-th thermostat is moderated by the $(n+1)$-th.

$$\begin{aligned} \dot{x} = \frac{p}{m} \quad , \qquad \dot{p} &= F - p\frac{p_{\theta 0}}{Q} \quad , \qquad \dot{\theta}_n = \frac{p_{\theta n}}{Q} \quad , \\ \dot{p}_{\theta 0} &= \left[\frac{p^2}{m} - kT\right] - p_{\theta 0}\frac{p_{\theta 1}}{Q} \quad , \\ \dot{p}_{\theta n} &= \left[\frac{p^2_{\theta(n-1)}}{Q} - kT\right] - p_{\theta n}\frac{p_{\theta(n+1)}}{Q} \quad , \\ \dot{p}_{\theta(N-1)} &= \left[\frac{p^2_{\theta(N-2)}}{Q} - kT\right] \quad , \qquad \text{where} \quad Q = \frac{kT}{\omega^2} \, , \end{aligned}$$

where ω is a ballpark estimate of the thermostat pumping frequency. This dynamics has the following energy-like quantity (which is not a true Hamiltonian),

$$H' = V(x) + \frac{p^2}{2m} + \sum_n \left[\frac{p^2_{\theta n}}{2Q} - kT\theta_n\right] .$$

The Nosé Hoover Chain method does improve the sampling of the state-space, by making it more chaotic, but it is still not ergodic.

8.4.1 Molecular Dynamics on a Cyclic Polymer

In the quantum limit, the partition function is much more complicated, but, when it is written down as a path integral, it suggests an analogy to Molecular Dynamics. For example, for a position operator $\hat{A}$,

$$\begin{aligned}\langle A\rangle &= \frac{1}{Z}\int \mathrm{d}^P x \left[\frac{1}{P}\sum_i A(x_i)\right]\left[\prod_i \rho(x_i, x_{i+1};\tau)\right] \\ &= \frac{1}{Z}\int_{\mathbf{x}} A_{\text{eff}}(\mathbf{x})\; e^{-\beta\phi_{\text{eff}}(\mathbf{x})} \quad ,\end{aligned}$$

$$\begin{aligned}\text{where}\quad \mathbf{x} &= (x_0, x_1, \cdots, x_{\mathrm{P}-1}) \quad , \qquad \int_{\mathbf{x}} = \int \mathrm{d}x_0\, \mathrm{d}x_0\, \ldots\, \mathrm{d}x_{P-1} \quad , \\ A_{\text{eff}}(\mathbf{x}) &= \frac{1}{P}\sum_i A(x_i) \quad , \\ \phi_{\text{eff}}(\mathbf{x}) &= -\frac{1}{\beta}\log\left[\prod_i \rho(x_i, x_{i+1};\tau)\right] \\ &= -\frac{1}{\beta}\sum_i \log\rho(x_i, x_{i+1};\tau) \quad .\end{aligned}$$

This works for distinguishable particles and bosons, because the density matrix is always positive definite, so that the logarithm is always meaningful.

We can add momenta to the partition function by inserting gaussian integrals of the momenta via the following identity

$$1 = \sqrt{\frac{\beta}{2\pi m_{\text{eff}}}} \times \int_p \exp\left\{-\beta\frac{p^2}{2m_{\text{eff}}}\right\} \quad ,$$

$$Z_P = \left[\tfrac{\beta}{2\pi m_{\text{eff}}}\right]^{\frac{P}{2}} \int_{\mathbf{x}}\int_{\mathbf{p}} \exp\left\{-\beta\tfrac{\mathbf{p}^2}{2m_{\text{eff}}} - \beta\phi_{\text{eff}}(\mathbf{x})\right\} \quad .$$

This implies a pseudo-hamiltonian, and the partition function can be renormalised to remove the prefactor.

$$Z_P = \int_{\mathbf{x}}\int_{\mathbf{p}} \exp\left\{-\beta H_{\text{eff}}(\mathbf{x},\mathbf{p})\right\}, \qquad \text{where} \quad H_{\text{eff}}(\mathbf{x},\mathbf{p}) = \frac{\mathbf{p}^2}{2m_{\text{eff}}} + \phi_{\text{eff}}(\mathbf{x}).$$

In other words, a molecular dynamics simulation of a system in N-coordinates becomes a molecular dynamics simulation in NP dimensions; a chain of pseudo-classical simulations that influence each other. This is analogous to sampling the states of a cyclic polymer except that there are no steric effects; the parts of the chain can both overlap and pass through each other.

Thermostatting for PIMD

Fortunately in PIMD we can apply these thermostats simply to the sampling momenta in exactly the same way as to a classical simulation; the only difference being that there are now P times as many momenta. Performing NVT simulations with our code, one massive Nosé Hoover Chain was added to thermostat every bead of every coordinate of the chain, which is massive thermostatting. That is, $d \times N \times P$ thermostat-chains, where d is the number of dimensions, N the number of particles, and P the number of beads. In general, the more thermostats there are, the more 'random' the heat-bath is and thus the better the sampling of the state-space. Even for Nosé Hoover Chains, for a multidimensional system dominated by oscillators, it is inadequate to use only one chain for the whole system; it produces unrealistic correlations across the system [MKT92]. For small systems, such as the dimer, large numbers of thermostats are relatively expensive to run, taking up to 80% of the runtime, but on large systems with many interaction forces to calculate, they are not computationally intensive.

Effective Potential for Free-Particle reference

In the simplest form of operator-splitting, the free particle reference, where there is a momentum-dependent kinetic term, and position-dependent 'external' potential,

$$\phi_{\text{eff}}(\mathbf{x}) = \frac{mP}{2\hbar^2\beta^2}\sum_i (x_i - x_{i+1})^2 + \frac{1}{P}\sum_i V(x_i)\ .$$

Note there are tension-like terms and external potential terms present; we have a chain of beads connected by harmonic bonds (with strength propotional to P to maintain the length of the chain with P), in which each bead moves on the external potential (softened by a factor P to maintain the total influence of the potential with P). P is only a discretisation parameter and should not affect the behaviour of the chain, at least in the large P limit.

Effective Potential for Harmonic reference

The basic reference that was used by Martyna and Whitfield [WM07], is the harmonic oscillator density matrix. Again there are tension-like and potential-like terms.

$$\rho_{\text{h.o.}}(x,x';\tau) = \exp\left\{\tfrac{1}{2}\mathcal{N} - \tfrac{1}{2}\mathcal{A}(x-x')^2 - \tfrac{1}{2}\mathcal{B}(x^2+x'^2) - \frac{\beta}{P}\sum_i \phi_{\text{ext}}(x_i)\right\},$$

$$\begin{aligned} \text{where} \quad f &= \tau\hbar\omega &&= \tfrac{1}{P}\beta\hbar\omega \quad , \\ \mathcal{N} &= \log\left[\tfrac{m\omega}{2\pi\hbar\sinh f}\right] \quad , \\ \mathcal{A} &= \tfrac{m\omega}{\hbar}\tfrac{1}{\sinh f} &&= \tfrac{mP}{\hbar^2\beta}\times\tfrac{f}{\sinh f} \quad , \\ \mathcal{B} &= \tfrac{m\omega}{\hbar}\tanh\tfrac{f}{2} &&= \tfrac{mP}{\hbar^2\beta}\times f\tanh\tfrac{f}{2} \quad , \end{aligned}$$

$$\Longrightarrow$$

$$\phi_{\text{eff}}(\mathbf{x}) = -\tfrac{NP}{2\beta}\mathcal{N} + \tfrac{1}{2\beta}\mathcal{A}\sum_i (x_i - x_{i+1})^2 + \tfrac{1}{\beta}\mathcal{B}\sum_i x_i^2 + \tfrac{1}{P}\sum_i \phi_{\text{ext}}(x_i).$$

Staging: Faux-masses

Above, we derived a pseudo-hamiltonian.

$$H_{\text{eff}}(\mathbf{x},\mathbf{p}) = \frac{\mathbf{p}^2}{2m_{\text{eff}}} + \phi_{\text{eff}}(\mathbf{x}) \quad .$$

Only ϕ_{eff} is fixed, and although the pseudo-kinetic energy has to be controlled by β, because the momenta are pseudo-momenta we are free to choose a pseudo-mass m_{eff} to maximise efficiency. These choices constitute a particular *staging* scheme. This can include evolving the dynamics in a transformed coordinate basis, (which is equivalent to choosing m_{eff} to be a non-diagonal matrix - that is $\mathbf{p}^2/2m \to \mathbf{p}^{\text{T}}\mathbf{M}^{-1}\mathbf{p}/2$), or by setting a different mass for each coordinate.

Staging: Coordinate transformation

Following the same approach as Ceperley used for the free-particle reference density matrix [PC84], Martyna and Whitfield used the following coordinate transformation

for the harmonic-oscillator reference [WM06, WM07]:

$$\begin{aligned}
\prod_{i=0}^{\mathrm{P}-1} \rho(x_i, x_{i+1}; \tau) &= \rho(x_0, x_P; \beta) \times \prod_{i=1}^{\mathrm{P}-1} \left[\frac{\rho(x_0, x_i; i\tau)\rho(x_i, x_{i+1}; \tau)}{\rho(x_0, x_{i+1}; (i+1)\tau)} \right] \\
&= \rho(x_0, x_0; \beta) \times \prod_{i=1}^{\mathrm{P}-1} \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left(-\frac{(x_i - x_i^*)^2}{2\sigma_i^2} \right) , \\
\text{where} \quad x_i^* &= \frac{\sinh f}{\sinh(i+1)f} \times x_0 + \frac{\sinh if}{\sinh(i+1)f} \times x_{i+1} , \\
\sigma_i^2 &= \frac{\hbar}{m\omega} \frac{\sinh f \, \sinh if}{\sinh(i+1)f} , \qquad f \equiv \tau\hbar\omega .
\end{aligned}$$

The special case is $i = 0$, which now simply obeys the harmonic oscillator density distribution:

$$u_0 = x_0 \quad , \qquad \sigma_0^2 = \frac{\hbar}{m\omega} \frac{1}{2 \tanh \beta\hbar\omega/2} \quad .$$

This suggests a coordinate transformation

$$\begin{aligned}
u_0 &= x_0 , \\
u_i &= x_i - \tfrac{\sinh f}{\sinh(i+1)f} \, x_0 - \tfrac{\sinh if}{\sinh(i+1)f} \, x_{i+1} , \qquad \text{where} \quad x_{i+P} \equiv x_i .
\end{aligned}$$

Note that if one refers to index 0 as index P, then this transformation could be represented by a triangular matrix. The determinant of a triangular matrix is the product of its diagonal elements; therefore the Jacobian of the transformation $x \to u$ is 1. In the code, we define two arrays S^1 and S^i to speed the calculation:

$$\begin{aligned}
u_i &= x_i - S_i^1 \, x_0 - S_i^i \, x_{i+1} , \\
S_i^1 &\equiv \tfrac{\sinh f}{\sinh(i+1)f} , \qquad S_i^i \equiv \tfrac{\sinh if}{\sinh(i+1)f} .
\end{aligned}$$

This transformation can be reversed, but needs a little care. It can be done step-wise by cycling down the indices from $i = 0$ $(+P)$ to $i = 1$:

$$\begin{aligned}
x_0 &= u_0 , \\
x_i &= u_i + S_i^1 \, x_0 + S_i^i \, x_{i+1} .
\end{aligned}$$

Therefore, we can also introduce momenta conjugate to the 'u' and obtain the staging fictitious classical Hamiltonian. This staging ansatz is useful as it diagonalises the harmonic part of the action, as well as reducing the range of frequencies in the problem, reducing the range of timescales involved, and thus giving us improved sampling of the

slowest timescales [TBJK93].

Internal MD-Forces and Momenta

The off-diagonal 'tension' terms have been transformed away, leaving a set of harmonic oscillators, that are independent until re-coupled by an external potential.

$$\begin{aligned} \phi_i^{(u)} &= \tfrac{u_i^2}{2\beta\sigma_i^2} = \tfrac{1}{2}k_i u_i^2 \,, \qquad \text{where} \quad k_i = \tfrac{1}{\beta\sigma_i^2} \,, \\ F_i^{(u)} &= -k_i\, u_i \,. \end{aligned}$$

This makes it very simple to evolve the 'internal' dynamics of the chain.

The effective masses m_{eff} still need to be chosen. In order to maximise sampling efficiency, each u coordinate should vary on the same timescale, and as we are not using any true physical dynamics, we can choose an arbitrary sampling frequency $\omega_{\text{PIMD}} = \gamma\,\omega_{\text{MD}}$, where ω_{MD} is the fundamental frequency of the nuclear motion.

$$\begin{aligned} \omega_{\text{PIMD}} &= \gamma\,\omega_{\text{MD}} \,, \\ k_i &= m_i\,\omega_{\text{PIMD}}^2 \,, \\ m_i &= \tfrac{k_i}{\omega_{\text{PIMD}}^2} = \tfrac{1}{\beta\sigma_i^2\omega_{\text{PIMD}}^2} \,. \end{aligned} \tag{8.1}$$

The momenta can be sampled from a gaussian distribution in a manner analogous to classical MD:

$$p_i^{(u)} = \sqrt{\tfrac{m_i}{\beta}} \times \text{Gaussian}() \,.$$

External MD-Forces

The 'external' forces are more complicated because they have to be calculated in the normal x-basis.

$$\begin{aligned} F_i^{(x)} &= -\frac{\partial}{\partial x_i}\phi_{\text{ext}}(\mathbf{x}) \,, \\ F_i^{(u)} &= -\frac{\partial}{\partial u_i}\phi_{\text{ext}}\left[\mathbf{u}(\mathbf{x})\right] = -\frac{\partial x_j}{\partial u_i}\frac{\partial}{\partial x_j}\phi_{\text{ext}}(\mathbf{x}) = \frac{\partial x_j}{\partial u_i}F_j^{(x)} \,. \end{aligned}$$

This is difficult to evaluate directly (the matrix $\frac{\partial x_j}{\partial u_i}$ is very complicated), but there is another good iterative method we can use. First we calculate the $i = 0$ term, which

behaves differently from the others, and defines a new array D is suggested.

$$\begin{aligned} u_0 = x_0 \qquad & \Longrightarrow \; D_0 \equiv \frac{\partial x_0}{\partial u_0} = 1 \; , \\ x_i = u_i + S_i^1 \, u_0 + S_i^i \, x_{i+1} \qquad & \Longrightarrow \; D_i \equiv \frac{\partial x_i}{\partial u_0} = S_i^1 + S_i^i \frac{\partial x_{i+1}}{\partial u_0} \; . \end{aligned}$$

This array D can be constructed iteratively. It is used to transform forces, and also in the calculation of the staging virial estimator discussed in the next section.

$$\begin{aligned} & D_0 = 1 \quad , \qquad D_i = S_i^1 + S_i^i \, D_{i+1} \quad \text{[cycling } i \text{ down from (P-1) to 1]} \; , \\ & \frac{\partial \phi_{\rm ext}}{\partial u_0} \equiv \sum_i D_i \frac{\partial \phi_{\rm ext}}{\partial x_i} \; . \end{aligned}$$

The remaining terms are calculated by a different iteration.

$$\begin{aligned} u_j &= x_j - S_j^1 \, x_0 - S_j^i \, x_{j+1} \; , \\ \frac{\partial u_j}{\partial x_i} &= \delta_{i,j} + 0 - S_j^i \, \delta_{i,j+1} \qquad \text{[for} \quad i \neq 0] \; , \\ \frac{\partial \phi_{\rm ext}}{\partial x_i} &= \frac{\partial u_j}{\partial x_i} \frac{\partial \phi_{\rm ext}}{\partial u_j} \\ &= \left[\delta_{i,j} - S_j^i \, \delta_{i,j+1} \right] \frac{\partial \phi_{\rm ext}}{\partial u_j} = \frac{\partial \phi_{\rm ext}}{\partial u_i} - S_{i-1}^i \, \frac{\partial \phi_{\rm ext}}{\partial u_{i-1}} \; , \\ \frac{\partial \phi_{\rm ext}}{\partial u_i} &= \frac{\partial \phi_{\rm ext}}{\partial x_i} + S_{i-1}^i \, \frac{\partial \phi_{\rm ext}}{\partial u_{i-1}} \qquad \text{(rearranged)} \; . \end{aligned}$$

Together these allow the transformation of forces from the x-basis to the u-basis where the dynamics are calculated.

$$\begin{aligned} F_0^{(u)} &= \sum_i D_i \, F_i^{(x)} \; , \\ F_i^{(u)} &= F_i^{(x)} + S_{i-1}^i \, F_{i-1}^{(u)} \qquad \text{[cycling } i \text{ up from 1 to (P-1)]} \; . \end{aligned}$$

Summary of Staging Transformation

$$\begin{aligned} x_0 &= u_0 \; , & x_i &= u_i + S_i^1 \, u_0 + S_i^i \, x_{i+1} \; , \\ F_0^{(u)} &= \textstyle\sum_i D_i F_i^{(x)} \; , & F_i^{(u)} &= F_i^{(x)} + S_{i-1}^i \, F_{i-1}^{(u)} \; , \\ \sigma_0^2 &= \frac{\hbar}{m\omega} \frac{1}{2 \tanh \beta\hbar\omega/2} \; , & \sigma_i^2 &= \frac{\hbar}{m\omega} \frac{\sinh f \, \sinh if}{\sinh (i+1) f} \; . \end{aligned}$$

Summary of Dynamics

$$
\begin{aligned}
F_i^{(u)} &= F_{\mathrm{int},i}^{(u)} + F_{\mathrm{ext},i}^{(u)}(\mathbf{x}) \,, \\
F_{\mathrm{int},i}^{(u)} &= -k_i \, u_i \,, \qquad\qquad \text{where} \quad k_i = \frac{1}{\beta \sigma_i^2} \,, \\
\dot{u}_i &= \frac{p_i^{(u)}}{m_i} \,, \qquad\qquad \text{where} \quad m_i = \frac{k_i}{\omega_{\mathrm{PIMD}}^2} \,.
\end{aligned}
$$

To summarise the thermostatting, the subscript i is dropped for clarity,

$$
\begin{aligned}
\dot{p}^{(u)} &= F - p^{(u)} \frac{p_{\theta 0}}{Q} \,, \\
\dot{p}_{\theta 0} &= \left[\frac{\left[p^{(u)}\right]^2}{m} - kT \right] - p_{\theta 0} \frac{p_{\theta 1}}{Q} \,, \\
\dot{p}_{\theta n} &= \left[\frac{p_{\theta (n-1)}^2}{Q} - kT \right] - p_{\theta n} \frac{p_{\theta (n+1)}}{Q} \,, \\
\dot{p}_{\theta (N-1)} &= \left[\frac{p_{\theta (N-2)}^2}{Q} - kT \right] \,, \\
\dot{\theta} &= \sum_n \frac{p_{\theta n}}{Q} \qquad \left(\text{where} \quad \theta = \sum_n \theta_n \right) \,, \\
\text{with} \quad Q &= \frac{kT}{\omega_{\mathrm{PIMD}}^2} = \frac{1}{\beta \omega_{\mathrm{PIMD}}^2} \,.
\end{aligned}
$$

They are initialised as follows:

$$
\begin{aligned}
u_i &= \sqrt{\sigma_i^2} \times \mathrm{Gaussian}() \,, \\
p_i^{(u)} &= \sqrt{\frac{m_i}{\beta}} \times \mathrm{Gaussian}() \,,
\end{aligned}
$$

$$
\begin{aligned}
\theta &= 0 \,, \\
p_{\theta i,n} &= \sqrt{\frac{Q}{\beta}} \times \mathrm{Gaussian}() \,.
\end{aligned}
$$

8.5 Evolving the Drude Atom Positions

The main purpose of the Drude model is to calculate the motion of the atoms moving on a good approximation of the ground state Born-Oppenheimer surface, an energy surface due to quantum charge-distribution effects.

Each atom is therefore represented by one coordinate, which all the beads share (it is effectively a centroid of the chain), that evolves classically. Here we denote heavy/classical coordinates by r. The coordinates of the quantum path are again represented by $\mathbf{x}$, but as will be explained below, they have been chosen to be relative coordinates that do not behave according to classical physical expectations, and so require some care.

The quantum charge-distribution effects can be thought of as 'flesh' on a skeleton of classical point coordinates. The flesh is shaped by the skeleton, and the motion of the skeleton is itself influenced by the resulting flexing of the flesh.

There are some components to the energy, which vary with r, but are constant with respect to the quantum dynamics (this includes the repulsive energy correction), and others which must be calculated by a potential of mean force.

$$
\begin{aligned}
Z(r,\beta) &= \int_x e^{-\beta\phi_{\text{eff}}(\mathbf{x};r,\beta)} \approx e^{-\beta E_0(r)} \ , \\
E(r) &\approx E_{\text{C}}(r) - \frac{1}{\beta}\log Z(r,\beta) \ , \\
F(r) &= -\frac{\partial}{\partial r}E(r) \\
&= -\frac{\partial}{\partial r}E_{\text{C}}(r) + \frac{1}{\beta}\frac{\partial}{\partial r}\log Z(r,\beta) \\
&= -\frac{\partial}{\partial r}E_{\text{C}}(r) - \frac{\int_x \left[\frac{\partial}{\partial r}\phi_{\text{eff}}\right] e^{-\beta\phi_{\text{eff}}}}{\int_x e^{-\beta\phi_{\text{eff}}}} \ .
\end{aligned}
$$

Thus the forces experienced by the atoms need to be averaged over the drude-states. The sampling is done by the drude-dynamics, which therefore has to be faster than the atom dynamics.

$$
m\ddot{r} = F_{\text{C}}(r) + \left\langle F_{\text{eff}}(\mathbf{x};r,\beta)\right\rangle_x \ .
$$

8.5.1 Sampling Efficiency: the faux-mass

In a PIMD simulation, this space-filling 'flesh' is represented by the rapidly-fluctuating quantum path or chain. However, the classical particles should feel only the mean influence of the statistical distribution of the quantum path, not the instantaneous

position, so if dynamics are important then it is important to ensure that any influence of the quantum path on the classical coordinates is sufficiently well averaged. The Q-Drude mass we use is not the true mass, but a fake mass that we have chosen appropriately; it will generally be several times smaller than those of the classical coordinate, so the velocities will generally be much faster; the Q-Drudes should buzz around the atoms several times before the atoms move significantly. The Q-Drude mass is closely related to the mass parameter used in Car-Parinello simulations and is chosen in a similar manner [TS01].

8.5.2 A False Temperature for the Q-Drudes

There is another major potential cost to P.I.M.D. ; the temperatures at which classical MD is performed are very very low for Quantum Drudes. For example a temperature of 300°K corresponds to $\beta\hbar\omega \equiv \hbar\omega/kT \approx 500$. The required bead number P is proportional to $\beta\hbar\omega$, and is even greater. This number of beads would be a prohibitive computational overhead. We would prefer to use many fewer.

Instead we note that for $\beta\hbar\omega \equiv \hbar\omega/kT = 10$ ($P \approx 10$, $T \approx 30,000$°K), a Q-Drude system will already be very close to its ground state, as we show below. This means that we make an approximation that vastly reduces the required bead number P; an artificially high temperature for the Q-Drude coordinates, while keeping the classical coordinates at a physically realistic temperature. This technique works well provided that it does not introduce any unwanted coupling between the two parts of the dynamics, for example the transfer of kinetic energy (heat) from hotter to colder. Given also that the commonly used thermostats, including Nosé Hoover Chains, are designed for equilibrium contexts, this technique requires that there be effectively an adiabatic separation between the Q-Drude and the classical coordinates.

Unfortunately adiabatic separation requires a large separation of mass- and thus time-scales. For an elastic collision, the quantity of kinetic energy transferred depends on the ratio of the masses. This is a problem also for Car-Parrinello simulations [TS01].

$$\frac{\partial}{\partial t}E_{1\to 2} \;\propto\; \frac{m_1 m_2}{(m_1 - m_2)^2} \;\approx\; \frac{m_<}{m_>}\,.$$

The total kinetic energy transfer rate depends also on the frequency of collisions, which

is proportional to ω_{PIMD}:

$$\begin{aligned}
\omega_{\mathrm{PIMD}} &\propto \gamma \qquad \text{(from eqn. 8.1 on pg. 150)} , \\
m_{\mathrm{PIMD}} &\propto \frac{1}{\omega_{\mathrm{PIMD}}^2} \propto \frac{1}{\gamma^2} , \\
\frac{\mathrm{d}E}{\mathrm{d}t} &\propto \omega_{\mathrm{PIMD}} \times m_{\mathrm{PIMD}} \propto \frac{1}{\gamma} .
\end{aligned} \tag{8.2}$$

In practice, one uses the biggest mass that can be tolerated empirically. The two sets of coordinates are kept at their appropriate temperatures by thermostats. The one thing that must be the same for both is the ground-state potential surface, which we approximate by moving on the free energy surface.

$$\begin{aligned}
Z_{\mathrm{drude}}(r) &= e^{-\beta_{\mathrm{drude}} E_0} + \sum_{i>0} e^{-\beta_{\mathrm{drude}} E_i} \\
&= e^{-\beta_{\mathrm{drude}} E_0} \left(1 + \sum_{i>0} e^{-\beta_{\mathrm{drude}} \Delta_i} \right) , \\
&\qquad \text{where} \quad \Delta_i = E_i - E_0 ,
\end{aligned}$$

$$\begin{aligned}
E_{\mathrm{free}}(r) &= -\frac{1}{\beta_{\mathrm{drude}}} \log Z(r, \beta_{\mathrm{drude}}) \\
&= E_0(r) - \frac{1}{\beta_{\mathrm{drude}}} \log \left(1 + \sum_{i>0} e^{-\beta_{\mathrm{drude}} \Delta_i} \right) \\
&\approx E_0(r) , \qquad \text{for} \quad \beta_{\mathrm{drude}} \Delta_i \gg 1 .
\end{aligned}$$

The atomic partition function uses this energy function:

$$\begin{aligned}
Z_{\mathrm{atom}} &= \int_r e^{-\beta_{\mathrm{atom}} E_{\mathrm{free}}(r)} \\
&= \int_r e^{\frac{\beta_{\mathrm{atom}}}{\beta_{\mathrm{drude}}} \log Z(r, \beta_{\mathrm{drude}})} \\
&= \int_r e^{-\beta_{\mathrm{atom}} E_0(r) + \frac{\beta_{\mathrm{atom}}}{\beta_{\mathrm{drude}}} \log\left(1 + \sum_{i>0} e^{-\beta_{\mathrm{drude}} \Delta_i}\right)} \\
&= \int_r e^{-\beta_{\mathrm{atom}} E_0(r)} \left(1 + \sum_{i>0} e^{-\beta_{\mathrm{drude}} \Delta_i} \right)^{\frac{\beta_{\mathrm{atom}}}{\beta_{\mathrm{drude}}}} , \\
&\approx \int_r e^{-\beta_{\mathrm{atom}} E_0(r)} \left(1 + \frac{\beta_{\mathrm{atom}}}{\beta_{\mathrm{drude}}} \sum_{i>0} e^{-\beta_{\mathrm{drude}} \Delta_i} \right) \\
&\approx \int_r e^{-\beta_{\mathrm{atom}} E_0(r)} .
\end{aligned}$$

So we can see that Z_{drude}, and thus Z_{atom}, converges to the ground state exponentially fast in β_{drude}.

8.5.3 Relative Coordinates for the Q-Drudes

Next, consider what happens when an isolated atom, with an attached Q-Drude, is moving in empty space, and experiences an impulse. If the Q-Drude path coordinates are treated as normal physical coordinates, the distribution of the Q-Drude will lag behind the atom which tugs it along. The true physical distribution would have a much smaller lag, if any. This can be corrected for by making the Q-Drude coordinates always relative to the centre of the harmonic bond, so that they never need to re-equilibrate their distributions relative to this dominant 'internal' part of the potential. They will not be in perfect equilibrium with the 'external' potential but this should be less significant. This distinction roughly corresponds to the difference between intramolecular and intermolecular forces in classical MD simulations.

This becomes important when calculating forces on the atoms: it means that all forces experienced by the Q-Drudes are also directly experienced by the molecular skeleton to which they are attached (not mediated by the tug of a harmonic bond or spring):

$$\begin{aligned} F_{\text{atom}ij} &= F\left[(r_i + x_i) - r_j\right] \\ &\quad \text{acts at } r_i \text{ and } r_j \text{ but not at } (r_i + x_i). \end{aligned}$$

This also has consequences for the pressure tensor, below.

8.5.4 Pressure

The matrix $\mathbf{h}$ is commonly used in MD simulations to define the size and shape of the periodic volume. The pressure is the response of the system to squashing of this volume. It can be used as a measurement estimator in NVT calculations, or as a 'force' in NPT calculations. The definition of the pressure tensor is [FS96],

$$\langle P_{\alpha\beta} \rangle = \left\langle \frac{NkT}{\det \mathbf{h}} + \frac{kT}{\det \mathbf{h}} h_{\beta\gamma} \frac{\partial \log Z}{\partial h_{\alpha\gamma}} \right\rangle .$$

In a classical system, this means

$$\begin{aligned}\langle P_{\alpha\beta}\rangle &= \frac{NkT}{\det \mathbf{h}} + \frac{1}{\det \mathbf{h}}\,\frac{1}{\beta}\left[\frac{\int_r h_{\beta\gamma}\frac{\partial}{\partial h_{\alpha\gamma}} e^{-\beta\phi(r)}}{\int_r e^{-\beta\phi(r)}}\right] \\ &= \frac{NkT}{\det \mathbf{h}} - \frac{1}{\det \mathbf{h}}\left[\frac{\int_r h_{\beta\gamma}\frac{\partial\phi(r)}{\partial h_{\alpha\gamma}} e^{-\beta\phi(r)}}{\int_r e^{-\beta\phi(r)}}\right] \\ &= \left\langle \frac{NkT}{\det \mathbf{h}} - \frac{1}{\det \mathbf{h}}\, h_{\beta\gamma}\frac{\partial\phi(r)}{\partial h_{\alpha\gamma}}\right\rangle .\end{aligned}$$

Analogously, in the PIMD formulation of a quantum system, this implies

$$\begin{aligned}\langle P_{\alpha\beta}\rangle &= \left\langle \frac{NkT}{\det \mathbf{h}} - \frac{1}{\det \mathbf{h}}\, h_{\beta\gamma}\frac{\partial\phi_{\text{eff}}(\mathbf{x}, r, \beta)}{\partial h_{\alpha\gamma}}\right\rangle , \\ &\quad \text{where} \quad \phi_{\text{eff}}(\mathbf{x}, r, \beta) = -\frac{1}{\beta}\log\left[\prod_i \rho(x_i, x_{i+1}; \tau)\right] , \quad \text{as previously.}\end{aligned}$$

Thus, the form of the pressure estimator also depends on the choice of reference density matrix.

8.5.5 Pressure with Relative Q-Drude Coordinates

In the same way that a Q-Drude does not lag behind and therefore get tugged along by its molecular skeleton (as described in the previous section for atomic forces), a Q-Drude in a large periodic box would experience little change in its distribution even if the box height was halved. Therefore, Q-Drude coordinates should not depend explicitly on the **h** matrix. This is in contrast to atomic coordinates, which are typically scaled by the **h** matrix to allow compressions of the whole box. As with the atomic forces, this means that elements of pressure tensor act with the force dependent on the distance between a Q-Drude and another charge, but on a line between the harmonic centre of the Q-Drude, and the charge (or in the case of two Q-Drudes, on the line between their

harmonic centres).

$$\begin{aligned} F_{\text{atom}ij} &= F\left[(r_i + x_i) - r_j\right] \\ &\quad \text{acts at } r_i \text{ and } r_j \text{ but not at } (r_i + x_i), \end{aligned}$$

$$P_{\alpha\beta,ij} = (r_i - r_j)_\beta F_\alpha \left[(r_i + x_i) - r_j\right] \ .$$

8.6 Measurement Estimators

Performing Molecular Dynamics (including PIMD) or Monte Carlo simulations, and making measurements for certain variables over a long time, is equivalent to integrating a function over state space by statistical sampling. Any function that would be sampled this way has a variance, and there will be some uncertainty in the result.

However, any integration can be refactored by an integration by parts, so that the same result is guaranteed to be obtained by integrating over the new function. If the new integrand function has lower variance, then the uncertainty of the result due to finite sampling will be likewise reduced. We refer to the original function and the new one as different 'estimators' for the same measurable.

Quantum Drude Oscillators have a potential problem in that the energy and dynamics of the unperturbed oscillators are far greater than what we are really interested in; the subtle perturbations due to external forces.

The simplest energy estimator, which we call the 'Barker' estimator, has a huge variance because of this. Here we outline a two-step improvement to this estimator, via the 'virial' estimator to the 'staging' estimator as outlined by [WM07]. Here we review these results.

8.6.1 Barker Energy Estimator

The 'thermodynamic' or 'Barker' estimator is the most basic estimator, as it is simply the derivative of the partition function with respect to β:

$$\begin{aligned} Z(\beta, P) &= \int_x \exp\Big\{ \tfrac{NP}{2}\mathcal{N} - \tfrac{1}{2}\mathcal{A}\textstyle\sum_i (x_i - x_{i+1})^2 - \mathcal{B}\sum_i x_i^2 - \tfrac{\beta}{P}\sum_i V_{\text{ext}}(x_i) \Big\}, \\ E(\beta, P) &= -\frac{1}{Z}\frac{\partial Z}{\partial \beta} = -\frac{\partial}{\partial \beta}\log(Z) = \langle E_{\text{B}} \rangle \,, \end{aligned}$$

where

$$\begin{aligned} E_{\rm B} &= -\tfrac{NP}{2}\dot{\mathcal{N}} + \tfrac{1}{2}\dot{\mathcal{A}}\sum_i (x_i - x_{i+1})^2 + \dot{\mathcal{B}}\sum_i x_i^2 + \tfrac{1}{P}\sum_i V_{\rm ext}(x_i) \\ &= \mathcal{E}_{\rm B} + \tfrac{1}{2}\mathcal{A}_{\rm B}\sum_i (x_i - x_{i+1})^2 + \mathcal{B}_{\rm B}\sum_i x_i^2 + \tfrac{1}{P}\sum_i V_{\rm ext}(x_i)\ , \\ & \quad \text{where } \mathcal{E}_{\rm B} \equiv -\tfrac{NP}{2}\dot{\mathcal{N}}\ , \qquad \mathcal{A}_{\rm B} \equiv \dot{\mathcal{A}}\ , \qquad \mathcal{B}_{\rm B} \equiv \dot{\mathcal{B}}\ . \end{aligned}$$

Using the free particle density matrix, we have $\mathcal{A} = \frac{mP}{\hbar^2\beta}$, $\mathcal{N} = \log(\mathcal{A}/2\pi)$, $\mathcal{B} = 0$.

$$\begin{aligned} E_{\rm B} &= \tfrac{NP}{2}\tfrac{1}{\beta} - \tfrac{mP}{\hbar^2\beta^2}\sum_i (x_i - x_{i+1})^2 + \tfrac{1}{P}\sum_i V_{\rm ext}(x_i) \\ &= \tfrac{NP}{2}kT - \tfrac{mP}{\hbar^2\beta^2}\sum_i (x_i - x_{i+1})^2 + \tfrac{1}{P}\sum_i V_{\rm ext}(x_i)\ . \end{aligned}$$

Using the Harmonic Oscillator density matrix, $V_{\rm ext} = \frac{1}{P}\sum_i V(x_i) - \frac{1}{2}m\omega^2 x_i^2$ is the 'external' potential (the local harmonic potential is called the 'internal' potential of a Quantum Drude). Using $f = \frac{1}{P}\beta\hbar\omega$; we write

$$\begin{aligned} \mathcal{A} &= \tfrac{m\omega}{\hbar \sinh f}\ , & &= \tfrac{mP}{\hbar^2\beta} \times \tfrac{f}{\sinh f}\ , \\ \mathcal{B} &= \tfrac{m\omega}{2\hbar}\tanh\tfrac{f}{2}\ , & &= \tfrac{mP}{\hbar^2\beta} \times f\tanh\tfrac{f}{2}\ , \\ \mathcal{N} &= \log(\mathcal{A}/2\pi)\ . \end{aligned}$$

$$\begin{aligned} E_{\rm B} &= \tfrac{N}{2}\hbar\omega\coth f - \tfrac{m\omega^2}{\tanh f \sinh f}\sum_i (x_i - x_{i+1})^2 + \tfrac{m\omega^2}{\cosh^2 f}\sum_i x_i^2 + \tfrac{1}{P}\sum_i V_{\rm ext}(x_i) \\ &= \tfrac{NP}{2\beta} f\coth f - \tfrac{mP}{\hbar^2\beta^2}\tfrac{f^2\coth f}{\sinh f}\sum_i (x_i - x_{i+1})^2 + \tfrac{mP}{\hbar^2\beta^2}\tfrac{f^2}{\cosh^2 f}\sum_i x_i^2 + \tfrac{1}{P}\sum_i V_{\rm ext}(x_i). \end{aligned}$$

8.6.2 Virial Energy Estimator

Following [HBB82], we can use the properties of the path integral partition function to express the integral in a different way. Define U as follows. Where possible the Einstein

summation convention is used on i.

$$\begin{aligned} Z &= \int_x e^{-U(\mathbf{x})} , \\ U &= -\tfrac{NP}{2}\mathcal{N} + \tfrac{1}{2}\mathcal{A}\left(x_i - x_{i+1}\right)^2 + \mathcal{B}x_i^2 + \tfrac{\beta}{P}\textstyle\sum_i V_{\text{ext}}(x_i) , \\ x_i\tfrac{\partial U}{\partial x_i} &= 2 \times \left[\tfrac{1}{2}\mathcal{A}\left(x_i - x_{i+1}\right)^2 + \mathcal{B}x_i^2\right] + \tfrac{\beta}{P}x_i\tfrac{\partial V_{\text{ext}}}{\partial x_i} , \end{aligned}$$

$$\begin{aligned} \left\langle x_i\tfrac{\partial U}{\partial x_i}\right\rangle &= \tfrac{1}{Z}\int_x x_i\tfrac{\partial U}{\partial x_i}e^{-U(\mathbf{x})} = \tfrac{1}{Z}\left(-\int_x x_i\tfrac{\partial}{\partial x_i}e^{-U(\mathbf{x})}\right) \\ &= \tfrac{1}{Z}\left(-\oint_{|x|\to\infty}|x|e^{-U(\mathbf{x})} + \int_x \tfrac{\partial x_i}{\partial x_i}e^{-U(\mathbf{x})}\right) \\ &= \tfrac{1}{Z}\left(0 + \delta_{ii}Z\right) \qquad \text{(for a localised particle)} \\ &= NP \qquad \text{(the trace of the identity matrix),} \end{aligned}$$

$$\begin{aligned} NP = \left\langle x_i\tfrac{\partial U}{\partial x_i}\right\rangle &= \mathcal{A}\left\langle\left(x_i - x_{i+1}\right)^2\right\rangle + 2\mathcal{B}\left\langle x_i^2\right\rangle + \tfrac{\beta}{P}\left\langle x_i\tfrac{\partial V_{\text{ext}}}{\partial x_i}\right\rangle , \\ \left\langle\left(x_i - x_{i+1}\right)^2\right\rangle &= \tfrac{NP}{\mathcal{A}} - 2\tfrac{\mathcal{B}}{\mathcal{A}}\left\langle x_i^2\right\rangle - \tfrac{\beta}{P\mathcal{A}}\left\langle x_i\tfrac{\partial V_{\text{ext}}}{\partial x_i}\right\rangle . \end{aligned}$$

Plug this into the Barker estimator to obtain the Virial estimator:

$$\begin{aligned} \langle E_{\text{B}}\rangle &= -\tfrac{NP}{2}\dot{\mathcal{N}} + \tfrac{1}{2}\dot{\mathcal{A}}\left\langle\left(x_i - x_{i+1}\right)^2\right\rangle + \dot{\mathcal{B}}\left\langle x_i^2\right\rangle + \tfrac{1}{P}\left\langle\textstyle\sum_i V_{\text{ext}}(x_i)\right\rangle , \\ \langle E_{\text{V}}\rangle &= -\tfrac{NP}{2}\dot{\mathcal{N}} + \tfrac{NP}{2}\tfrac{\dot{\mathcal{A}}}{\mathcal{A}} + \left(\dot{\mathcal{B}} - \mathcal{B}\tfrac{\dot{\mathcal{A}}}{\mathcal{A}}\right)\left\langle x_i^2\right\rangle + \tfrac{1}{P}\left\langle\textstyle\sum_i V_{\text{ext}}(x_i)\right\rangle - \tfrac{1}{2P}\tfrac{\beta\dot{\mathcal{A}}}{\mathcal{A}}\left\langle x_i\tfrac{\partial V_{\text{ext}}}{\partial x_i}\right\rangle , \end{aligned}$$

$$\begin{aligned} E_{\text{V}} &= \mathcal{E}_{\text{V}} + \mathcal{B}_{\text{V}}\sum_i x_i^2 + \tfrac{1}{P}\sum_i V_{\text{ext}}(x_i) - \mathcal{A}_{\text{V}}\tfrac{1}{P}\sum_i x_i\tfrac{\partial V_{\text{ext}}}{\partial x_i} , \\ &\text{where} \quad \mathcal{E}_{\text{V}} \equiv -\tfrac{NP}{2}\dot{\mathcal{N}} + \tfrac{NP}{2}\tfrac{\dot{\mathcal{A}}}{\mathcal{A}} , \quad \mathcal{A}_{\text{V}} \equiv \tfrac{1}{2}\tfrac{\beta\dot{\mathcal{A}}}{\mathcal{A}} , \quad \mathcal{B}_{\text{V}} \equiv \left(\dot{\mathcal{B}} - \mathcal{B}\tfrac{\dot{\mathcal{A}}}{\mathcal{A}}\right) . \end{aligned}$$

There is no free-particle density virial estimator, because the above derivation only works for a localised particle. In the case of the harmonic density matrix with an external potential, some terms cancel or simplify in the virial estimator. Using $f \equiv \frac{1}{P}\beta\hbar\omega$ as usual:

$$\begin{aligned} E_{\text{V}} &= \tfrac{m\omega^2}{P}\left\langle x_i^2\right\rangle + \tfrac{1}{P}\sum_i V_{\text{ext}}(x_i) + \tfrac{1}{2}f\coth(f)\tfrac{1}{P}\sum_i x_i\tfrac{\partial V_{\text{ext}}}{\partial x_i} \\ &= \tfrac{mP}{\hbar^2\beta^2}f^2\left\langle x_i^2\right\rangle + \tfrac{1}{P}\sum_i V_{\text{ext}}(x_i) + \tfrac{1}{2}f\coth(f)\tfrac{1}{P}\sum_i x_i\tfrac{\partial V_{\text{ext}}}{\partial x_i} . \end{aligned}$$

8.6.3 Staging Energy Estimator

The harmonic staging virial estimator [WM07] is the least-obvious estimator, and it is useful because it allows the unperturbed oscillator energy to be integrated out. Care needs to be taken because it uses the same potential as for MD in the staging basis, which may not be the same as the remainder/external potential used so far in the estimators (for example when we derive the dipole-approximation density matrix expansion below). We denote these ϕ_{ext} (the PIMD external potential for dynamics) and V_{ext} (the physical external potential) respectively.

In the staging coordinate basis, we can write down the partition function:

$$\begin{aligned} Z(\beta,P) &= \int_u \left\{ \prod_{i\neq 0} \sqrt{\tfrac{1}{2\pi\sigma_i^2}} \exp\left(\tfrac{u_i^2}{2\sigma_i^2}\right) \right\} \mathcal{N}_0 \exp\left(-\mathcal{B}_0 u_0^2 - \beta\phi_{\text{ext}}\left[\mathbf{x}(\mathbf{u})\right]\right) \ , \\ &\quad \text{where} \quad \mathcal{N}_0 = \sqrt{\tfrac{m\omega}{2\pi\hbar\sinh(\beta\hbar\omega)}} \ , \quad \mathcal{C}_0 = \tfrac{m\omega}{2\hbar}\tanh\left(\beta\hbar\omega/2\right) \ . \end{aligned}$$

Because $x_0 = u_0$, we can integrate out the quantity $\left\langle x_0^2 \right\rangle$, through another integration by parts, this time on u_0,

$$\begin{aligned} \left\langle x_0^2 \right\rangle = \left\langle u_0^2 \right\rangle &= \tfrac{1}{Z}\int_u \left\{ \prod_{i\neq 0} \sqrt{\tfrac{1}{2\pi\sigma_i^2}} \exp\left(\tfrac{u_i^2}{2\sigma_i^2}\right) \right\} \sqrt{\tfrac{m\omega}{2\pi\hbar\sinh(\beta\hbar\omega)}} \times \\ &\quad \left\{u_0 \exp\left(-\mathcal{C}_0 u_0^2\right)\right\} \times \left\{u_0 \exp\left(-\beta\phi_{\text{ext}}\right)\right\} \ , \end{aligned}$$

$$\begin{aligned} &\int_{u_0} \left\{u_0 \exp\left(-\mathcal{C}_0 u_0^2\right)\right\} \left\{u_0 \exp\left(-\beta\phi_{\text{ext}}\right)\right\} \\ &= \oint_{|u_0|=\infty} -\tfrac{1}{2\mathcal{C}_0}|u_0| \exp\left(-\mathcal{C}_0 u_0^2\right) \exp\left(-\beta\phi_{\text{ext}}\right) \\ &\quad - \int_{u_0} \left(-\tfrac{1}{2\mathcal{C}_0}\right) \left(1 - \beta u_0 \tfrac{\partial\phi_{\text{ext}}}{\partial u_0}\right) \exp\left(-\mathcal{C}_0 u_0^2\right) \exp\left(-\beta\phi_{\text{ext}}\right) \ , \end{aligned}$$

$$\begin{aligned} \left\langle x_0^2 \right\rangle &= 0 + \tfrac{1}{2\mathcal{C}_0} - \tfrac{\beta}{2\mathcal{C}_0}\left\langle u_0 \tfrac{\partial\phi_{\text{ext}}}{\partial u_0}\right\rangle = \tfrac{1}{2\mathcal{C}_0} - \tfrac{\beta}{2\mathcal{C}_0}\left\langle u_0 \tfrac{\partial x_i}{\partial u_0}\tfrac{\partial\phi_{\text{ext}}}{\partial x_i}\right\rangle \\ &= \tfrac{1}{2\mathcal{C}_0} - \tfrac{\beta}{2\mathcal{C}_0}\left\langle x_0 \mathrm{D}_i \tfrac{\partial\phi_{\text{ext}}}{\partial x_i}\right\rangle \ , \end{aligned}$$

where the $D_i = \frac{\partial x_i}{\partial u_0}$ were derived previously in the section on staging forces.

The $\langle x_0^2 \rangle$ term was easy to calculate because $x_0 = u_0$, but it is not immediately obvious what can be done about $\langle x_i^2 \rangle$ for $i \neq 0$. The next step is to observe that although the u's are not equivalent, they depend on an arbitrary choice of labelling for the x's. The x's *are* equivalent before any transformation, so we could rotate the chain to choose any another bead to be x_0. This is an equally valid choice for the transformation $\mathbf{x} \rightarrow \mathbf{u}$, which would give us a different $\mathbf{u}$, which could give an alternative set of dynamics for MD. As the estimators do not depend on the actual dynamics, we can take the same estimator, rotate it in P, and apply it to each x coordinate in turn: $x_i \rightarrow x_{i+\imath}$, or $\mathrm{D}_i \rightarrow \mathrm{D}_{i-\imath}$.

$$\langle x_i^2 \rangle \quad = \quad \tfrac{1}{2\mathcal{C}_0} - \tfrac{\beta}{2\mathcal{C}_0} \tfrac{1}{P} \left\langle x_\imath \mathrm{D}_{(i-\imath)} \tfrac{\partial \phi_{\mathrm{ext}}}{\partial x_i} \right\rangle \ .$$

Inserting this into the expression for the virial estimator,

$$E_{\mathrm{V}} = \mathcal{E}_{\mathrm{V}} + \mathcal{B}_{\mathrm{V}} \sum_i x_i^2 + \tfrac{1}{P} \sum_i \phi_{\mathrm{ext}}(x_i) - \mathcal{A}_{\mathrm{V}} \tfrac{1}{P} \sum_i x_i \tfrac{\partial \phi_{\mathrm{ext}}}{\partial x_i} \ ,$$
$$\text{where} \quad \mathcal{E}_{\mathrm{V}} \equiv -\tfrac{NP}{2} \dot{\mathcal{N}} + \tfrac{NP}{2} \tfrac{\dot{\mathcal{A}}}{\mathcal{A}} \ , \qquad \mathcal{A}_{\mathrm{V}} \equiv \tfrac{1}{2} \tfrac{\beta \dot{\mathcal{A}}}{\mathcal{A}} \ , \qquad \mathcal{B}_{\mathrm{V}} \equiv \left(\dot{\mathcal{B}} - \mathcal{B} \tfrac{\dot{\mathcal{A}}}{\mathcal{A}} \right) \ .$$

gives the staging virial estimator:

$$E_{\mathrm{St}} = \mathcal{E}_{\mathrm{V}} + \mathcal{E}_{\mathrm{St}} + \tfrac{1}{P} \textstyle\sum_i \phi_{\mathrm{extr}}(x_i) - \mathcal{A}_{\mathrm{V}} \tfrac{1}{P} \sum_i x_i \tfrac{\partial \phi_{\mathrm{ext}}}{\partial x_i} - \mathcal{B}_{\mathrm{St}} \tfrac{1}{P} \sum_{i\imath} x_\imath \mathrm{D}_{(i-\imath)} \tfrac{\partial \phi_{\mathrm{ext}}}{\partial x_i} \ ,$$
$$\text{where} \quad \mathcal{B}_{\mathrm{St}} \equiv \tfrac{\beta}{2P\mathcal{C}_0} \left(\dot{\mathcal{B}} - \mathcal{B} \tfrac{\dot{\mathcal{A}}}{\mathcal{A}} \right) \ .$$

For the harmonic density matrix, this means:

$$\begin{aligned} E_{\mathrm{St}} \quad = \quad & \tfrac{1}{2} \hbar\omega \coth\left(\tfrac{\beta\hbar\omega}{2} \right) + \tfrac{1}{P} \sum_i \phi_{\mathrm{ext}}(x_i) + \tfrac{1}{2} f \coth(f) \tfrac{1}{P} \sum_i x_i \tfrac{\partial \phi_{\mathrm{ext}}}{\partial x_i} \\ & - \tfrac{1}{2} \beta\hbar\omega \coth\left(\tfrac{\beta\hbar\omega}{2} \right) \tfrac{1}{P} \sum_{i\imath} x_\imath \mathrm{D}_{(i-\imath)} \tfrac{\partial \phi_{\mathrm{ext}}}{\partial x_i} \ . \end{aligned}$$

8.6.4 Summary of Energy Estimators

If we have a partition function of the following form:

$$Z = \exp\Big\{ \tfrac{1}{2}\mathcal{N} - \tfrac{1}{2}\mathcal{A}\textstyle\sum_i (x_i - x_{i+1})^2 - \mathcal{B}\sum_i x_i^2 - \frac{\beta}{P}\sum_i V(x_i) \Big\} ,$$

$$E_{\rm B}(\beta,P) = \mathcal{E}_{\rm B} + \tfrac{1}{2}\mathcal{A}_{\rm B}\textstyle\sum_i (x_i - x_{i+1})^2 + \mathcal{B}_{\rm B}\sum_i x_i^2 + \frac{1}{P}\sum_i V(x_i) ,$$

$$E_{\rm V}(\beta,P) = \mathcal{E}_{\rm V} + \mathcal{B}_{\rm V}\textstyle\sum_i x_i^2 + \frac{1}{P}\sum_i V(x_i) - \mathcal{A}_{\rm V}\frac{1}{P}\sum_i x_i \nabla_i V ,$$

$$\begin{aligned} E_{\rm St}(\beta,P) &= \mathcal{E}_{\rm V} + \tfrac{1}{P}\textstyle\sum_i V(x_i) - \mathcal{A}_{\rm V}\frac{1}{P}\sum_i x_i \nabla_i V , \\ &\quad + \big(\mathcal{E}_{\rm St} - \mathcal{B}_{\rm St}\frac{1}{P}\sum_{ii} x_i {\rm D}_{(i-i)} \nabla_i \phi_{\rm ext} \big) . \end{aligned}$$

For the harmonic oscillator reference density matrix, these evaluate as:

$$\begin{aligned} \mathcal{E}_{\rm B} &= \tfrac{N}{2}\hbar\omega \coth f &&= \tfrac{NP}{2\beta} f \coth f , \\ \mathcal{E}_{\rm V} &= 0 , \\ \mathcal{E}_{\rm St} &= \tfrac{N}{2}\hbar\omega \coth\Big(\tfrac{\beta\hbar\omega}{2}\Big) , \\ \mathcal{A}_{\rm B} &= -\tfrac{m\omega^2}{\tanh f \sinh f}\tfrac{1}{P} &&= -\tfrac{mP}{\hbar^2\beta^2}\tfrac{f^2}{\tanh f \sinh f} , \\ \mathcal{A}_{\rm V} &= -\tfrac{1}{2} f \coth(f) , \\ \mathcal{B}_{\rm B} &= \tfrac{m\omega^2}{\cosh^2 f}\tfrac{1}{P} &&= \tfrac{mP}{\hbar^2\beta^2}\tfrac{f^2}{\cosh^2 f} , \\ \mathcal{B}_{\rm V} &= \tfrac{m\omega^2}{P} &&= \tfrac{mP}{\hbar^2\beta^2} f^2 , \\ \mathcal{B}_{\rm St} &= \tfrac{1}{2}\beta\hbar\omega \coth\Big(\tfrac{\beta\hbar\omega}{2}\Big)\tfrac{1}{P} . \end{aligned}$$

On the right hand side, these expressions have also been written with the ω's incorporated into f's in order to make it simpler to do convert them into a dipole quadratic expansion in λ, which we will do in the next section.

Chapter 9

New High-T Density Matrices for the QDO

We briefly discuss the challenge of providing more accurate discretisations of the quantum-path density matrix; to reduce the number of 'beads' required to obtain accurate results. We present a new dipole-dipole density matrix derived by analogy to the matrix diagonalisation methods of chapter 3. We also sketch some approaches to porting the high-quality DMC 'Coulomb' wavefunction into a density matrix for PIMD, reasoning that it will result in a similar improvement in convergence.

With DMC, we found that a suitable choice of wavefunction dramatically improved the probability distribution, reduced its reliance on the clumsy external potential operator (the population operator), and improving the energy-convergence by reducing the variance of the estimator. We also found that improved propagators markedly improved the time-step dependence.

PIMD does not map exactly to DMC, but it would be good if we could use some of the insights gained, as well as any new ones, to improve the behaviour of the density matrix, perhaps allowing for a reduction in the bead number P. In particular, we would like to reduce the number of expensive long-range calculations made around the chain.

Another general approach would be 'bead RESPA', where we use the method of time-scale separation used by Martyna and Tuckerman[TBM92], but applied in an imaginary time coordinate around the chain. This would allow us to calculate fewer of the expensive, but slowly-varying, long-range energy components. This has a disadvantage that not all beads are equivalent, reducing the number of beads from which the energy can be sampled, which also means that every n_{respa}-th bead experiences a

different potential, causing it to yank on its neighbours that are constrained differently. As such, it would effectively also reduces the overall P number, and thus fail to get round the inherent limitation of discretised path integrals, that to capture the full quantum behaviour well, one needs $P > \hbar\omega/kT$.

Yet another general approach, one which might get around the bead-number limitation, for long range forces, is that of Manolopoulos[MM08]. To calculate long-range forces, this method first transforms the chain of P beads by removing all but the first p ($p < P$) of its normal modes, reducing it to simplified chain of the same length in imaginary time, but with only p beads. The energies and forces thus calculated are then transformed back to the full P-mode representation, where they can be merged with the cheap short-range forces.

In this work, we leave aside the 'bead-respa' and normal-mode-contraction ideas, for now, to see if we can improve energy convergence using better high-T (small-τ) density matrix approximations. Initially, we focussed on the dipole-limit because it is relatively simple; it requires an perturbation expansion of the quadratic simple harmonic density-matrix in terms of the dipole-dipole interaction tensor, in a similar manner to the expansion in the Dipole-Limit chapter. It is also possible to produce a similar expansion of the simple harmonic virial (and thus also the staging-virial) estimator. We test these ideas using our existing model of Xenon.

Finally, we sketch two possible approaches towards a coulomb-cancelling 'F'-density matrix by analogy to the method we already proved for DMC.

9.1 New Dipole-Expansion Density Matrix

9.1.1 Derivation of the Density Matrix

The dipole-dipole interaction between two drudes is a quadratic function.

$$V_{\rm dpl} \;=\; -\tfrac{1}{2}q^2 x\mathbf{T}x \;=\; -\tfrac{1}{2}m\omega^2\alpha x\mathbf{T}x\;.$$

This method transfers its influence from the 'perturbing' operator $V_{\rm ext}$ to the trial operator H_0. This time the Hamiltonian is decomposed into two new operators.

$$\begin{aligned}
&\hat{H} \;=\; \hat{H}_{\rm dpl} \;+\; \hat{V}_{\rm rem}\;,\\
&\text{where}\quad \hat{H}_{\rm dpl} = \tfrac{\hat{p}^2}{2m} + \tfrac{1}{2}m\omega^2x^2 + V_{\rm dpl}\;,\\
&\qquad\quad \hat{V}_{\rm rem} = V_{\rm ext}(x) - V_{\rm dpl} \qquad \text{(the remainder)}.
\end{aligned}$$

The density operator is decomposed symmetrically as before.

$$\rho(x,x';\tau) \;\;\rightarrow\;\; e^{-\frac{\tau}{2}V_{\rm rem}(x)}\;\rho_{\rm dpl}(x,x';\tau)\;e^{-\frac{\tau}{2}V_{\rm rem}(x')}\;,\quad \text{also correct to } \mathcal{O}\left(\tau^2\right).$$

In the systems we would like to investigate, there may be many dipoles coupled, and they may also be moving, making it awkward to generate an exact solution for $\rho_{\rm dpl}$. But the above operator splitting is correct only to $\mathcal{O}\left(\tau^2\right)$, so it is only necessary to make sure that $\rho_{\rm dpl}$ is correct to the same order, which in turn requires that the logarithm be correct to the same order (ultimately the logarithm is the only thing we need to manipulate).

Starting from the harmonic oscillator density matrix

$$\begin{aligned}
\log\rho_{\rm h.o.} \;&=\; \tfrac{1}{2}\mathcal{N} - \tfrac{1}{2}\mathcal{A}(x-x')^2 - \tfrac{1}{2}\mathcal{B}(x^2+x'^2) - \tau V(x)\;,\\
\text{where}\qquad
&\mathcal{A} \;=\; \tfrac{m\omega}{\hbar\sinh f} \;=\; \tfrac{mP}{\hbar^2\beta}\times\tfrac{f}{\sinh f}\;,\\
&\mathcal{B} \;=\; \tfrac{m\omega}{\hbar}\tanh\tfrac{f}{2} \;=\; \tfrac{mP}{\hbar^2\beta}\times f\tanh\tfrac{f}{2}\;,\\
&\mathcal{N} \;=\; \log\left(\mathcal{A}/2\pi\right)\;.
\end{aligned}$$

We expand it in $\lambda = -\alpha\mathbf{T}$ as before. The functions $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{N}$ above were rearranged so that all the λ dependence appears in $f = \beta\hbar\omega/P$. It is not immediately obvious that a 2nd-order expansion in λ would also be 2nd-order in τ, so that had to be checked (and fortunately, it is):

$$\omega \;\;\rightarrow\;\; \omega\left(1+\lambda\right)^{1/2}\;,\quad f\rightarrow f\left(1+\lambda\right)^{1/2}\;,\quad \text{where}\quad \lambda = -\alpha\mathbf{T}\;,$$

$$\begin{aligned}\log\rho_{\rm dpl} &= \sum_{n=0}^{\infty}\tfrac{\lambda^n}{n!}\tfrac{\partial^n}{\partial\lambda^n}\left[\log\rho_{\rm h.o.}\right] \\ &\approx \log\rho_{\rm h.o.} + \lambda\tfrac{\partial}{\partial\lambda}\left(\log\rho_{\rm h.o.}\right) + \tfrac{\lambda^2}{2}\tfrac{\partial^2}{\partial\lambda^2}\left(\log\rho_{\rm h.o.}\right)\ .\end{aligned}$$

For a system in N dimensions, note $N = \mathrm{Tr}(\delta)$.

$$\log\rho_{\rm h.o.} = \quad \tfrac{N}{2}\mathcal{N} \qquad -\tfrac{1}{2}\mathcal{A}(\mathbf{x}-\mathbf{x}')^2 \qquad -\tfrac{1}{2}\mathcal{B}(\mathbf{x}^2+\mathbf{x}'^2)\ ,$$

$$\begin{aligned}\log\rho_{\rm dpl} = \quad & \tfrac{N}{2}\mathcal{N}_0 && -\tfrac{1}{2}\mathcal{N}_1\alpha\mathrm{Tr}(\mathbf{T}) && +\tfrac{1}{2}\mathcal{N}_2\tfrac{\alpha^2}{2}\mathrm{Tr}\left(\mathbf{T}^2\right) \\ & -\tfrac{1}{2}\mathcal{A}_0(\mathbf{x}-\mathbf{x}')^2 && +\tfrac{1}{2}\mathcal{A}_1\alpha(\mathbf{x}-\mathbf{x}')\mathbf{T}(\mathbf{x}-\mathbf{x}') && -\tfrac{1}{2}\mathcal{A}_2\tfrac{\alpha^2}{2}(\mathbf{x}-\mathbf{x}')\mathbf{T}\mathbf{T}(\mathbf{x}-\mathbf{x}') \\ & -\tfrac{1}{2}\mathcal{B}_0(\mathbf{x}^2+\mathbf{x}'^2) && +\tfrac{1}{2}\mathcal{B}_1\alpha(\mathbf{x}\mathbf{T}\mathbf{x}+\mathbf{x}'\mathbf{T}\mathbf{x}') && -\tfrac{1}{2}\mathcal{B}_2\tfrac{\alpha^2}{2}(\mathbf{x}\mathbf{T}\mathbf{T}\mathbf{x}+\mathbf{x}'\mathbf{T}\mathbf{T}\mathbf{x}')\ ,\end{aligned}$$

$$\text{where}\quad \mathcal{A}_n \equiv \tfrac{\partial^n}{\partial\lambda^n}A \quad \text{and so on.}$$

Here are the expansion coefficients explicitly:

$$\begin{aligned}\mathcal{N}_0 &= \log\left(\tfrac{mP}{2\pi\hbar^2\beta}\right) + \log\left(\tfrac{f}{\sinh f}\right)\ , \\ \mathcal{N}_1 &= \tfrac{1}{2}\left(1 - f\coth f\right)\ , \\ \mathcal{N}_2 &= \tfrac{1}{4}\left(f\coth f - 2 + \tfrac{f^2}{\sinh^2 f}\right)\ , \\ \mathcal{A}_0 &= \tfrac{mP}{\hbar^2\beta}\times\tfrac{f}{\sinh f}\ , \\ \mathcal{A}_1 &= \tfrac{mP}{\hbar^2\beta}\times\tfrac{1}{2}\tfrac{f}{\sinh f}\left(1 - f\coth f\right)\ , \\ \mathcal{A}_2 &= \tfrac{mP}{\hbar^2\beta}\times -\tfrac{1}{4}\tfrac{f}{\sinh f}\left[1 + f\coth f + f^2\left(1 - 2\coth^2 f\right)\right]\ , \\ \mathcal{B}_0 &= \tfrac{mP}{\hbar^2\beta}\times f\tanh\tfrac{f}{2} \\ \mathcal{B}_1 &= \tfrac{mP}{\hbar^2\beta}\times\left[\tfrac{1}{2}f\tanh\tfrac{f}{2} + \tfrac{1}{4}f^2\left(1 - \tanh^2\tfrac{f}{2}\right)\right]\ , \\ \mathcal{B}_2 &= \tfrac{mP}{\hbar^2\beta}\times\left[-\tfrac{1}{4}f\tanh\tfrac{f}{2} + \tfrac{1}{8}f^2\left(1 - \tanh^2\tfrac{f}{2}\right) - \tfrac{1}{8}f^3\tanh\tfrac{f}{2}\left(1 - \tanh^2\tfrac{f}{2}\right)\right]\ .\end{aligned}$$

9.1.2 Dipole-Expansion P.I.M.D.

When this is used to construct a path integral, the path-density function becomes

$$\begin{aligned}
\log \rho_{\mathrm{dpl}}(\mathbf{x}) = \quad & \tfrac{NP}{2}\mathcal{N}_0 && -\tfrac{1}{2}\mathcal{A}_0 \textstyle\sum_i (x_i - x_{i+1})^2 && -\mathcal{B}_0 \textstyle\sum_i x_i^2 \\
& -\tfrac{P}{2}\mathcal{N}_1 \alpha \mathrm{Tr}(\mathbf{T}) && +(\mathcal{A}_1 + \mathcal{B}_1)\,\alpha \textstyle\sum_i x_i \mathbf{T} x_i && -\mathcal{A}_1 \alpha \textstyle\sum_i x_i \mathbf{T} x_{i+1} \\
& +\tfrac{P}{2}\mathcal{N}_2 \tfrac{\alpha^2}{2} \mathrm{Tr}\left(\mathbf{T}^2\right) && -(\mathcal{A}_2 + \mathcal{B}_2)\,\tfrac{\alpha^2}{2} \textstyle\sum_i x_i \mathbf{T}\mathbf{T} x_i && +\mathcal{A}_2 \tfrac{\alpha^2}{2} \textstyle\sum_i x_i \mathbf{T}\mathbf{T} x_{i+1} \\
\\
= \quad & \log \rho_{\mathrm{h.o.}}(\mathbf{x}) \\
& -\tfrac{P}{2}\mathcal{N}_1 \alpha \mathrm{Tr}(\mathbf{T}) && +(\mathcal{A}_1 + \mathcal{B}_1)\,\alpha \textstyle\sum_i x_i \mathbf{T} x_i && -\mathcal{A}_1 \alpha \textstyle\sum_i x_i \mathbf{T} x_{i+1} \\
& +\tfrac{P}{2}\mathcal{N}_2 \tfrac{\alpha^2}{2} \mathrm{Tr}\left(\mathbf{T}^2\right) && -(\mathcal{A}_2 + \mathcal{B}_2)\,\tfrac{\alpha^2}{2} \textstyle\sum_i x_i \mathbf{T}\mathbf{T} x_i && +\mathcal{A}_2 \tfrac{\alpha^2}{2} \textstyle\sum_i x_i \mathbf{T}\mathbf{T} x_{i+1} \,.
\end{aligned}$$

The path-weights becomes an effective potential as before.

$$\begin{aligned}
\phi_{\mathrm{eff}}(\mathbf{x}) \quad = \quad & \tfrac{NP}{2\beta}\mathcal{N}_0 && +\tfrac{1}{2\beta}\mathcal{A}_0 \textstyle\sum_i (x_i - x_{i+1})^2 && +\tfrac{1}{\beta}\mathcal{B}_0 \textstyle\sum_i x_i^2 \\
& +\tfrac{P}{2\beta}\mathcal{N}_1 \alpha \mathrm{Tr}(\mathbf{T}) && -\tfrac{1}{\beta}(\mathcal{A}_1 + \mathcal{B}_1)\,\alpha \textstyle\sum_i x_i \mathbf{T} x_i && +\tfrac{1}{\beta}\mathcal{A}_1 \alpha \textstyle\sum_i x_i \mathbf{T} x_{i+1} \\
& -\tfrac{P}{2\beta}\mathcal{N}_2 \tfrac{\alpha^2}{2} \mathrm{Tr}\left(\mathbf{T}^2\right) && +\tfrac{1}{\beta}(\mathcal{A}_2 + \mathcal{B}_2)\,\tfrac{\alpha^2}{2} \textstyle\sum_i x_i \mathbf{T}\mathbf{T} x_i && -\tfrac{1}{\beta}\mathcal{A}_2 \tfrac{\alpha^2}{2} \textstyle\sum_i x_i \mathbf{T}\mathbf{T} x_{i+1} \\
& +\tfrac{1}{P}\textstyle\sum_i \left[V_{\mathrm{rem}}(x_i) + \tfrac{q^2}{2} x_i \mathbf{T} x_i \right] \,.
\end{aligned}$$

Part of this potential is the 'internal' potential that comes to the harmonic behaviour. This part does not change at all, and is treated by the staging transformation already described. The remaining part is the 'external' potential that comes from interactions between the particles. The difference is that previously this term contained simple, diagonal potential terms. Now, with the dipole-perturbation reference density matrix, the function is smoothed with respect to the beads, and this includes the introduction of some off-diagonal terms as well.

$$\phi_{\mathrm{eff}}(\mathbf{x}) \quad = \quad \tfrac{NP}{2\beta}\mathcal{N}_0 + \tfrac{1}{2\beta}\mathcal{A}_0 \sum_i (x_i - x_{i+1})^2 + \tfrac{1}{\beta}\mathcal{B}_0 \sum_i x_i^2 - \phi_{\mathrm{ext}}(\mathbf{x}) \,,$$

$$\begin{aligned}
\phi_{\mathrm{ext}}(\mathbf{x}) \quad = \quad & \tfrac{P}{2\beta}\mathcal{N}_1 \alpha \mathrm{Tr}(\mathbf{T}) && -\tfrac{1}{\beta}(\mathcal{A}_1 + \mathcal{B}_1)\,\alpha \textstyle\sum_i x_i \mathbf{T} x_i && +\tfrac{1}{\beta}\mathcal{A}_1 \alpha \textstyle\sum_i x_i \mathbf{T} x_{i+1} \\
& -\tfrac{P}{2\beta}\mathcal{N}_2 \tfrac{\alpha^2}{2} \mathrm{Tr}\left(\mathbf{T}^2\right) && +\tfrac{1}{\beta}(\mathcal{A}_2 + \mathcal{B}_2)\,\tfrac{\alpha^2}{2} \textstyle\sum_i x_i \mathbf{T}\mathbf{T} x_i && -\tfrac{1}{\beta}\mathcal{A}_2 \tfrac{\alpha^2}{2} \textstyle\sum_i x_i \mathbf{T}\mathbf{T} x_{i+1} \\
& +\tfrac{1}{P}\textstyle\sum_i \left[V_{\mathrm{ext}}(x_i) + \tfrac{q^2}{2} x_i \mathbf{T} x_i \right] \,.
\end{aligned}$$

Beware! For small systems we often wish to include interactions with additional images (repetitions) of the box, outside of the box. A dipole interacts not just with the other dipoles in a box, but also with all their images outside the box (and its own images). This means that the dipole tensor (and anything derived from it, such as $g\mathbf{T}$ and $p\mathbf{T}$ explained below) has to be summed over all the images that are included. This makes no difference for most parts of the calculation, but it does matter when working with $\mathrm{Tr}\left(\mathbf{T}^2\right)$. We need to be careful to sum the matrix over all images first, and then perform the trace.

$$\cancel{\left[\sum_n \mathrm{Tr}\left(\mathbf{T}_n^2\right)\right]} \quad \longrightarrow \quad \mathrm{Tr}\left[\left(\sum_n \mathbf{T}_n\right)^2\right] \qquad \checkmark$$

9.1.3 Dipole-Expansion Drude-Forces

The dipole potential involves the drude coordinates x and the atomic coordinates r in two different ways.

$$\begin{aligned}\phi_{\rm dpl} &= -\tfrac{\alpha}{2}\sum_{jk} x_j\, T_{jk}\, x_k \qquad (j,k \text{ are particle indices}) \\ &= -\tfrac{\alpha}{2}\sum_{jk} \frac{1}{|r_{jk}|^5}\left[3(x_j\cdot r_{jk})(x_k\cdot r_{jk}) - r_{jk}^2(x_i\cdot x_j)\right] \ .\end{aligned}$$

The forces on the drude, and the forces on the atoms are much more different than for the simplest external potential. The drude-forces are the simplest to calculate, but they have an extra index, the bead index, which is denoted i here in order to be consistent; j,k,ℓ, etc are used to reference particles.

The forces on the drudes,

$$\begin{aligned}F^{(x)}_{{\rm ext}ij} &= -\nabla_{ij}\phi_{\rm ext}\left(\mathbf{x}\right) \\ &= +\tfrac{2}{\beta}\left(\mathcal{A}_1+\mathcal{B}_1\right)\alpha\left[\mathbf{T}x\right]_{ij} \qquad -\tfrac{1}{\beta}\mathcal{A}_1\alpha\left\{\left[\mathbf{T}x\right]_{i-1}+\left[\mathbf{T}x\right]_{i+1}\right\} \\ &\quad -\tfrac{2}{\beta}\left(\mathcal{A}_2+\mathcal{B}_2\right)\tfrac{\alpha^2}{2}\left[\mathbf{T}\mathbf{T}x\right]_{ij} \quad +\tfrac{1}{\beta}\mathcal{A}_2\tfrac{\alpha^2}{2}\left\{\left[\mathbf{T}\mathbf{T}x\right]_{i-1}+\left[\mathbf{T}\mathbf{T}x\right]_{i+1}\right\} \\ &\quad -\tfrac{1}{P}\nabla_{ij}\,V_{\rm ext}\left(x_i\right) \qquad -\tfrac{q^2}{P}\left[\mathbf{T}x\right]_{ij}\ ,\end{aligned}$$

$$\begin{aligned}\text{where}\quad &\left[\mathbf{T}x\right]_{ij} = \sum_k T_{jk}x_{ik}\ , \\ &\left[\mathbf{T}\mathbf{T}x\right]_{ij} = \sum_k T_{jk}\left[\mathbf{T}x\right]_{ik} = \sum_{k\ell} T_{jk}T_{k\ell}x_{i\ell} \\ &\text{(both are } P\times N_{\rm drude} \text{ arrays of 3-vectors).}\end{aligned}$$

9.1.4 Dipole-Expansion Atomic Forces and Pressure

The forces and pressures on the atoms are somewhat different. Define two shorthand operators: the gradient with respect to atomic coordinates,

$$\hat{g}_j \;=\; \nabla_{r,j}\;, \qquad \hat{g}_{j\gamma}T_{jk\alpha\beta} = \nabla_\gamma\nabla_\alpha\nabla_\beta\tfrac{1}{|r_{jk}|}\;,$$

$$\begin{aligned}
\hat{g}\mathrm{Tr}\left(\mathbf{T}\right) &= 0\;,\\
\hat{g}\mathrm{Tr}\left(\mathbf{T}^2\right) &= 2\textstyle\sum_{jk}T_{jk}\left(\hat{g}T_{kj}\right)\;,\\
\hat{g}_j x_i\mathbf{T}x_{i'} &= \hat{g}_j\textstyle\sum_{k\ell}x_{ik}T_{k\ell}x_{i'\ell} = \sum_{k\ell}x_{ij}\left(\hat{g}_jT_{j\ell}\right)x_{i'\ell} + x_{ik}\left(\hat{g}_jT_{kj}\right)x_{i'j}\;,
\end{aligned}$$

$$\begin{aligned}
\hat{g}_j x_i\mathbf{T}x_{i'} &= x_{ij}\left[g\mathbf{T}x\right]_{i'j} + x_{i'j}\left[g\mathbf{T}x\right]_{ij}\;,\\
\hat{g}_j x_i\mathbf{T}x_i &= 2x_{ij}\left[g\mathbf{T}x\right]_{ij}\;,\\
\hat{g}_j x_i\mathbf{T}\mathbf{T}x_{i'} &= x_{ij}\left[g\mathbf{T}\mathbf{T}x\right]_{i'j} + \left[x\mathbf{T}\right]_{ij}\left[g\mathbf{T}x\right]_{i'j} + x_{i'j}\left[g\mathbf{T}\mathbf{T}x\right]_{ij} + \left[x\mathbf{T}\right]_{i'j}\left[g\mathbf{T}x\right]_{ij}\;,\\
\hat{g}_j x_i\mathbf{T}\mathbf{T}x_i &= 2x_{ij}\left[g\mathbf{T}\mathbf{T}x\right]_{ij} + 2\left[x\mathbf{T}\right]_{ij}\left[g\mathbf{T}x\right]_{ij}\;,
\end{aligned}$$

$$\begin{aligned}
\text{where}\quad \left[g\mathbf{T}x\right]_{ij} &\equiv \sum_k\left(\hat{g}_jT_{jk}\right)x_{ki}\;,\\
\left[g\mathbf{T}\mathbf{T}x\right]_{ij} &\equiv \sum_{k\ell}\left(\hat{g}_jT_{jk}\right)T_{k\ell}x_{i\ell} \qquad \text{(arrays of } 3\times3 \text{ matrices).}
\end{aligned}$$

and the pressure operator,

$$\hat{p}_{\alpha\beta} \;=\; \textstyle\sum_j\frac{1}{\det h}r_{j,\beta}\hat{g}_{j,\alpha}\;, \qquad \hat{p}_{\alpha\beta}T_{jk} = \frac{1}{2}\left(r_{j\beta}-r_{k\beta}\right)\nabla_\alpha T_{jk}\;,$$

$$\begin{aligned}
\hat{p}\mathrm{Tr}\left(\mathbf{T}\right) &= 0\;,\\
\hat{p}\mathrm{Tr}\left(\mathbf{T}^2\right) &= 2\textstyle\sum_{jk}T_{jk}\left(\hat{p}T_{kj}\right)\;,\\
\hat{p}x_i\mathbf{T}x_{i'} &= \hat{p}\textstyle\sum_{jk}x_{ij}T_{jk}x_{i'k} = \sum_{jk}x_{ij}\hat{p}T_{jk}x_{i'k}\;,
\end{aligned}$$

$$\begin{aligned}
\hat{p}x_i\mathbf{T}x_{i'} &= \textstyle\sum_j x_{ij}\left[p\mathbf{T}x\right]_{i'j}\;,\\
\hat{p}x_i\mathbf{T}\mathbf{T}x_{i'} &= 2\left[x\mathbf{T}\right]_{ij}\left[p\mathbf{T}x\right]_{i'j}\;,
\end{aligned}$$

$$\text{where}\quad \left[p\mathbf{T}x\right]_{ij} = \sum_k\left(\hat{p}T_{jk}\right)x_{ki} \qquad \text{(array of } 3\times3\times3 \text{ matrices).}$$

The forces on the atoms,

$$\begin{aligned}
F_j^{(r)} &= -\hat{g}_j\,\phi_{\text{ext}}\left(\mathbf{x}\right) \\
&= -\tfrac{P}{2\beta}\mathcal{N}_1\alpha\,\hat{g}_j\mathrm{Tr}(\mathbf{T}) + \tfrac{P}{2\beta}\mathcal{N}_2\tfrac{\alpha^2}{2}\,\hat{g}_j\mathrm{Tr}\left(\mathbf{T}^2\right) \\
&\quad +\tfrac{2}{\beta}\left(\mathcal{A}_1+\mathcal{B}_1\right)\alpha\sum_i x_{ij}\left[g\mathbf{T}x\right]_{ij} \\
&\quad -\tfrac{1}{\beta}\mathcal{A}_1\alpha\sum_i\left(x_{i-1,j}+x_{i+1,j}\right)\left[g\mathbf{T}x\right]_{ij} \\
&\quad -\tfrac{2}{\beta}\left(\mathcal{A}_2+\mathcal{B}_2\right)\tfrac{\alpha^2}{2}\sum_i\left\{x_{ij}\left[g\mathbf{T}\mathbf{T}x\right]_{ij}+\left[x\mathbf{T}\right]_{ij}\left[g\mathbf{T}x\right]_{ij}\right\} \\
&\quad +\tfrac{1}{\beta}\mathcal{A}_2\tfrac{\alpha^2}{2}\sum_i\left\{\left(x_{i-1,j}+x_{i+1,j}\right)\left[g\mathbf{T}\mathbf{T}x\right]_{ij}+\left(x\mathbf{T}_{i-1,j}+x\mathbf{T}_{i+1,j}\right)\left[g\mathbf{T}x\right]_{ij}\right\} \\
&\quad -\tfrac{1}{P}\nabla_{ij}\,V_{\text{ext}}\left(x_i\right)-\sum_i\tfrac{q^2}{P}x_{ij}\left[g\mathbf{T}x\right]_{ij}\ .
\end{aligned}$$

The pressure estimator,

$$\begin{aligned}
\mathbf{P} &= \hat{p}\,\phi_{\text{ext}}\left(\mathbf{x}\right) \\
&= +\tfrac{P}{2\beta}\mathcal{N}_1\alpha\,\hat{p}\,\mathrm{Tr}(\mathbf{T}) - \tfrac{P}{2\beta}\mathcal{N}_2\tfrac{\alpha^2}{2}\,\hat{p}\,\mathrm{Tr}\left(\mathbf{T}^2\right) \\
&\quad -\tfrac{1}{\beta}\left(\mathcal{A}_1+\mathcal{B}_1\right)\alpha\sum_{ij}x_{ij}\left[p\mathbf{T}x\right]_{ij} \\
&\quad +\tfrac{1}{\beta}\mathcal{A}_1\alpha\sum_{ij}x_{ij}\left[p\mathbf{T}x\right]_{i+1,j} \\
&\quad +\tfrac{1}{\beta}\left(\mathcal{A}_2+\mathcal{B}_2\right)\tfrac{\alpha^2}{2}\sum_{ij}\left\{2\left[x\mathbf{T}\right]_{ij}\left[p\mathbf{T}x\right]_{ij}\right\} \\
&\quad -\tfrac{1}{\beta}\mathcal{A}_2\tfrac{\alpha^2}{2}\sum_{ij}\left\{2\left[x\mathbf{T}\right]_{ij}\left[p\mathbf{T}x\right]_{i+1,j}\right\} \\
&\quad +\tfrac{1}{P}\nabla_{ij}\,V_{\text{ext}}\left(x_i\right)+\sum_{ij}\tfrac{q^2}{2P}x_{ij}\left[p\mathbf{T}x\right]_{ij}\ .
\end{aligned}$$

These structures are efficient for computational implementation. Although some of the sums appear to be $\mathcal{O}\left(N^3\right)$, they can actually be done by passing a vector twice through a function that costs only $\mathcal{O}\left(N^2\right)$. In our code, a single 'dipole-matrix' function simultaneously performs 3 different operations (3 different outputs) on a single input vector:

$$\begin{array}{rcl} x & \rightarrow & \{ \quad \mathbf{T}x \quad , \quad (g\mathbf{T})\,x \quad , \quad (p\mathbf{T})\,x \quad \}\,, \\ \mathbf{T}x & \rightarrow & \{ \quad \mathbf{T}\mathbf{T}x \quad , \quad (g\mathbf{T})\,\mathbf{T}x \quad , \quad (p\mathbf{T})\,\mathbf{T}x \quad \}\,. \end{array}$$

For a 2nd order expansion in $\mathbf{T}$, it turns out that $p\mathbf{T}\mathbf{T}x$ is not needed, so there is an option to omit it, or any of the other operations.

9.1.5 Dipole-Expansion Energy Estimators

The Barker estimator is the simplest to derive. It can be derived two ways: by taking derivatives in β of all the terms in ϕ_{eff} above, or by taking derivatives in $\lambda = -\alpha\mathbf{T}$ from the standard Barker estimator for the harmonic oscillator.

The virial estimator involves a particular partial integration, for which there are two choices; we can choose to cancel only the harmonic cross-terms, or we can choose to cancel all the dipole-expansion cross-terms.

At present the staging estimator involves cancelling the diagonal terms from the harmonic part of the virial estimator, so there are two choices for this depending on which choice of virial integration was used.

Barker Estimator

Here is the harmonic-reference barker estimator again

$$E_{\text{B}} \;=\; \mathcal{E}_{\text{B}} + \tfrac{1}{2}\mathcal{A}_{\text{B}}\textstyle\sum_i (x_i - x_{i+1})^2 + \mathcal{B}_{\text{B}}\sum_i x_i^2 + \frac{1}{P}\sum_i V(x_i)\,,$$

$$\begin{array}{rcccl} & \mathcal{E}_{\text{B}} & = & \frac{N}{2}\hbar\omega\coth f & = \frac{NP}{2\beta} f\coth f\,, \\ & \mathcal{E}_{\text{V}} & = & 0\,, & \\ \text{with} & \mathcal{E}_{\text{St}} & = & \frac{N}{2}\hbar\omega\coth\left(\frac{\beta\hbar\omega}{2}\right)\,. & \end{array}$$

The dipole-approximation estimator can be derived by expanding in $\lambda = -\alpha\mathbf{T}$

via $f \to f\sqrt{1-\alpha\mathbf{T}}$ as before.

$$\begin{aligned} E_{\mathrm{B,dpl}}(\beta,P) &= \mathcal{E}_{\mathrm{B}0} - \mathcal{E}_{\mathrm{B}1}\alpha\mathrm{Tr}(\mathbf{T}) + \mathcal{E}_{\mathrm{B}2}\tfrac{\alpha^2}{2}\mathrm{Tr}\left(\mathbf{T}^2\right) \\ &\quad + \tfrac{1}{2}\mathcal{A}_{\mathrm{B}0}\textstyle\sum_i (x_i - x_{i+1})^2 + \mathcal{B}_{\mathrm{B}0}\sum_i x_i^2 \\ &\quad - \left(\mathcal{A}_{\mathrm{B}1} + \mathcal{B}_{\mathrm{B}1}\right)\alpha\textstyle\sum_i x_i\mathbf{T}x_i + \mathcal{A}_{\mathrm{B}1}\alpha\sum_i x_i\mathbf{T}x_{i+1} \\ &\quad + \left(\mathcal{A}_{\mathrm{B}2} + \mathcal{B}_{\mathrm{B}2}\right)\tfrac{\alpha^2}{2}\textstyle\sum_i x_i\mathbf{T}\mathbf{T}x_i - \mathcal{A}_{\mathrm{B}2}\tfrac{\alpha^2}{2}\sum_i x_i\mathbf{T}\mathbf{T}x_{i+1} \\ &\quad + \tfrac{1}{P}\textstyle\sum_i V_{\mathrm{rem}}(x_i)\ . \end{aligned}$$

Now note that the top line is an expansion of dispersion energy terms in the dipole limit. As $\mathrm{Tr}\left(\mathbf{T}^n\right)$ is the n-body term, it might seem worrying that we truncate it at the two body term, when we know that 3-body and higher interactions are important, but the whole series can be cancelled by a suitable choice of virial estimator below.

Virial Estimator

There are two choices here: to exactly cancel some of the cross terms $(x_i - x_{i+1})^2$ or to approximately cancel the full dipole expansion of these terms. It turns out that the latter is actually simplest even though at first it sounds more complicated. Like the barker extimator, it can also be found also by simply expanding the harmonic-reference virial estimator in $\lambda = -\alpha\mathbf{T}$ via $f \to f\sqrt{1-\alpha\mathbf{T}}$ as before. Here is the harmonic-reference virial estimator

$$E_{\mathrm{V}} = \mathcal{E}_{\mathrm{V}} + \mathcal{B}_{\mathrm{V}}\textstyle\sum_i x_i^2 + \frac{1}{P}\sum_i V_{\mathrm{ext}}(x_i) - \mathcal{A}_{\mathrm{V}}\frac{1}{P}\sum_i x_i\nabla_i V_{\mathrm{ext}}\ ,$$

$$\begin{aligned} E_{\mathrm{V,dpl}}(\beta,P) &= \mathcal{E}_{\mathrm{V}0} - \mathcal{E}_{\mathrm{V}1}\alpha\mathrm{Tr}(\mathbf{T}) + \mathcal{E}_{\mathrm{V}2}\tfrac{\alpha^2}{2}\mathrm{Tr}\left(\mathbf{T}^2\right) \\ &\quad + \mathcal{B}_{\mathrm{V}0}\textstyle\sum_i x_i^2 - \mathcal{A}_{\mathrm{V}0}\frac{1}{P}\sum_i x_i\nabla_i V_{\mathrm{rem}} + \frac{1}{P}\sum_i V_{\mathrm{rem}}(x_i) \\ &\quad - \mathcal{B}_{\mathrm{V}1}\alpha\textstyle\sum_i x_i\mathbf{T}x_i + \mathcal{B}_{\mathrm{V}2}\tfrac{\alpha^2}{2}\sum_i x_i\mathbf{T}\mathbf{T}x_i \\ &\quad + \mathcal{A}_{\mathrm{V}1}\alpha\tfrac{1}{P}\textstyle\sum_i \left[x\mathbf{T}\right]_i\nabla_i V_{\mathrm{rem}} - \mathcal{A}_{\mathrm{V}2}\tfrac{\alpha^2}{2}\tfrac{1}{P}\sum_i \left[x\mathbf{T}\mathbf{T}\right]_i\nabla_i V_{\mathrm{rem}}\ , \end{aligned}$$

and because $\mathcal{E}_{\rm V} = 0$, this simplifies to

$$\begin{aligned} E_{\rm V,dpl}(\beta,P) \;=\;& \mathcal{B}_{\rm V0}\textstyle\sum_i x_i^2 - \mathcal{A}_{\rm V0}\frac{1}{P}\sum_i x_i\nabla_i V_{\rm rem} + \frac{1}{P}\sum_i V_{\rm rem}(x_i) \\ & -\mathcal{B}_{\rm V1}\alpha\textstyle\sum_i x_i\mathbf{T}x_i + \mathcal{B}_{\rm V2}\frac{\alpha^2}{2}\sum_i x_i\mathbf{T}\mathbf{T}x_i \\ & +\mathcal{A}_{\rm V1}\frac{\alpha}{P}\textstyle\sum_i \left[x\mathbf{T}\right]_i\nabla_i V_{\rm rem} - \mathcal{A}_{\rm V2}\frac{\alpha^2}{2P}\sum_i \left[x\mathbf{T}\mathbf{T}\right]_i\nabla_i V_{\rm rem}\;. \end{aligned}$$

Staging Estimator

This is the harmonic-reference staging estimator

$$\begin{aligned} E_{\rm St} \;=\;& \tfrac{1}{P}\textstyle\sum_i V(x_i) - \mathcal{A}_{\rm V}\frac{1}{P}\sum_i x_i\nabla_i V \\ & + \left(\,\mathcal{E}_{\rm St} \;-\; \mathcal{B}_{\rm St}\frac{1}{P}\sum_{ii}\, x_\imath \mathrm{D}_{(i-\imath)}\nabla_i\phi_{\rm ext}\right)\;. \end{aligned}$$

One possibility is to simply follow the same procedure as for the barker and virial estimators; expanding the harmonic-reference estimator in $\mathbf{T}$. That would give

$$\begin{aligned} E_{\rm St,dpl}(\beta,P) = \;& \tfrac{1}{P}\textstyle\sum_i V_{\rm rem}(x_i) \quad +\frac{q^2}{2}\sum_i x_i\mathbf{T}x_x \\ & +N\mathcal{E}_{\rm St0} \quad -\mathcal{E}_{\rm St1}\alpha\mathrm{Tr}\left(\mathbf{T}\right) \quad +\mathcal{E}_{\rm St2}\frac{\alpha^2}{2}\mathrm{Tr}\left(\mathbf{T}^2\right) \\ & -\mathcal{A}_{\rm V0}\frac{1}{P}\textstyle\sum_i x_i\nabla_i V_{\rm rem} \qquad\qquad -\frac{1}{P}\sum_{ii}\, x_\imath\mathcal{D}_{0,(i-\imath)}\nabla_i V_{\rm rem} \\ & +\mathcal{A}_{\rm V1}\frac{\alpha}{P}\textstyle\sum_i \left[x\mathbf{T}\right]_i\nabla_i V_{\rm rem} \qquad +\frac{\alpha}{P}\sum_{ii}\, \left[x\mathbf{T}\right]_\imath\mathcal{D}_{1,(i-\imath)}\nabla_i V_{\rm rem} \\ & -\mathcal{A}_{\rm V2}\frac{\alpha^2}{2P}\textstyle\sum_i \left[x\mathbf{T}\mathbf{T}\right]_i\nabla_i V_{\rm rem} \quad -\frac{\alpha^2}{2P}\sum_{ii}\, \left[x\mathbf{T}\mathbf{T}\right] x_\imath\mathcal{D}_{2,(i-\imath)}\nabla_i V_{\rm rem}\;, \end{aligned}$$

$$\text{where} \quad \mathcal{D}_{n,(i-\imath)} \equiv \tfrac{\partial^n}{\partial\lambda^n}\left[\mathcal{B}_{\rm St}\mathrm{D}_{(i-\imath)}\right]\;.$$

However, this expansion is not used because it has two problems. Firstly, derivatives of the vector D_i are complicated to calculate and they should be avoided if possible. Secondly and more seriously, the new energy term $\mathcal{E}_{\rm St}$ is the actual energy of an unperturbed harmonic oscillator at that temperature. Thus it does not disappear with increasing P. If we expand it in $\mathbf{T}$, then we will obtain all the n-body dipole-limit dispersion terms that came out in the dipole-limit chapter. These cannot be simply discarded, or the energy estimator would be incorrect.

For these two reasons, it is preferable to try a different approach for the dipole-reference staging estimator, cancelling only the on-site, unperturbed harmonic contribution. Using exactly the same derivation as for the harmonic-reference staging estimator,

the $\sum_i x_i^2$ can be integrated out using the external effective potential.

$$\begin{aligned} E_{\mathrm{St,dpl}}(\beta,P) &= \left(\mathcal{E}_{\mathrm{St}} - \mathcal{B}_{\mathrm{St}} \tfrac{1}{P} \textstyle\sum_{i\imath} x_\imath \mathrm{D}_{(i-\imath)} \nabla_i \phi_{\mathrm{ext}} \right) \\ &\quad - \mathcal{A}_{\mathrm{V0}} \tfrac{1}{P} \textstyle\sum_i x_i \nabla_i V_{\mathrm{rem}} + \tfrac{1}{P} \textstyle\sum_i V_{\mathrm{rem}}(x_i) \\ &\quad - \mathcal{B}_{\mathrm{V1}} \alpha \textstyle\sum_i x_i \mathbf{T} x_i + \mathcal{B}_{\mathrm{V2}} \tfrac{\alpha^2}{2} \textstyle\sum_i x_i \mathbf{T}\mathbf{T} x_i \\ &\quad + \mathcal{A}_{\mathrm{V1}} \tfrac{\alpha}{P} \textstyle\sum_i \left[x\mathbf{T}\right]_i \nabla_i V_{\mathrm{rem}} - \mathcal{A}_{\mathrm{V2}} \tfrac{\alpha^2}{2P} \textstyle\sum_i \left[x\mathbf{T}\mathbf{T}\right]_i \nabla_i V_{\mathrm{rem}} \, . \end{aligned}$$

Note that there are two different potential-like terms now in the estimator, V_{rem} and ϕ_{ext}. Fortunately, this does not create any extra work, as the new terms simply use the 'external' drude-forces. Note also, that $\phi_{\mathrm{ext}} \neq V_{\mathrm{ext}}$ for the dipole-expansion case.

$$\nabla_i \phi_{\mathrm{ext}} = -F^{(x)}_{\mathrm{ext}i} \, .$$

9.1.6 Analytic Tests for the Density Matrix

It is possible to obtain the discretisation-error of the path-integral approximation very precisely for quadratic potentials in 1D. This can be extended to general dipole-limit interactions, which we can calculate by diagonalising an dipole-dipole interaction matrix, and then summing the energies of the uncoupled modes.

One method is to analytically integrate a series of coupled Gaussians. The coupling between neighbouring beads on the chain (which is caused by the kinetic energy operator) can be represented by a tridiagonal matrix, and the integration involves calculating the determinant of that matrix[SSCW81]. For a density matrix with coefficient n, a, b and an external potential with coefficient c, we simply state the result,

$$\begin{aligned} Z_P &= \prod_i^P \int_{x_i} \tfrac{n}{\sqrt{2\pi}} \exp\left\{-\tfrac{1}{2}a(x_i - x_{i+1})^2 - \tfrac{1}{2}b(x_i^2 - x_{i+1}^2) - \tfrac{1}{2}c(x_i^2 - x_{i+1}^2)\right\} \\ &= \left[\tfrac{n}{\sqrt{a}}\right]^P \tfrac{A^{P/2}}{A^P - 1} , \qquad \text{where} \quad A = \tfrac{a+(b+c)+\sqrt{2a(b+c)+(b+c)^2}}{a} , \\ F_P &= -\tfrac{1}{\beta}\log Z_P , \qquad E_P = -\tfrac{\partial}{\partial\beta}\log Z_P . \end{aligned}$$

The free-energy F_P can be calculated trivially, and we can use numerical derivatives to obtain the energy E_P. We found we needed quadruple-precision to calculate these. Here we will refer to this method as Gaussian Integration.

The other method, Numerical Matrix Multiplication[TBB83] is to numerically integrate (convolve) two spatially-discretised density matrices to form a third, by spatially-discretising the following formula: For convenience, we elected to recursively square the density matrix each time. This is to use the choice $p_1 = p_2$.

$$\begin{aligned} \int_y \rho_{p1}(x, y; p_1\tau)\, \rho_{p2}(y, x'; p_2\tau) &= \rho_{(p1+p2)}\left(x, x'; (p_1+p_2)\tau\right) , \\ \int_y \rho_P(x, y; P\tau)\, \rho_P(y, x'; P\tau) &= \rho_{2P}(x, x'; 2P\tau) , \\ E_P &= -\tfrac{1}{Z_P}\tfrac{\partial}{\partial\beta} Z_P , \end{aligned}$$

where ρ_p means the density matrix that has been formed from a p-fold discretisation of the density operator $\hat{\rho}$.

Fig. 9.1 was made via the analytical Gaussian Integration method, studying a harmonic perturbation to the potential, which is λ times as strong as that of the unperturbed potential.

For the simple-harmonic reference (triangles), the error always falls like P^{-2}. This is because of the higher-order commutators between the simple harmonic reference-operator, and the external potential, which come in at order τ^3 or P^{-3} (where $\tau = \beta/P$). But as the chain is P beads long, the total error is P times bigger, at P^{-2}.

The dipole-expansion density-matrix (circles) converges better than this because, in this ideal case, the entire external potential has been folded into the reference propagator, so there is no commutator. As the error falls like P^{-4}, we can infer that the first errors in the dipole-expansion density matrix come in at order τ^5. Unfortunately this is of no use to us unless we can improve the operator decomposition to remove the higher-order commutators. There is a decomposition that would allow us to do this [Suz86, Suz94, Chi97], but it is complicated to implement, so we have not used it at this stage. The overall error using our conventional splitting method, $\left(e^{-\tau H'/2} e^{\tau H_0} e^{-\tau H'/2}\right)^P$, should fall like P^{-2}.

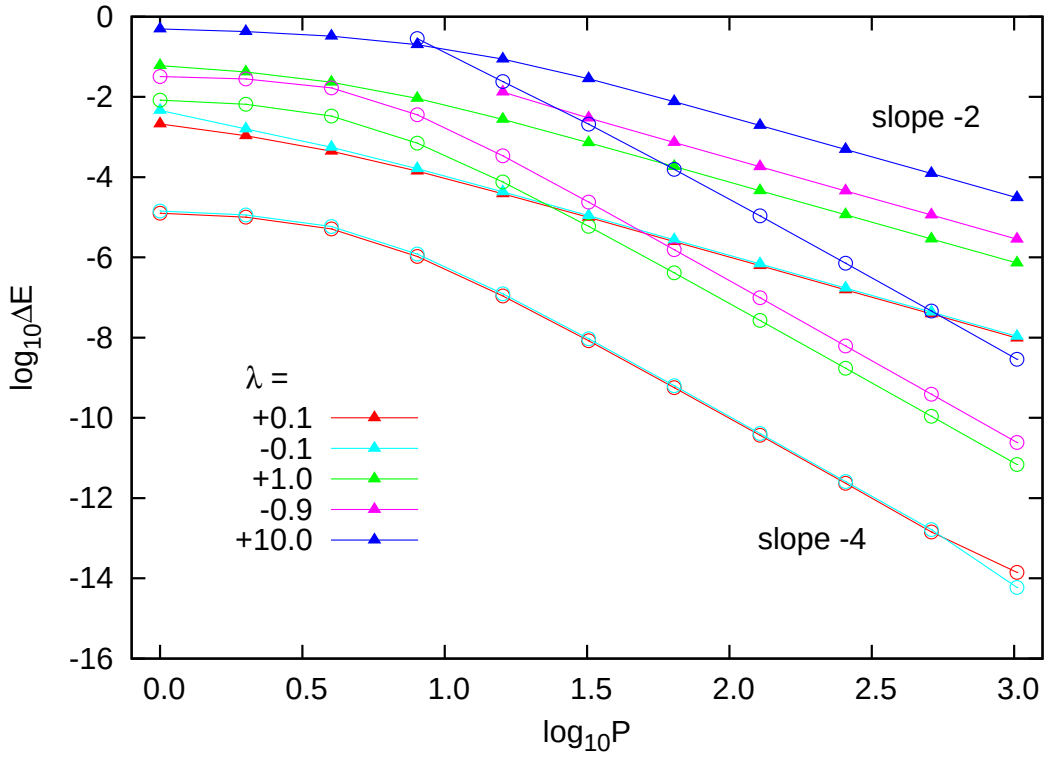


Figure 9.1: The log discretisation-error of the free energy with beads, at $\hbar\omega/kT = 10.0$, under the perturbation potential $\frac{\lambda}{2}m\omega^2x^2$, for various values of λ (see color key), using the simple-harmonic reference (triangles) and the new dipole-expansion reference (circles), but without the new dipole-expansion-virial estimator, demonstrating that the new reference is more accurate and converges as $\mathcal{O}\left(\tau^4\right)$ compared $\mathcal{O}\left(\tau^2\right)$ for the old reference.

Fig. 9.2 shows a similar graph, but this time the discretisation error in the energy calculated for a periodic 108-atom FCC crystal in the dipole-limit. It displays similar convergence, for exactly the same reasons, even though there are now 324 degrees of freedom. It was calculated by diagonalising the dipole-dipole interaction matrix, and summing the energy over each independent mode. The dipole-expansion density-matrix permits either a minimum 100-fold improvement in accuracy, or a minimum 10-fold improvement in bead number.

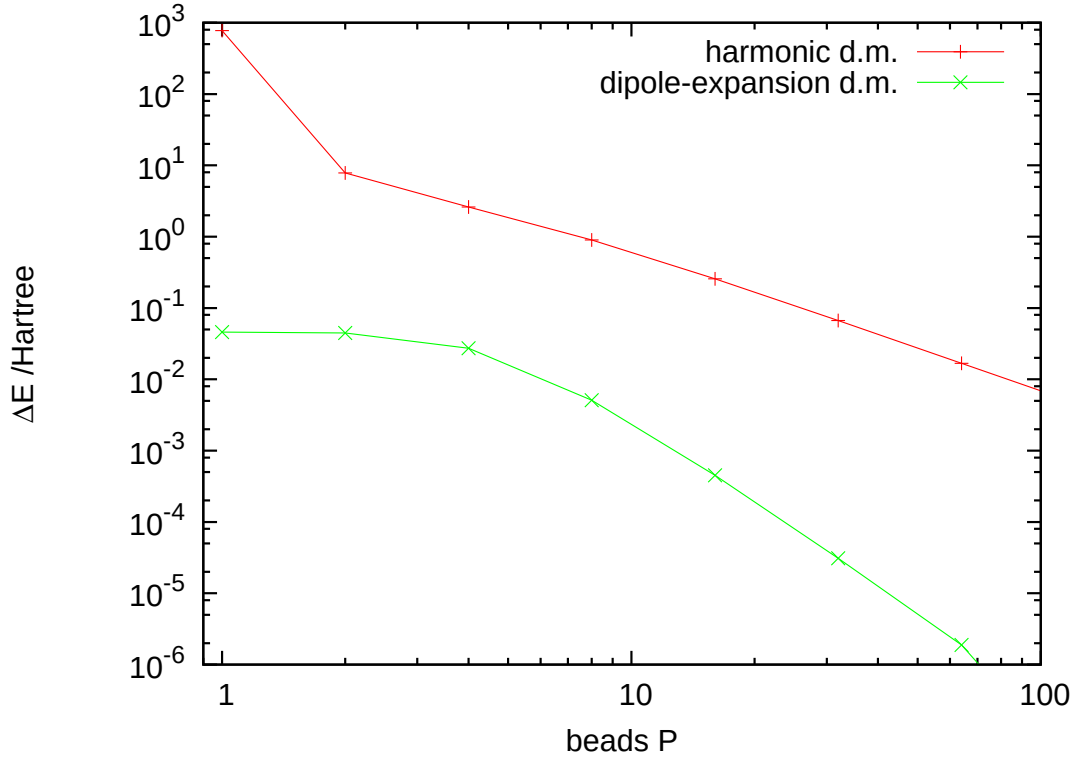


Figure 9.2: The discretisation-error of the energy with beads, at $\hbar\omega/kT = 10.0$, for the dipole-limit periodic FCC crystal, with 108 atoms, calculated using NMM, using the simple-harmonic reference (red) and the new dipole-expansion density matrix with dipole-expansion-virial estimator (green). This shows that the dipole-expansion reference remains the most accurate and quickly-converging, when applied to a complex many-body system.

Fig. 9.3 shows the convergence in terms of absolute energies, making it more obvious how much better the dipole-expansion density-matrix is. Note, however, that convergence does not really get under way until $P \approx \hbar\omega/kT = 10.0$. This is a general observation when discretising density matrices [MM08]. However, it converges very quickly after that.

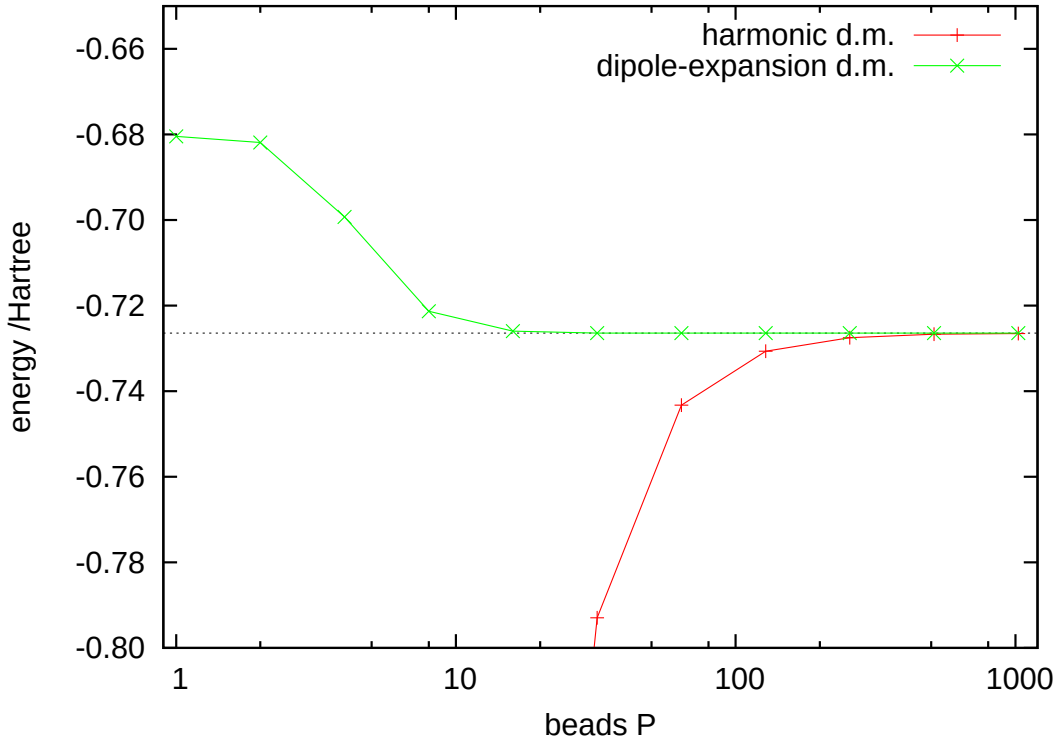


Figure 9.3: Another view of the previous result, showing the total energy with beads, at $\hbar\omega/kT = 10.0$, for the dipole-limit FCC crystal with 108 atoms (and periodic images), calculated by NMM, using the simple-harmonic reference (red) and the new dipole-expansion density matrix with dipole-expansion-virial estimator (green), This graph makes it more obvious how much the bead number can be decreased by using the dipole-expansion reference propagator for systems in the dipole-limit.

Fig. 9.4 shows how the new dipole-expansion-virial energy estimator (section 9.1.5) converges compared to the naive virial estimator (which we highlight by showing that it is equivalent to the Gaussian integration result). The convergence rate is very similar (slightly better) except from the opposite direction, showing that this estimator is viable (as good as any other). As we already showed, it can be extended trivially to the staging-virial method, but that detail is not relevant to these tests, which do not require statistical sampling.

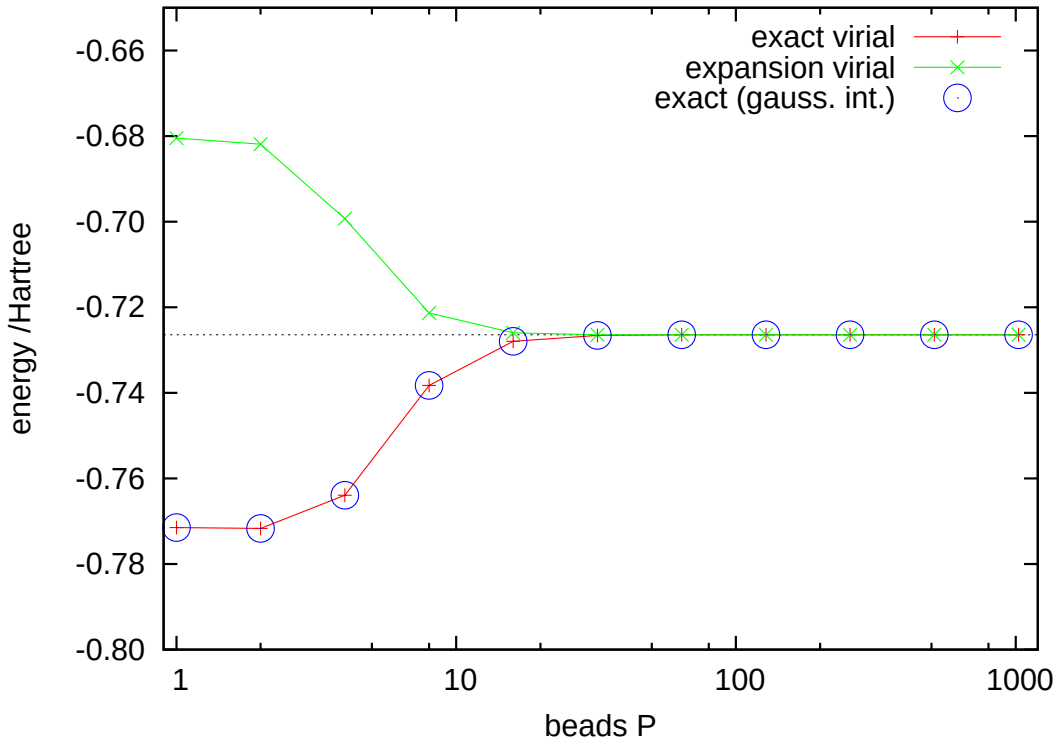


Figure 9.4: The total energy with beads, for the dipole-limit FCC crystal, at $\hbar\omega/kT = 10.0$, with 108 atoms (and periodic images), using the new dipole-expansion density-matrix, calculated using NMM. (red) is the naive virial energy estimator for this density matrix. (green) is the new dipole-expansion-virial energy estimator. (blue) is the exact Gaussian Integration result for comparison. This shows that the dipole-expansion-virial energy estimator converges no worse, and possibly better, than the problematic naive estimator.

9.2 New on-site plus Coulomb Density Matrix

When we tested the dipole-expansion density matrix for the QDO full model, we found that gave very little improvement in terms of bead number. This is not that surprising when one considers that the dipole-limit wavefunction captured only about 50% of the variational energy when compared to the harmonic wavefunction, at the energy minimum of the dimer.

The coulomb-expansion trial wavefunction described in section 6.3 was found be vastly superior to the dipole trial wavefunction. If a density matrix could be constructed along similar lines, it could perhaps lead to real improvements in convergence with bead number.

There are three overlapping approaches we can take; we can attempt to generate a coulomb correction to the harmonic oscillator density matrix by (a) solving an equation similar to eqn.6.2, (b) writing down and symmetrising the DMC propagator that is implied by the DMC operators described above, or (c) writing down the DMC trial hamiltonian that is implied by the DMC wavefunction (and thus operators).

9.2.1 Approach (a): Solving a PDE directly, or partially

For simplicity following Sec. 6.3, we begin with the base Hamiltonian,

$$H = -\frac{\hbar^2}{2m}\nabla^2 + \frac{m\omega^2}{2}r^2 - \frac{Q^2}{\mid r-R \mid} + \frac{Q^2}{\mid R \mid}\,, \tag{9.1}$$

and write an ansatz for the density matrix associated with that Hamiltonian as

$$\rho(r,r';\tau) = \rho_{\mathrm{H.O.}}(r,r';\tau)\exp[-F(r,r';\tau)]\,. \tag{9.2}$$

All the perturbation is now buried in $F(r,r';\tau)$. Inserting the ansatz into the Bloch equation yields

$$\begin{aligned}\frac{\partial F(r,r';\tau)}{\partial\tau} = &-\frac{Q^2}{\mid r-R \mid} + \frac{Q^2}{\mid R \mid} - \frac{\hbar^2}{2m}\left[\mid \nabla F(r,r';\tau) \mid^2 - \nabla^2 F(r,r';\tau)\right] \\ &-\hbar\omega\left[\frac{(r-r')}{\sinh(\tau\hbar\omega)} + \tfrac{1}{2}r\tanh\left(\frac{\tau\hbar\omega}{2}\right)\right]\cdot\nabla F(r,r';\tau)\,.\end{aligned} \tag{9.3}$$

Following the example of Sec. 6.3, we simplify the expression by neglecting the $\nabla^2 F$ term, which was approximately zero, and the $|\nabla F|^2$ term, which is related to 3-body

correlations, leaving an approximate relation,

$$\frac{\partial F(r,r';\tau)}{\partial \tau} = -\frac{Q^2}{\mid r-R \mid} + \frac{Q^2}{\mid R \mid} \\ -\hbar\omega \left[\frac{(r-r')}{\sinh(\tau\hbar\omega)} + \tfrac{1}{2} r \tanh\left(\frac{\tau\hbar\omega}{2}\right) \right] \cdot \nabla F(r,r';\tau) \; . \tag{9.4}$$

Next, we can make the following approximation

$$F(r,r';\tau) = \tfrac{1}{2}\left[\tilde{F}(r;\tau) + \tilde{F}(r';\tau) \right] \; , \tag{9.5}$$

which will give $\mathcal{O}\left(\tau^3\right)$ convergence provided that $\tilde{F}(r;\tau) \rightarrow \tau V(r)$ in the small-τ limit, so that the simple Trotter-Suzuki decomposition reappears,

$$\rho(r,r';\tau) = \rho_{\text{H.O.}}(r,r';\tau) \exp(-\frac{\tau}{2}\left[V(r)+V(r')\right]) \; , \quad \text{correct to } \mathcal{O}\left(\tau^3\right).$$

If the large-τ limit, we would like to recover the zero-temperature limit of the density matrix as a product of ground-state wavefunctions

$$\rho(r,r';\tau) \quad \rightarrow \quad \psi_0^{(HO)}(r)\psi_0^{(HO)}(r') \exp(-[F_{\text{T}}(r)+F_{\text{T}}(r')]) \exp(-\tau E_0) \; ,$$

where $F_{\text{T}}(r)$ is the function we derived in Sec. 6.3.

In order to derive an equation for $\tilde{F}(r;\tau)$, we insert the Eq. 9.5 into Eq. 9.4 and set $r = r'$ to yield

$$\frac{\partial \tilde{F}(r;\tau)}{\partial \tau} = -\frac{Q^2}{\mid r-R \mid} + \frac{Q^2}{\mid R \mid} \\ -\frac{\hbar\omega}{2} \tanh\left(\frac{\tau\hbar\omega}{2}\right) \frac{r\partial F(r;\tau)}{\partial r} \; ,$$

which can be solved using Laplace Transforms subject to the boundary condition,

$$\begin{aligned} \tilde{F}(r,0) &= \tau V(r) = 0 \; , \\ V(r) &= V(r,\theta,\phi) = -\frac{Q^2}{\mid r-R \mid} + \frac{Q^2}{\mid R \mid} \; , \end{aligned}$$

to yield the elliptic integral

$$\begin{aligned}\tilde{F}(r;\tau) &= \frac{2}{\hbar\omega}\int_{rp(\tau)}^{r} dy \frac{V(y,\theta,\phi)}{y}\left[\frac{1}{1-\left(\frac{rp(\tau)}{y}\right)^2}\right] , \\ &\text{where} \quad p(\tau) = \operatorname{sech}\left[\tau\hbar\omega/2\right] .\end{aligned}$$

It is easy to see that $\tilde{F}(r;\tau)$ behaves as specified reducing to $\tau V(r) + \mathcal{O}(\tau^3)$ at small τ and to $F_T(r)$ as τ goes to infinity ($p(\tau) \to 0$).

Since we cannot evaluate the eliptic integral in closed form, we seek approximations:

$$\begin{aligned}\tilde{F} &= -\frac{2}{\hbar\omega}\int_{rc(\tau)}^{r} \mathrm{d}r' \, \frac{\phi(r')}{r'} \quad = F_{\mathrm{DMC}}[r] - F_{\mathrm{DMC}}[rc(\tau)] , \\ \frac{\partial}{\partial\tau}\tilde{F} &= +\frac{1}{\hbar\omega}\frac{1}{c(\tau)}\frac{\partial c}{\partial\tau} \times \phi\left[rc(\tau)\right] , \\ -\hbar\omega r\frac{\partial}{\partial r}\tilde{F} &= \phi\left[r\right] - \phi\left[rc(\tau)\right] .\end{aligned}$$

If we make an approximation that $\frac{\hbar^2}{m}\nabla\log\rho_0$ is a constant equal to $-\hbar\omega r$, then we have a solution if

$$\begin{aligned}\frac{\partial}{\partial\tau}\log c &= -\hbar\omega/2 , \\ c(t) &= e^{-\tau\hbar\omega/2} .\end{aligned}$$

This solution obeys two limits nicely; for small τ, $c(t) \to (1-\tau\hbar\omega/2)$, and F converges to $-\tau\phi$:

$$\begin{aligned}\tilde{F} &\to -\frac{2}{\hbar\omega}\frac{r\tau\hbar\omega}{2} \times \left[\frac{\phi(r)}{r}\right] \\ &= -\tau\phi(r) ,\end{aligned}$$

then for large τ, $c(t) \to 0$, and F converges on the DMC ground-state trial wavefunction:

$$\begin{aligned}\tilde{F} &\to 2F_{\mathrm{DMC}}[r] - 2\cancel{F_{\mathrm{DMC}}[0]} \\ &= 2F_{\mathrm{DMC}}[r] .\end{aligned}$$

In reality, however, $\frac{\hbar^2}{m}\nabla\log\rho_0$ is not a constant, and so we would have to add correction terms, order by order.

9.2.2 Approach (b): Symmmetrise the DMC propagator

The DMC propagator is not the same as the density matrix, but we know that if the DMC propagator $P(x \to x';\tau)$ was exact, then it would be related to the density matrix as follows

$$P(x \to x';\tau) \quad = \quad \rho(x,x';\tau) \times \frac{\Psi_{\rm T}(x')}{\Psi_{\rm T}(x)} \; .$$

We can symmetrise the DMC propagator by multiplying it by the same propagator in the reverse direction, yielding an expression for the density matrix itself, in terms of the DMC propagator.

$$P(x \to x';\tau)\, P(x' \to x;\tau) = \left[\rho(x,x';\tau)\right]^2 \; ,$$

$$\rho(x,x';\tau) = \sqrt{P(x \to x';\tau)\, P(x' \to x;\tau)} \; .$$

This gives a general form,

$$\begin{aligned}
P(x \to x';\tau) \quad &= \quad \exp\left\{ -\frac{\tau}{2}\left[\frac{\hat{H}\Psi_{\rm T}}{\Psi_{\rm T}} - \bar{E} \right] \right\} \\
&\qquad \times \exp\left\{ -\frac{\left[x' - (x + \Delta(x))\right]^2}{2\sigma^2(x)} \right\} \\
&\qquad\quad \times \exp\left\{ -\frac{\tau}{2}\left[\frac{\hat{H}\Psi_{\rm T}(x')}{\Psi_{\rm T}(x')} - \bar{E} \right] \right\} \; , \\
\rho(x,x';\tau) \quad &= \quad \exp\left\{ -\frac{\tau}{2}\left[\frac{\hat{H}\Psi_{\rm T}(x)}{\Psi_{\rm T}(x)} + \frac{\hat{H}\Psi_{\rm T}(x')}{\Psi_{\rm T}(x')} - 2\bar{E} \right] \right. \\
&\qquad \left. - \frac{\left[x' - (x + \Delta(x))\right]^2}{4\sigma^2(x)} - \frac{\left[x - (x' + \Delta(x'))\right]^2}{4\sigma^2(x')} \right\} \; ,
\end{aligned}$$

where Δ and σ are functions of τ as well as position.

However, we only know the DMC propagator to 1st order in τ. That means for PIMD we have to design 2nd order in τ correction via more careful analysis, perhaps by looking at the operator decomposition.

9.2.3 Approach (c): Use DMC trial-Hamiltonian

DMC splits the Hamiltonian into two parts; $\hat{A}$, the diffuse-drift operator, and B the external potential, population operator. In the most naive splitting, with no importance sampling or trial wavefunction, the splitting is $\hat{A} = \hat{T}$ vs $\hat{B} = V$. For the simple harmonic trial wavefunction, this changes to $\hat{A} = \hat{T} + \frac{1}{2}m\omega^2x^2$, vs $\hat{B} = V - \frac{1}{2}m\omega^2x^2$. At first sight, this looks as if terms in the potential are simply being transferred across from one operator to another, but this is naive. In fact, what is happening, is that $\hat{A}$ is a fictitious- or trial-Hamiltonian which is the exact eigenoperator for a particular new trial wavefunction that we have chosen. In general, $\hat{A}$ and $\hat{B}$ are defined as follows (for now we neglect constant terms such as the normalisation $\bar{E}$).

$$\begin{aligned} \hat{A} &= \hat{T} - \frac{\hat{T}\Psi_{\mathrm{T}}}{\Psi_{\mathrm{T}}} &&= -\frac{\hbar^2}{2m}\nabla^2 + \frac{\hbar^2}{2m}\frac{\nabla^2\Psi_{\mathrm{T}}}{\Psi_{\mathrm{T}}} \,, \\ \hat{B} &= +\frac{\hat{T}\Psi_{\mathrm{T}}}{\Psi_{\mathrm{T}}} + V &&= \frac{\hat{H}\Psi_{\mathrm{T}}}{\Psi_{\mathrm{T}}} \,. \end{aligned}$$

Investigating this, may help to better understand the first two approaches, or generate new insights.

Chapter 10

Xenon - PIMD simulations

We use the same Quantum Drude model of Xenon to test the new density matrix in PIMD, primarily to test its efficacy on a realistic system, presenting convergence tests in terms of bead-number, and timesteps, for static configurations, and faux-mass for a liquid. Finally we test the new density matrix on two fluid state-points.

10.1 Dipole-Limit 'machinery tests'

The dipole limit is useful for testing our PIMD machinery, because we can obtain analytic results (or at least, in the case of NMM, numerically very accurate results). The tests in this section converged quickly because they were run with fewer beads than would be needed for an accurate simulation, and because we used a higher termperature ($\beta\hbar\omega = 5.152$ instead of $\beta\hbar\omega = 10.0$, which we use for all the simulations after this section), which also reduces the required number of beads.

We simulated a dipole-limit Xenon dimer (Fig. 10.1) for various discretisations P, and for two kinds of density matrix; the basic harmonic-reference density matrix and the dipole-expansion density matrix, for the dipole-expansion density matrix, we also used the approximate dipole-expansion-virial estimator derived above. For each of these density matrices, we calculated exact energies by integrating the estimator numerically, using Numerical Matrix Multiplication (see sec. 9.1.6) in conjunction with matrix-diagonalisation, which is trivial for the dipole-dimer case (see sec. 3.3). As the dipole-expansion-virial estimator is only approximate, the normal relation $E = -\partial/\partial\beta \log Z$ is not valid, so the Gaussian Integration method (also sec. 9.1.6) is not an option.

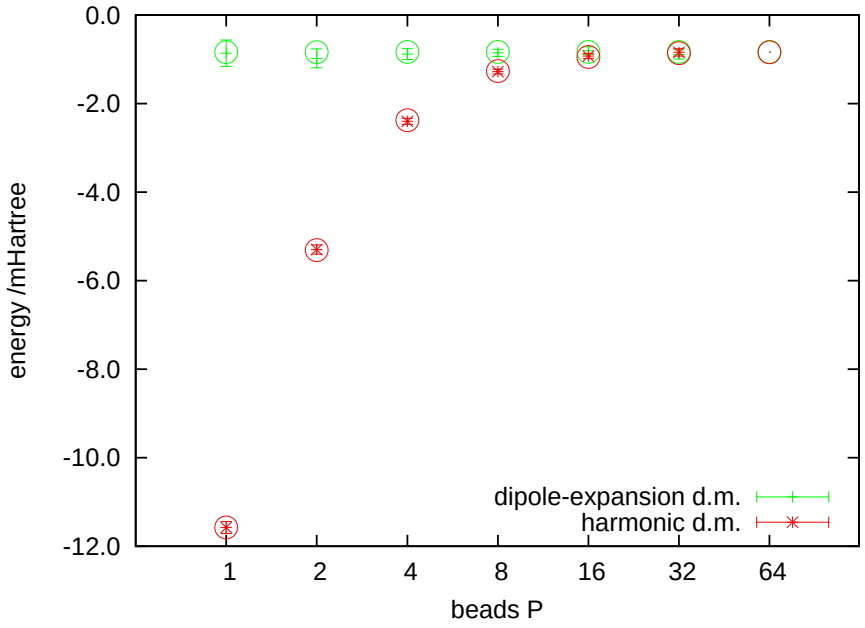


Figure 10.1: Convergence of the interaction energy with bead number P, for a static dipole-limit Xenon dimer, simulated using PIMD, using the harmonic density matrix (red) and the dipole-expansion density matrix (green), compared with values calculated by NMM (circles), with $\beta\hbar\omega = 5.152$, demonstrating that the PIMD code predicts correctly (with converged timestep and respa number).

Next, we did essentially the same thing, but for a dipole-limit Xenon periodic FCC solid, containing 32 atoms (Fig. 10.2). The only difference here was that we had to use the exact periodic sum of dipole-interaction tensors for our dipole-expansion reference propagator. This is fine for static configurations, because it is not necessary to calculate $\mathrm{Tr}(\mathbf{T}^2)$ in this case, as it appears only as a normalisation factor in the partition function, and does not appear at all in the dipole-expansion-virial estimator. However, moving the atoms would require calculating gradients of $\mathrm{Tr}(\mathbf{T}^2)$, periodically summed which is not as trivial. The reason for this choice was that it is needed for performing the matrix diagonalisation that is fed into NMM. It was necessary to compares NMM apples with PIMD apples and not oranges.

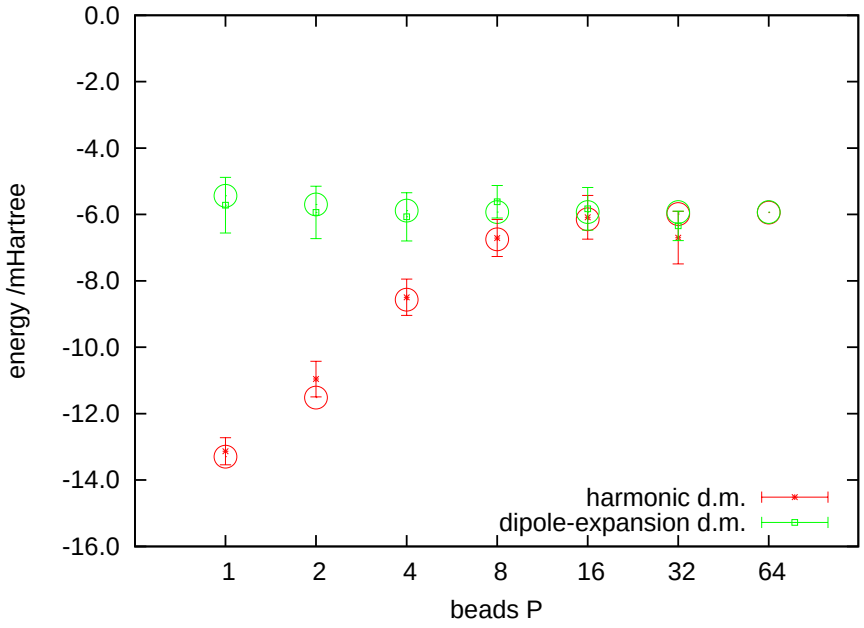


Figure 10.2: Convergence of the interaction energy with bead number P, for a static dipole-limit Xenon FCC periodic solid with 32 atoms, simulated using PIMD, using the harmonic density matrix (red) and the dipole-expansion density matrix including reciprocal-space terms (green), each compared with values calculated by NMM (circles), with $\beta\hbar\omega = 5.152$, demonstrating that the PIMD code predicts correctly (with converged timestep and respa number).

10.2 Xenon dimer (static atoms)

In figure 10.3, we present a timestep-study for the Xenon dimer, using $P = 80$ beads, as a particular tough case. Hereafter in this section, we use the converged timestep $t = 0.01/\omega_{\mathrm{PIMD}}$. For the dimer case, the potential is especially cheap, and the computational time is dominated by the internal chain dynamics, especially the thermostats. Therefore there was little to be gained from an N_{respa} study, so we omitted it, simply choosing $N_{\mathrm{respa}} = 1$.

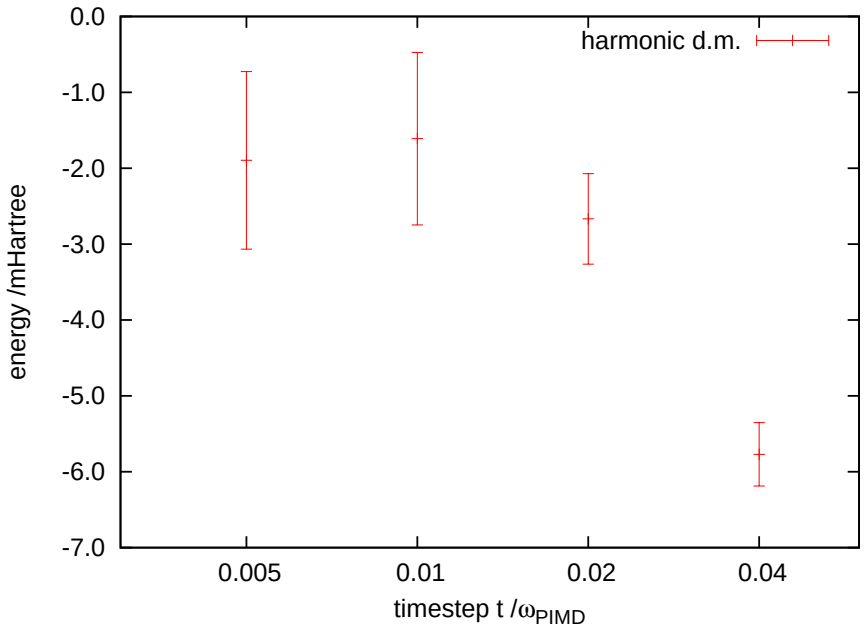


Figure 10.3: Convergence of the interaction energy with timestep, for a static Xenon dimer, simulated using PIMD, using the harmonic density matrix (red) with $P = 80$ beads, $\beta\hbar\omega = 10.0$ and separation $R = 8$bohr (near the minimum)

Next we performed two convergence-with-P studies, one with a separation of 8bohr (Fig. 10.4), near the minimum of the total interatomic potential, and the second with a separation of 12bohr (Fig. 10.5), a little further out. At 8bohr, the dipole-expansion density-matrix seems to cause slightly improved convergence, but at 12bohr, it makes a significantly greater difference, because the interaction is more strongly dominated by the dipole-dipole interaction at that range. However, in neither case is the convergence particularly good. This suggests we may need to further improve the density matrix.

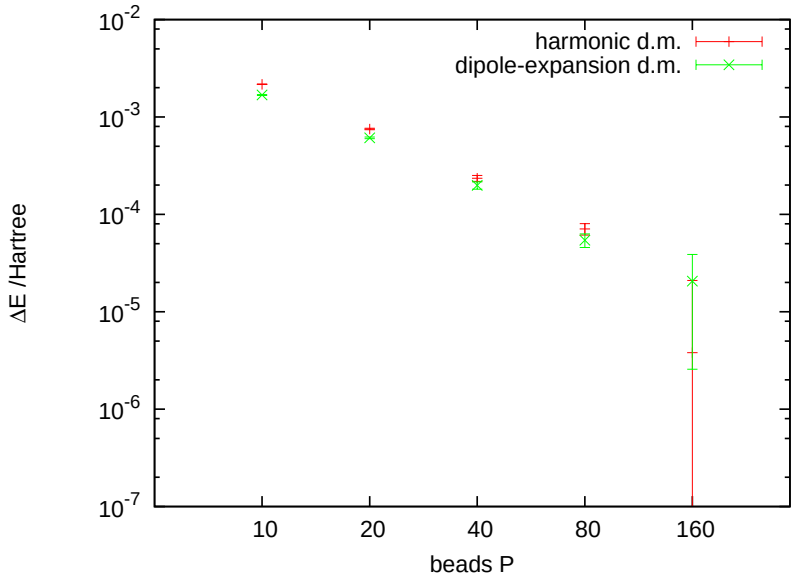


Figure 10.4: Convergence of the interaction energy error with bead number P, for a static Xenon dimer, separated by 8bohr, simulated using PIMD, using the harmonic density matrix (red) and the dipole-expansion density matrix (green).

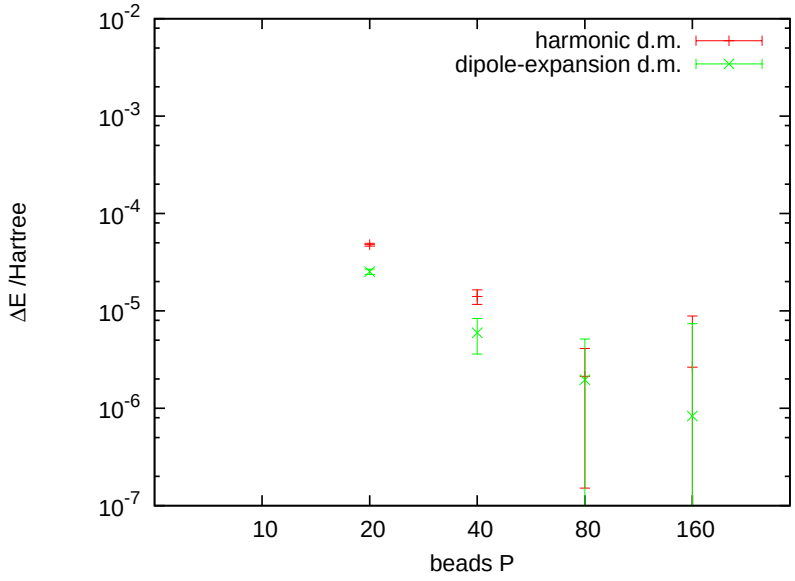


Figure 10.5: Convergence of the interaction energy error with bead number P, for a static Xenon dimer, separated by 12bohr, simulated using PIMD, using the harmonic density matrix (red) and the dipole-expansion density matrix (green).

However figure 10.6 shows that, at least when PIMD is converged with P, for example at $P = 160$, and where $kT = 1/\beta$ is small enough, it reproduces the exact ground state energy as calculated with very accurate DMC, at least within its error-bars. Therefore PIMD, is a viable way to simulate the ground-state Born-Oppenheimer surface.

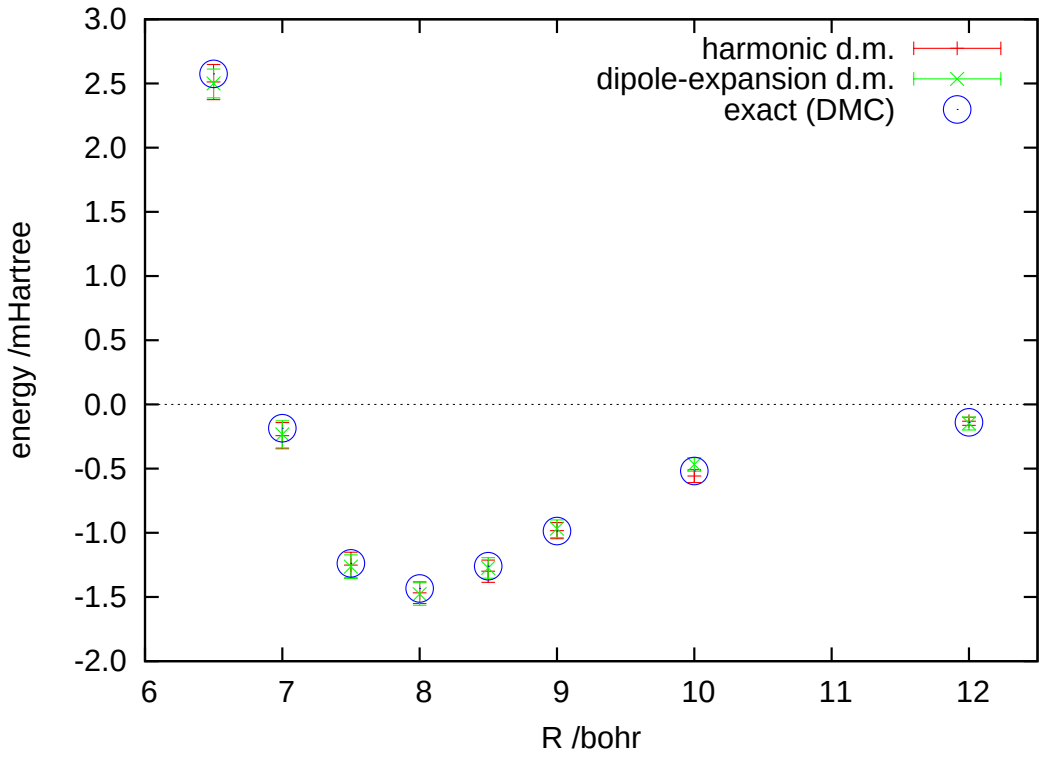


Figure 10.6: Interaction energy vs separation R, for static Xenon dimers, simulated using PIMD, $P = 160$ beads, using the harmonic density matrix (red) and the dipole-expansion density matrix (green), compared with very accurate DMC results (blue circles).

10.3 Xenon FCC solid (static atoms)

We performed some tests in order to choose parameters for the Xenon liquid simulations. We use two different timesteps; a short one for the internal dynamics of the chain (this can be thought of as a kind of intramolecular dynamics), and a long one for the external (intermolecular) forces / dynamics. We do this because the internal dynamics are relatively cheap, but the intermolecular forces tend to be more expensive. For the dimer (above), the intermolecular forces are cheap, so we simply used $N_{\text{respa}} = 1$. We perform a number of short 'internal' steps before the forces can be calculated for the 'external' step, and so the ratio of timesteps them must be a positive integer, which we denote N_{respa}.

First we performed a timestep study for the 'internal' timestep, and then chose a timestep of $t = 0.02/\omega_{\text{PIMD}}$, which seems cautious, but these forces/dynamics are not expensive so it does not matter too much.

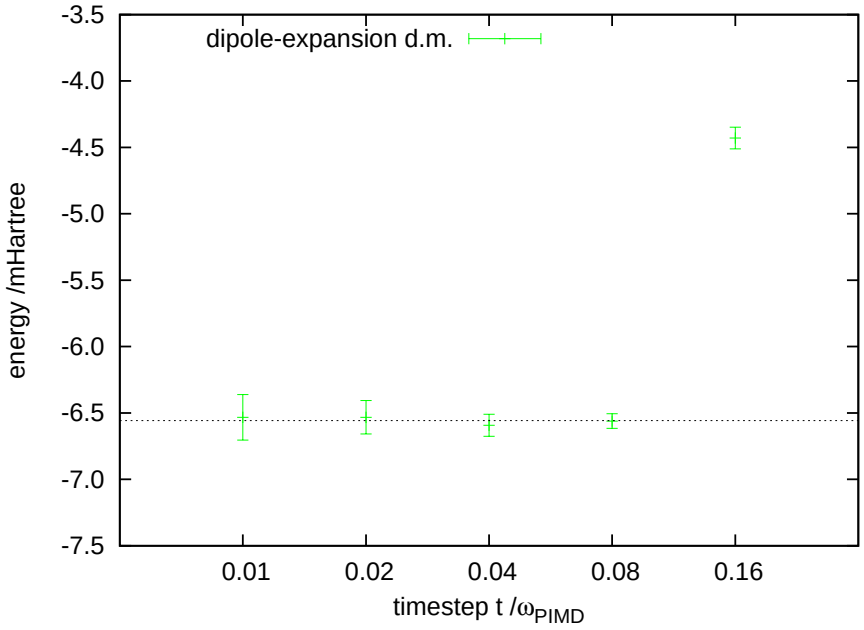


Figure 10.7: Convergence of the interaction energy per atom with timestep with $N_{\text{respa}} = 1$, for a static Xenon FCC periodic solid in the full Quantum Drude model, with 108 atoms, simulated using PIMD, using the dipole-expansion density matrix (green) with $P = 160$ beads, $\beta\hbar\omega = 10.0$. Results are compared to the ground-state energy previously calculated via DMC (dotted line).

Then we try to push up N_{respa} as far as possible in order to increase the total timestep evolved between calculations of the expensive external force. We chose $N_{\text{respa}} = 4$, giving a long timestep of $t = 0.08/\omega_{\text{PIMD}}$. Note that the scale of this graph is very much smaller and thus more detailed than the previous graph.

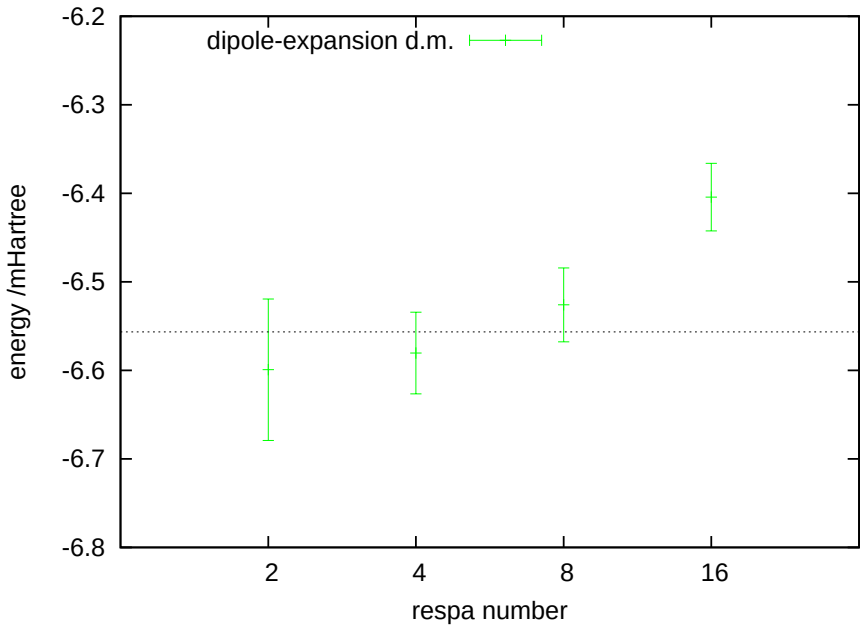


Figure 10.8: Convergence of the interaction energy per atom with N_{respa}, with short timestep $t = 0.02/\omega_{\text{PIMD}}$, for a static Xenon FCC periodic solid in the full Quantum Drude model, with 108 atoms, simulated using PIMD, using the dipole-expansion density matrix (green) with $P = 160$ beads, $\beta\hbar\omega = 10.0$ Results are compared to the ground-state energy previously calculated via DMC (dotted line).

10.4 Xenon Liquid (moving atoms)

To study the liquid, we need to be able to move the atoms. But first we need to choose a suitable faux-mass for the Quantum Drude dynamics, or equivalently by choosing a faux-frequency ω_{MD} for them. We do this using the relation $\omega_{\text{MD}} = \gamma\omega_{\text{nuc}}$, where ω_{nuc} is the characteristic frequency of the xenon atoms, estimated from the standard Lennard Jones potential. A corrolary of this is that γ captures the ratio of the LJ timestep to the PIMD timestep, and is thus a factor in the overall computational cost (alongside the bead number P). We first try $\gamma = 1$ (and found that it corresponded to a timestep of $2\Delta t_{\text{LJ}}$, and then increase γ, thus increasing the the frequency of the Quantum Drudes (and reducing the corresponding timestep to $2\Delta t_{\text{LJ}}/\gamma$), until the measurable (in this case the compressibility factor) shows convergence. When that happens it indicates that the nuclei are effectively adiabatically separated; the rate of heat flow from Quantum Drudes to the nuclei is sufficiently slow as to be negligible compared to the action of the thermostats on the atoms. Unfortunately this graph shows that, with our current thermostatting method, convergence is excruciatingly slow.

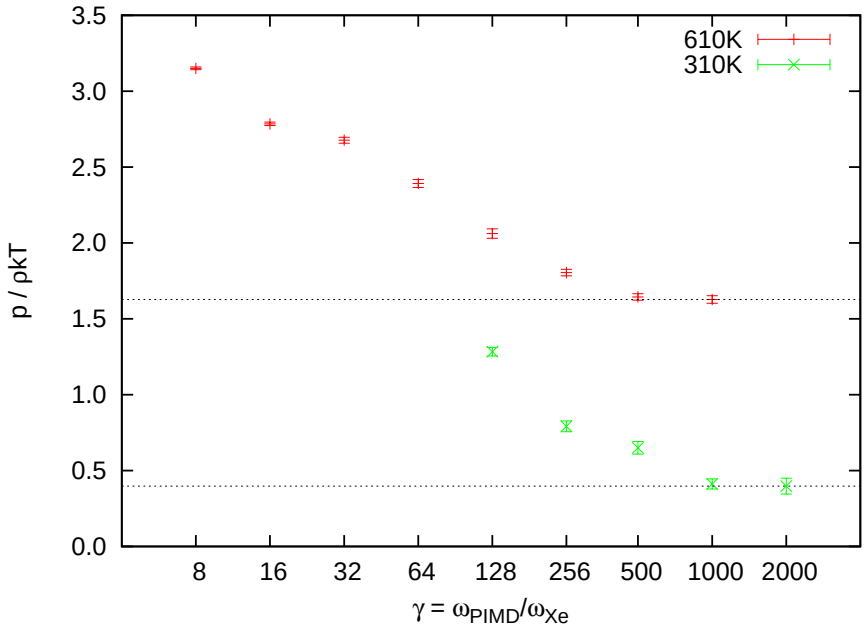


Figure 10.9: Convergence of the compressibility factor (pressure) as a function of γ, with short timestep $t = 0.02/\omega_{\text{PIMD}}$, and $N_{\text{respa}} = 4$, for a periodic Xenon liquid in the full Quantum Drude model, with 108 atoms, simulated using PIMD, using the dipole-expansion density matrix (green) with $P = 160$ beads, $\beta\hbar\omega = 10.0$

This figure explains the problem: the atom temperature is too hot. If we assume that the heat-flow in is inversely proportional to γ (see eqn. 8.2 on pg. 155), and assume that the heat-flow out is proportional to the temperature perturbation ΔT of the atoms, then once a steady state is reached, and the heat flows are balanced, we get $c_1/\gamma = c_2 \Delta T \implies \Delta T \propto 1/\gamma$. We include a fit to the data, which agrees with this reasoning at least for small perturbations (large γ).

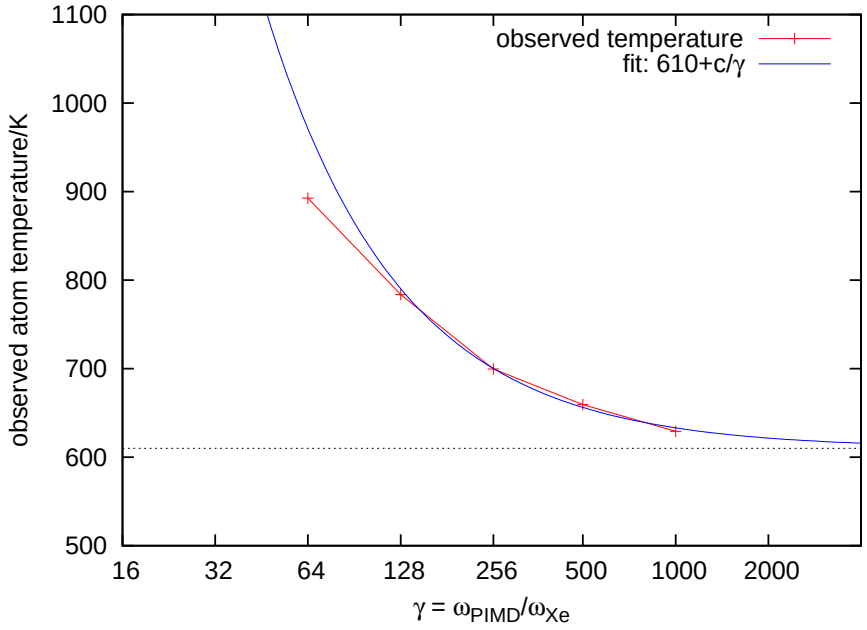


Figure 10.10: Convergence of the observed atom temperature as a function of γ, plus a fitted curve, thermostatted by Nosé Hoover Chains at T=610K, using short timestep $t = 0.02/\omega_{\mathrm{PIMD}}$, and $N_{\mathrm{respa}} = 4$, for a periodic Xenon liquid in the full Quantum Drude model, with 108 atoms, simulated using PIMD, using the dipole-expansion density matrix (green) with $P = 160$ beads, $\beta\hbar\omega = 10.0$

The problem is that Nosé Hoover Chain (NHC) thermostats are designed for equilibrium contexts, and cannot cope with non-equilibrium (non-adiabatic) conditions, even in a steady state. Ordinary Nosé Hoover Dynamics (NHD - chains of length 1) happen to be simple enough to deal with steady states of heat flow, but the whole point of NHC was to randomize and therebyy soften the rather harsh damping effect of NHD which can be non-ergodic. This suggests we need to forget about strict adiabaticity and try thermostats designed for non-equilibrium situations. We leave this for future work.

This final graph shows that the converged PIMD result did indeed come out very close to the experimental value.

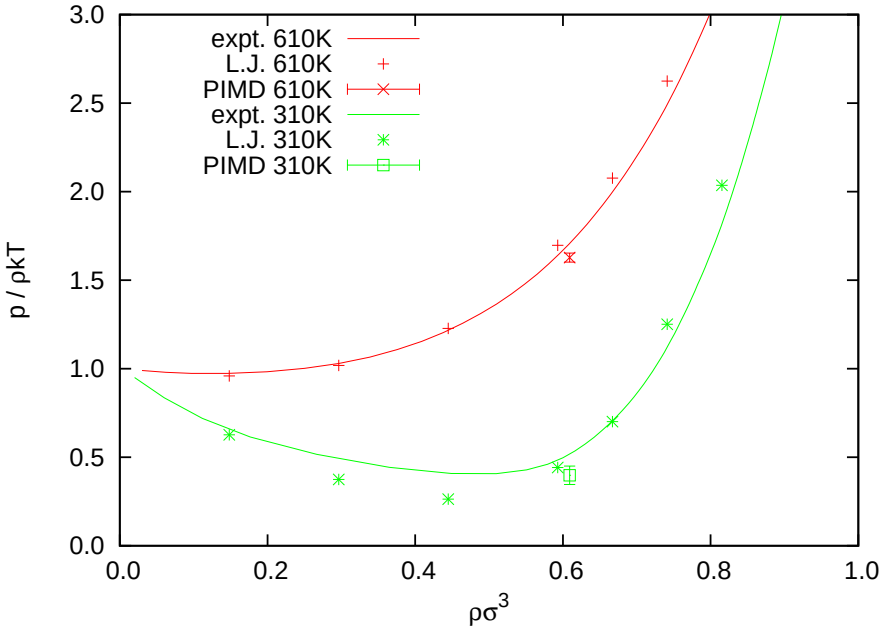


Figure 10.11: Compressibility factor as a function of density (LennardJones reduced units are used) for xenon at: 610 K (red) and 310 K (green), with short timestep $t = 0.02/\omega_{\text{PIMD}}$, and $N_{\text{respa}} = 4$, for a periodic Xenon liquid in the full Quantum Drude model, with 108 atoms, simulated using PIMD, compared with Lennard-Jones potential and experimental data from NIST. with $P = 160$ beads, $\beta\hbar\omega = 10.0$

10.5 Conclusions on the new Density Matrices

In these two chapters, we have introduced the dipole-expansion approximate density matrix and estimator, and found that it is feasible in PIMD simulation, that it also reproduces the very close approximation to the ground-state energy for static configurations of Xenon, and reproduces the same results for the liquid Xenon in the large bead number limit, compared to the simple harmonic density matrix.

We found that the new dipole-expansion density matrix is a vast improvement on the simple harmonic approximate density matrix, in the context of a dipoles-only or dipole-limit potential, but that it adds very little in the context of the full Coulomb potential, owing to the fact that higher-multipole effects are significant and converge slowly with bead number P.

This provided motivation to attempt to develop a density matrix specifically for the full Coulomb potential (with damping), and we sketched some promising approaches, but have not developed it yet to the extent that it can be coded and evaluated.

Chapter 11

Future Work: Water - a Quantum Drude Model

Water is of very wide interest from earth-sciences [CCS+99, GD09] to molecular biology, so a transferable model that accurately reproduces its detailed properties [EK69], over a wide range of states and phases, will be of interest to researchers across many fields. Therefore a major target of the Quantum Drude formalism is to create a model of water that is transferable from one context to another.

Bernal and Fowler [BF33] produced an early model of water, but Rahman and Stillinger [RS71] were the first to simulate water using the model ST2, and were followed by many others including SPC [BPvG+81], TIPS2, TIP3P, and TIP4P [JCM+83].

Sommerfeld and Jordan [SJ05], have used a limited Quantum Drude model of water with a perturbation theory approach, to simulate the polarisation of water in response to an excess electron. However, that model treated Drude-Drude interactions with a mean-field approach (classical polarisation plus dispersion in a Lennard-Jones form), and thereby neglected much of the many-body character of the Quantum interaction (some many-body character but only those interactions mediated by the electron itself). Furthermore, they fitted their model only to the dipole polarisability of water. Quantum Drudes, however, can be fit to best reproduce also the quadrupole polarisability, octopole polarisability, and C_6, C_8, and C_{10} dispersion coefficients. A consequence of this could be that the spatial distribution of their Quantum Drude oscillator was sub-optimal, and that in turn could mean that the effects of spatial confinement in clusters or condensed matter (that is, reduced effective polarisability) were incorrectly modelled.

Here we present the general outline for a new model of water based on a recent TIP4P model [AV05], augmented by a Quantum Drude to capture many-body dispersion and polarisation. We present the Quantum Drude parameters that best reproduce the polarisability properties of water (all three of the isotropic dipole-, quadrupole- and octopole-polarisabilities) and then compare its dispersion properties with those of the real model. We also sketch how the rest of the model will be fitted, including the short range repulsion, and how to generalise the use of damping terms from the Xenon model.

11.1 The Base Model

The Water molecule has a permanent or mean charge distribution as well as a fluctuating or polarisable part. We can base a new model of water on an existing model, for example TIP4P.

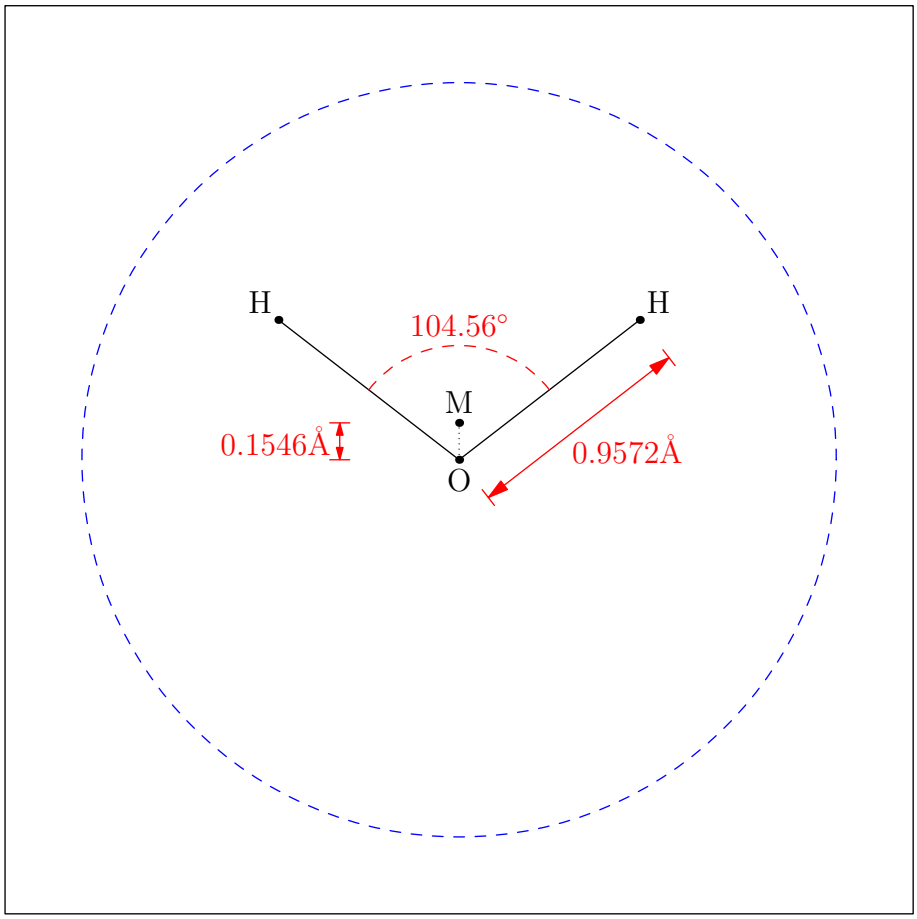


Figure 11.1: TIP4P model of water [AV05]. Blue represents the size of the molecule (half the Van-der-Waals radius)

This has a charge q_{M} placed at the M-site, and a charge q_{H} on the two hydrogen (H) sites. A water molecule is charge-neutral, fixing $q_{\mathrm{M}} = -2q_{\mathrm{H}}$. The oxygen (O) site interact only with the O-site in other molecules, as the centre for a repulsive potential V_{rep}, whereas the other sites interact via coulomb potentials. The H-M potential is attractive, and can get very close (a typical H-O distance is roughly 1.7Å), so it needs to be damped to prevent unphysical 'yanking', but the others, H-H and M-M are repulsive and, because they are heavy sites that behave classically, there is no need to

damp them.

$$\begin{aligned} V_{\text{H-H}} &= q_{\text{H}} q_{\text{H}} / \left| r_{\text{H-H}} \right| \ , \\ V_{\text{M-M}} &= q_{\text{M}} q_{\text{H}} / \left| r_{\text{M-H}} \right| \ , \\ V_{\text{H-M}} &= q_{\text{M}} q_{\text{M}} \operatorname{erf}\left(\alpha_H r_{\text{M-M}}\right) / \left| r_{\text{M-M}} \right| \ , \\ V_{\text{O-O}} &= V_{\text{rep}}\left(r_{\text{O-O}}\right) \qquad \text{(empirical repulsive potential to be fitted).} \end{aligned}$$

11.2 Adding a Drude

The Xenon model (previous chapter) has a mobile Drude with charge q_{D} whose harmonic potential is centred on a drude-centre with cancelling charge q_{d}. To simplify model building, we can start from a TIP4P (or other) potential as a framework for underlying charge distribution, and add a neutral (Drude)-(drude-centre) pair to capture the polarisation and dispersion behaviour. In the case of TIP4P we have a choice to add it to the O-site or the M-site.

11.2.1 Fitting Drude Parameters

Because for Xenon there are no permanent fields, the fitting focussed on getting the dispersion coefficients correct, rather than the polarisabilities, so the model was fitted using exact rules for α_1, C_6 and C_8.

For water, we care about polarisabilities just as much, if not more than dispersion coefficients, so we would like to find a best fit to all of α_1, α_2, α_3, C_6, C_8, C_{10}. However, we have only 3 unknowns to fit, so there will not be an exact solution in general, so one option is to minimise some fitting function, and weight the relative importance of each parameter.

Glenn Martyna and Troy Whitfield fitted Drude parameters for water using the following fitting function [unpublished work].

$$\begin{aligned} & f(q,m,\omega) \ = \\ & k_1 \left(\frac{\alpha_1(q,m,\omega)}{\alpha_{1\text{expt}}} - 1 \right)^2 + k_2 \left(\frac{\alpha_2(q,m,\omega)}{\alpha_{2\text{expt}}} - 1 \right)^2 + k_3 \left(\frac{\alpha_3(q,m,\omega)}{\alpha_{3\text{expt}}} - 1 \right)^2 \\ & + k_4 \left(\frac{C_6(q,m,\omega)}{C_{6\text{expt}}} - 1 \right)^2 + k_5 \left(\frac{C_8(q,m,\omega)}{C_{8\text{expt}}} - 1 \right)^2 + k_6 \left(\frac{C_{10}(q,m,\omega)}{C_{10\text{expt}}} - 1 \right)^2 \ , \end{aligned}$$

where $k_{1\text{-}6}$ are weight factors (a high value results in smaller deviation), which can be set to tighten particular parameters in preference to others, and where we treat the

following relations as functions of q, m, and ω.

$$\alpha_1 = \frac{q^2}{m\omega^2} \; , \qquad C_6 = \tfrac{3}{4}\alpha_1\alpha_1\hbar\omega \; ,$$
$$\alpha_2 = \tfrac{3}{4}\left(\tfrac{\hbar}{m\omega}\right)\alpha_1 \; , \qquad C_8 = 5\left(\tfrac{\hbar}{m\omega}\right) \times C_6 \; ,$$
$$\alpha_3 = \tfrac{5}{4}\left(\tfrac{\hbar}{m\omega}\right)^2\alpha_1 \; , \qquad C_{10} = \tfrac{245}{8}\left(\tfrac{\hbar}{m\omega}\right)^2 \times C_6 \; .$$

Another option is to follow a direct fitting procedure derived from these relations,

$$(1) \quad \omega = \frac{1}{\hbar}\frac{4C_6}{3\alpha_1^2} \; ,$$
$$(2) \quad m = \frac{\hbar}{\omega}\frac{3\alpha_1}{4\alpha_2} \; ,$$
$$(3) \quad q = \pm\sqrt{m\omega^2\alpha_1} \; .$$

Here we have the results for the direct parameterisation.

q	1.063
m	0.2881
ω	0.6287

Property	Value	Target	Dev.
α_1	9.92	9.92	0.0%
α_2	41.08	41.08	0.0%
α_3	377.96	377.00	0.3%
C_6	46.40	46.40	0.0%
C_8	1280.86	1141.70	12.2%
C_{10}	43313.24	32400.00	33.7%

Table 11.1: Drude parameters which have been fitted to best reflect the polarisability properties of the water molecule

11.2.2 Damping of Coulomb Interactions

For the Xenon model, there is only one kind of classical particle that carries charge, the 'nucleus' at the drude-centre that carries a charge opposite to that of the Drude particle.

For water, however, there are now potentially 3 different kinds of charge centres with which the Drude particle may interact, and each coulomb interaction will need to be damped to prevent the Drude from getting close to the centre of a potential well.

$$\begin{array}{cccc} V_{\mathrm{DD}} & V_{\mathrm{Dd}} & V_{\mathrm{dd}} & \qquad\qquad \gamma_{\mathrm{DD}} \quad \gamma_{\mathrm{Dd}} \quad \gamma_{\mathrm{dd}} \\ & \downarrow & \downarrow & \qquad\qquad\qquad\quad \downarrow \quad \downarrow \\ (\text{unique}) & V_{\mathrm{DH}} & V_{\mathrm{dH}} & \qquad\qquad\qquad\quad \gamma_{\mathrm{DH}} \quad \gamma_{\mathrm{dH}} \\ & V_{\mathrm{DM}} & V_{\mathrm{dM}} & \qquad\qquad\qquad\quad \gamma_{\mathrm{DM}} \quad \gamma_{\mathrm{dM}} \end{array}$$

$$\text{where} \qquad V_{\mathrm{yz}}(r) = \frac{q_{\mathrm{y}} q_{\mathrm{z}} \left[1 - \exp\left(-\gamma_{\mathrm{yz}}^4 r^4\right)\right]}{|r|} \ .$$

As a first estimate, we can assume $\gamma_{\mathrm{DM}} = \gamma_{\mathrm{DH}} = \gamma_{\mathrm{Dd}}$ and $\gamma_{\mathrm{dM}} = \gamma_{\mathrm{dH}} = \gamma_{\mathrm{dd}}$. Next, we can estimate these values by scaling the lengths from the Xenon model. The lengthscales can be compared in different ways. Most obvious is to look at the ratio of the Lennard-Jones length parameter, σ_{LJ}. Another option is to look at the characteristic lengthscale defined by the Drudes themselves, as they have been fitted independently of such considerations. Either way, because γ is an inverse length, we would get

$$\gamma_{\mathrm{H_2O}} = \frac{\sigma_{\mathrm{Xe}}}{\sigma_{\mathrm{H_2O}}} \times \gamma_{\mathrm{Xe}} \ .$$

From this initial estimate, the parameters could be tuned, alongside or before fitting the O-O repulsion potential below. In particular, it might be necessary to change the damping on the D-H attraction, as with the M-H damping previously, because the Drude particle can get even closer. In that case, we can vary the parameters γ_{DH} and γ_{dH} together, keeping their ratio the same.

11.2.3 Fine-tuning, and O-O repulsion

To add the short range repulsion and to fine-tune the model, we can begin by looking at the energies of different water-dimer geometries, with the energies calculated from high-accuracy quantum simulations in or parameterised potentials, already calculated in this way, such as that of Xantheas [BLXL99]. Meanwhile the geometry configurations themselves can be sampled from simulations of liquid water and ice that use simpler classical models of the water molecule (fitted to bulk behaviour), which will help ensure

that the samples are relevant to the configurations that actually exist in condensed-phase water. We can also perform further fine-tuning by examining the properties of water in the various phases of ice.

Chapter 12

Conclusions

The Quantum Drude model has been characterised with analytical theory, showing that coefficients for the polarisability and dispersion asymptotic series, agree with experiment. Fast simulation methods that are specifically suited to Quantum Drudes have been developed. The Quantum Drude model has been applied to a realistic system, Xenon, which is difficult to model because its long-range interactions are pure dispersion, obtaining very good results also in agreement with experiment. A path to the future has been outlined: the use of Quantum Drude models to treat water.

Using analytic theory, we showed that Quantum Drudes reproduce the properties of many real atoms and molecules fairly accurately, including polarisabilities, two-body and three-body dispersion coefficients. Through matrix diagonalisation (in the dipole-limit) and diagrammatic expansions, we showed Quantum Drudes can be expected to capture many-body, multipole polarisation, as well as many-body, multipole, mixed-species dispersion, and everything in between, in a realistic way. Although the Quantum Drude model does not treat repulsion inherently, we can simply add two-body repulsion corrections (or even many-body repulsion, should the need arise), in line with what is standard practice for the current state of the art. Therefore, Quantum Drudes are very promising for building high-accuracy force-fields for biomolecular simulation, and we have laid out the steps required to build such models.

We have made progress in developing high-accuracy methods for simulating Quantum Drudes with DMC, including improved propagators and improved population-conserving methods and the approximate 'Coulomb' trial wavefunction, which is a step beyond multipole-expansions. In the process, we have gained a better understanding of the wavefunctions, including the basic 'Coulomb' perturbation on one hand, and on the other hand the many-body multipole recursive expansion which inspired the diagrammatic expansion mentioned above. Some of these developments have in turn suggested new avenues for PIMD density matrices. Even though we have not yet completely solved the problem, we have also made good headway into improving the Path Integral discretisation, through the dipole-expansion density matrices, as a proof-of-concept, and our sketches towards a 'Coulomb'-reference density matrix.

Through DMC simulations and PIMD simulations, we have demonstrated that the Quantum Drude works, using Xenon as test model. The Quantum Drude model of Xenon is accurate in the condensed phase even though it was fit in the gas-phase, due to its ability to capture many-body dispersion. This demonstrates that it is possible to produce truly transferable potentials using the Quantum Drude formalism.

Finally, we sketched a model of water, which we believe will be of wide interest once it is finished and we begin simulating it. It will then be possible to extend the force-fields of proteins and other biomolecules, leading to a very open-ended research path. Thus we have made a good deal of progress, and there is plenty more still to be done.

Bibliography

[AAA+10] Jeremy J. Agrestia, Eugene Antipov, Adam R. Abate, Keunho Ahn, Amy C. Rowat, Jean-Christophe Baret, Manuel Marquez, Alexander M. Klibanov, Andrew D. Griffiths, and David A. Weitz. Ultrahigh-throughput screening in drop-based microfluidics for directed evolution. *PNAS*, 107:4004–4009, (2010).

[AdP02] Peter Atkins and Julio de Paula. *Atkins' Physical Chemistry*. Oxford University Press, (2002).

[AM03a] Andrés Aguado and Paul A. Madden. Ewald summation of electrostatic multipole interactions up to the quadrupolar level. *J.Chem.Phys.*, 119(14):7471, October (2003).

[AM03b] Andrés Aguado and Paul A. Madden. Oxide potentials from ab initio molecular dynamics: An assessment of their transferability. *J.Chem.Phys.*, 118:5718, (2003).

[AMK00] R. Assaraf, M.Caffarel, and A. Khelif. Diffusion monte carlo methods with a fixed number of walkers. *Phys.Rev.E*, 61:45664575, (2000).

[And75] James B. Anderson. A random-walk simulation of the schroedinger equation: H_3^+. *J.Chem.Phys.*, 63:1499, (1975).

[Asb05] Charles L Asbury. Kinesin: world's tiniest biped. *Current Opinion in Cell Biology*, 17:89–97, (2005).

[AT43] B. M. Axilrod and E. Teller. Interaction of the van der waals type between three atoms. *J.Chem.Phys.*, 11:299, (1943).

[AV05] J. L. F. Abascal and C. Vega. A general purpose model for the condensed phases of water: Tip4p/2005. *J.Chem.Phys.*, 123:234505, (2005).

[Axe04] D.D. Axe. Estimating the prevalence of protein sequences adopting functional enzyme folds. *J.Mol.Biol.*, 341:1295–1315, (2004).

[Bar79] J. A. Barker. A quantum-statistical Monte Carlo method; path integrals with boundary conditions. *J.Chem.Phys.*, 70:2914, (1979).

[Bar93] J.A. Barker. Surface tension and atomic interactions in simple liquids. *Molecular Physics*, 80:815–820, (1993).

[BB83] M.V. Bobetic and J.A. Barker. Solid state properites of Ar, Kr, Xe near 0K. *Phys. Rev. B*, 28:7317, (1983).

[BBL+95] James W. Bryson, Stephen F. Betz, Helen S. Lu, Daniel J. Suich, Hongxing X. Zhou, Karyn T. O'Neil, and William F. DeGrado. Protein design: A hierarchic approach. *Science*, 270:935–941, (1995).

[BBL08] Z.D. Blount, C.Z. Borland, and R.E. Lenski. Historical contingency and the evolution of a key innovation in an experimental population of escherichia coli. *PNAS*, 105:7899–7906, 2008.

[BBO+83] Bernard R. Brooks, Robert E. Bruccoleri, Barry D. Olafson, David J. States, S. Swaminathan, and Martin Karplus. CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *J.Comp.Chem.*, 4:187217, (1983).

[BD55] T.H.K. Barron and C. Domb. On the cubic and hexagonal close-packed lattices. *Proc. Royal Soc. A*, 227:447–465, (1955).

[Bel70] R.J. Bell. Multipolar expansion for the non-additive third-order interaction energy of three atoms. *J.Phys.B: Atomic and Molecular Physics*, 3:751, 1970.

[BF33] J.D. Bernal and R.H. Fowler. A theory of water and ionic solution, with particular reference to hydrogen and hydroxyl ions. *J.Chem.Phys.*, 1:515–548, (1933).

[BGBS05] Daniel N. Bolon, Robert A. Grant, Tania A. Baker, and Robert T. Sauer. Specificity versus stability in computational protein design. *PNAS*, pages 12724–12729, (2005).

[BK57] W. L. Bade and John G. Kirkwood. Drude-model calculation of dispersion forces. ii. the linear lattice. *J.Chem.Phys.*, 27:1284, (1957).

[BLXL99] Christian J. Burnham, Jichen Li, Sotiris S. Xantheas, and Maurice Leslie. The parametrization of a thole-type all-atom polarizable water model from first principles and its application to the study of water clusters (n=221) and the phonon spectrum of ice-ih. *J.Chem.Phys.*, 110:4566, (1999).

[BNR67] R.M. Bookchin, R.L. Nagel, and H.M. Ranney. Structure and properties of hemoglobin c-harlem, a human hemoglobin variant with amino acid substitutions in 2 residues of the beta-polypeptide chain. *J.Biol.Chem.*, 242:248–255, (1967).

[BO27] M. Born and J. R. Oppenheimer. Zur quantentheorie der molekeln. *Ann. Physik*, 84:457, (1927).

[BPvG+81] H.J.C. Berendsen, J.P.M. Postma, W.F. van Gunsteren, , and J. Hermans. Interaction models for water in relation to protein hydration. In *Intermolecular Forces*, page 331342. Reidel, (1981).

[BS98] M.C. Buonaura and S. Sorella. Numerical study of the two-dimensional heisenberg model using a green function monte carlo technique with a fixed number of walkers. *Phys.Rev.B*, 57:11446, (1998).

[BT91] Carl Branden and John Tooze. *Introduction to Protein Structure*. Garland Publishing Inc., (1991).

[BU70] A. D. Buckingham and B. D. Utting. Intermolecular forces. *Annu. Rev. Phys. Chem.*, (1970).

[Buc67] A.D. Buckingham. *Intermolecular Forces*. Halsted Press, (1967).

[BWL+74] J. A. Barker, R. O. Watts, Jong K. Lee, T. P. Schafer, and Y. T. Lee. Interatomic potentials for krypton and xenon. *J.Chem.Phys.*, 61:3081, (1974).

[BWZ09] Bruce J. Berne, John D. Weeks, and Ruhong Zhou. Dewetting and hydrophobic interaction in physical and biological systems. *Annu.Rev.Phys.Chem.*, 60:85–103, (2009).

[BXJ98] Enrique R. Batista, Sotiris S. Xantheas, and Hannes Jónsson. Molecular multipole moments of water molecules in ice ih. *J.Chem.Phys.*, 109:4546, (1998).

[CB92a] J. Cao and B.J. Berne. Many-body dispersion forces of polarizable clusters and liquids. *J.Chem.Phys.*, 97(11):8628, December (1992).

[CB92b] J. Cao and B.J. Berne. A new quantum propagator for hard sphere and cavity systems. *J.Chem.Phys.*, 97:2382, (1992).

[CCB+95] W.D. Cornell, P. Cieplak, C.I. Bayly, I.R. Gould, K.M. Merz, D.M. Ferguson, D.C. Spellmeyer, T. Fox, J.W. Caldwell, and P.A. Kollman. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.*, 117:51795197, (1995).

[CCS+99] C. Cavazzoni, G. L. Chiarotti, S. Scandolo, E. Tosatti, M. Bernasconi, and M. Parrinello. Superionic and Metallic States of Water and Ammonia at Giant Planet Conditions. *Science*, 283:44–46, (1999).

[CEO92] N J Cordova, B Ermentrout, and G F Oster. Dynamics of single-motor molecules: the thermal ratchet model. *PNAS*, 89:339–343, (1992).

[Cep95] D.M. Ceperley. Path integrals in the theory of condensed helium. *Rev.Mod.Phys.*, 67:279355, (1995).

[Cha02] C. Chakravarty. Path integral simulations of quantum lennard-jones solids. *J. Chem. Phys.*, 116:8938, (2002).

[Chi97] S.A. Chin. Symplectic integrators from composite operator factorizations. *Phys.Lett.A*, 226:344–348, (1997).

[CIKP03] B. Chen, I. Ivanov, M.L. Klein, and M. Parrinello. Hydrogen bonding in water. *Phys.Rev.Lett.*, 91:215503, (2003).

[CM09] Jiahao Chen and Todd J. Martinez. The dissociation catastrophe in fluctuating-charge models and its implications for the concept of atomic electronegativity. In *Advances in the Theory of Atomic and Molecular Systems*, pages 397–415. Springer Netherlands, (2009).

[Coc71] W. Cochran. Lattice dynamics of ionic and covalent crystals. *Critical Reviews in Solid State and Materials Sciences*, 2:1–44, (1971).

[dB48] J. de Boer. Quantum theory of condensed permanent gases I the law of corresponding states. *Physica*, 14:139–148, (1948).

[DCP+02] S. Dixit, J. Crain, W. C. K. Poon, J. L. Finney, and A. K. Soper. Molecular segregation observed in a concentrated alcoholwater solution. *Nature*, 416:829–832, (2002).

[dLPS80] S.W. de Leeuw, J.W. Perram, and E.R. Smith. Simulation of electrostatic systems in periodic boundary conditions. i. lattice sums and dielectric constants. *Proc. Roy. Soc. London*, 373:27–56, (1980).

[DMK94] Zhihong Deng, Glenn J. Martyna, and Michael L. Klein. Quantum simulation studies of metal-ammonia solutions. *J.Chem.Phys.*, 100:7590, (1994).

[DO58] B.G. Dick and A.W. OverHauser. Theory of the dielectric constants of alkali solids. *Physical Review*, 112:90, (1958).

[Dru00] P. Drude. *Lehrbuch der Optik*. S.Hirzel, Leipzig, (1900).

[DZ71] M.B. Doran and I.J. Zucker. Higher order multipole three-body van der waals interactions and stability of rare gas solids. *Journal of Physical Chemistry*, 4:307, (1971).

[EK69] D. Eisenberg and W. Kauzmann. *The Structure and Properties of Water*. Oxford Univ. Press, (1969).

[EL30] R. Eisenschitz and F. London. Über das Verhältnis der van der Waalsschen Kraefte zu den Homöopolaren Bindungskraeften. *Zeitschrift für Physik*, 60:491, 1930.

[EPB+95] Ulrich Essmann, Lalith Perera, Max L. Berkowitz, Tom Darden, Hsing Lee, and Lee G. Pedersen. A smooth particle mesh Ewald method. *J.Chem.Phys*, 103:8577–8593, (1995).

[ES90] Andrew D. Ellington and Jack W. Szostak. In vitro selection of RNA molecules that bind to specific ligands. *Nature*, 346:818, (1990).

[Fey72] Richard P. Feynman. *Statistical Mechanics*. Advanced Book Program, (1972).

[Flu71] S. Flugge. *Practical Quantum Mechanics*. Springer Verlag, Berlin, (1971).

[Fon61] Peter R. Fontana. Theory of Long-Range Interatomic Forces i. Dispersion Energies between Unexcited Atoms. *Physical Review*, 123:18651870, September (1961).

[FS96] Daan Frenkel and Berend Smit. *Understanding Molecular Simulation*. Academic Press, (1996).

[GD09] Charles A. Geiger and Edgar Dachs. Quasi-ice-like C_P behavior of molecular H_2O in hemimorphite $Zn_4Si_2O_7(OH)_2H_2O$: C_P and entropy of confined H_2O in microporous silicates. *American Mineralogist*, 94:634637, (2009).

[GEFS10] Ann K Gauger, Stephanie Ebnet, Pamela F Fahey, and Ralph Seelke. Reductive evolution can prevent populations from taking simple adaptive paths to high fitness. *Bio-complexity*, (2010).

[Goo04] David S. Goodsell. *Bionanotechnology: Lessons from Nature*. Wiley-Blackwell, (2004).

[Goo09] David S. Goodsell. *The Machinery of Life*. Springer, (2009).

[GR87] L. Greengard and V. Rokhlin. A fast algorithm for particle simulations. *J.Comp.Sim.*, 73:325–348, (1987).

[GS71] R. C. Grimm and R. G. Storer. Monte-carlo solution of schroedinger's equation. *J.Comp.Phys.*, 7:134–156, (1971).

[HB05] Henry Hess and George D. Bachand. Biomolecular motors. *Materials Today*, 8:22–29, (2005).

[HBB82] M.F. Herman, E.J. Bruskin, and B.J. Berne. On path integral monte carlo simulations. *J.Chem.Phys.*, 76(10):5150, May (1982).

[HCB54] J.O. Hirschfelder, C.F. Curtiss, and R.B. Bird. *The Molecular Theory of Gases and Liquids, John Wiley and Sons*. John Wiley and Sons, New York, (1954).

[Het84] J.H. Hetherington. Observations on the statistical iteration of matrices. *Phys.Rev.A*, 30:27132719, (1984).

[HJR94] B.L. Hammond, W.A. Lester Jr., and P.J. Reynolds. *Monte Carlo methods in ab initio quantum chemistry*. World Scientific, (1994).

[Hoo85] William G. Hoover. Canonical dynamics: Equilibrium phase-space distributions. *Phys. Rev. A*, 31(3):16951697, March (1985).

[HPDS05] David H. Herce, Lalith Perera, Thomas A. Darden, and Celeste Sagui. Surface solvation for an ion in a water cluster. *J.Chem.Phys*, 122:24513, (2005).

[JA91] Christopher J. Jones and Shin-Ichi Aizawa. The bacterial flagellum and flagellar motor: Structure, assembly and function. *Advances in Microbial Physiology*, 32:109–172, (1991).

[Jas55] Robert Jastrow. Many-body problem with strong forces. *Physics Review*, 98:14791484, 1955.

[JCM+83] William L. Jorgensen, Jayaraman Chandrasekhar, Jeffry D. Madura, Roger W. Impey, and Michael L. Klein. Comparison of simple potential functions for simulating liquid water. *J.Chem.Phys.*, 79:926, (1983).

[JMC+09] Andrew Jones, Glenn J. Martyna, Jason Crain, Martin Müser, and Andrew Thompson. Norm-conserving diffusion Monte Carlo method and diagrammatic expansion of interacting Drude oscillators: Application to solid xenon. *Phys. Rev. B*, 79:144119, (2009).

[JT02] Pavel Jungwirth and Douglas J. Tobias. Chloride anion on aqueous clusters, at the air-water interface, and in liquid water: Solvent effects on cl- polarizability. *J.Chem.Phys.A*, 106:379383, (2002).

[KLF+00] E. M. Knipping, M. J. Lakin, K. L. Foster, P. Jungwirth, D. J. Tobias, R. B. Gerber, D. Dabdub, and B. J. Finlayson-Pitts. Experiments and simulations of ion-enhanced interfacial chemistry on aqueous NaCl aerosols. *Science*, 288:301–306, (2000).

[KSS+96] Michael J. Kozal, Nila Shah, Naiping Shen, Robert Yang, Raymond Fucini, Thomas C. Merigan, Douglas D. Richman, Don Morris, Earl Hubbell, Mark Chee, and Thomas R. Gingeras. Extensive polymorphisms observed in hiv-1 clade b protease gene using high-density oligonucleotide arrays. *Nature Medecine*, 2:753–759, (1996).

[KW86] Malvin H. Kalos and Paula A. Whitlock. *Monte Carlo Methods.* John Wiley and Sons, New York, NY, (1986).

[LADMR03] Guillaume Lamoureux, Jr. Alexander D. MacKerell, and Benoît Roux. A simple polarizable model of water based on classical drude oscillators. *J.Chem.Phys.*, 119(10):5185, (2003).

[LBP74] Jong K. Lee, J. A. Barker, and G. M. Pound. Surface structure and surface tension: Perturbation theory and Monte Carlo calculation. *J.Chem.Phys.*, 60:1976, (1974).

[LCW99] K. Lum, D. Chandler, and J.D. Weeks. Hydrophobicity at small and large length scales. *Journal of Physical Chemistry B*, 103:45704577, (1999).

[LJ24] J.E. Lennard-Jones. On the Determination of Molecular Fields. *Proc.R.Soc.Lond.A*, 106:463, (1924).

[LLM09] Frederic Legoll, Mitchell Luskin, and Richard Möeckel. Non-ergodicity of Nose-Hoover dynamics. *Nonlinearity*, 22:1673, (2009).

[MB05] J. Mitroy and M.W.J. Bromley. Higher-order C_n dispersion coefficients for hydrogen. *Phys.Rev.A*, 71:32709, (2005).

[MD00] Cait E. MacPhee and Christopher M. Dobson. Formation of mixed fibrils demonstrates the generic nature and potential utility of amyloid nanostructures. *J. Am. Chem. Soc.*, 122:1270712713, (2000).

[MKT92] Glenn J. Martyna, Michael L. Klein, and Mark Tuckerman. Nosé-Hoover chains: The canonical ensemble via continuous dynamics. *J. Chem. Phys.*, 97(4):2635, August (1992).

[MM08] Thomas E. Markland and David E. Manolopoulos. An efficient ring polymer contraction scheme for imaginary time path integral simulations. *J.Chem.Phys.*, 129:024105, (2008).

[MO05] J. Mitroy and Vitali D. Ovsiannikov. Generating van der Waals coefficients to arbitrary orders of the atomatom interaction. *Chem.Phys.Lett.*, 412:76–81, (2005).

[MPL+01] John B. O. Mitchell, Sarah L. Price, Maurice Leslie, David Buttar, and Ron J. Roberts. Anisotropic repulsion potentials for cyanuric chloride ($C_3N_3Cl_3$) and their application to modeling the crystal structures of azaaromatic chlorides. *J. Phys. Chem. A*, 105:99619971, (2001).

[MW70] I.R. McDonald and L.V. Woodcock. Triple-dipole dispersion forces in dense fluids. *J.Phys.C: Solid State Physics*, 3:722, (1970).

[MW05] Simon K. Mencher and Long G. Wang. Promiscuous drugs compared to selective drugs (promiscuity can be a virtue). *BMC Clinical Pharmacology*, 5:3, (2005).

[NM09] R.A. Nistor and M.H. Mueser. Dielectric properties of solids in the regular and split-charge equilibration formalisms. *Phys.Rev.B*, 79:104303, (2009).

[OGL88] V. D. Ovsiannikov, A. V. Guilyarovskia, and O.Ya. Lopatko. Higher order effects in dispersion interaction of atoms. *Molecular Physics*, 64:111–123, (1988).

[OM06] Vitali D. Ovsiannikov and J. Mitroy. Regular approach for generating van der Waals C_s coefficients to arbitrary orders. *J. Phys. B: At. Mol. Opt. Phys.*, 39:159, (2006).

[PC84] E. L. Pollock and D. M. Ceperley. Simulation of quantum many-body systems by path-integral methods. *Phys. Rev. B*, 30(5):25552568, September (1984).

[PC03] Jay W. Ponder and David A. Case. Force fields for protein simulations. *Advances in Protein Chemistry*, 66:27–85, (2003).

[PISW49] L. Pauling, H.A. Itano, S.J. Singer, and I.C. Wells. Sickle cell anemia, a molecular disease. *Science*, 110:543, (1949).

[PR84] M. Parrinello and A. Rahman. Study of an F center in molten KCl. *J.Chem.Phys.*, 80:860, (1984).

[PT05] Priya V. Parandekar and John C. Tully. Mixed quantum-classical equilibrium. *J.Chem.Phys.*, 122:094102, (2005).

[RI91] A.K. Rappe and W.A. Goddard III. Charge equilibration for molecular dynamics simulations. *J. Phys. Chem.*, 95:3358, (1991).

[Rok85] V. Rokhlin. Rapid solution of integral equations of classical potential theory. *J.Comp.Sim.*, 60:187–207, (1985).

[RS71] A. Rahman and F. H. Stillinger. Molecular dynamics study of liquid water. *J.Chem.Phys.*, 55:3336, (1971).

[SC85] J.M. Standard and P.R. Certain. Bounds to two- and three-body long-range interaction coefficients for S-state atoms. *J.Chem.Phys.*, 83:3002, (1985).

[Sch81] L. S. Schulman. *Techniques and Applications of Path Integration*. Wiley, New York, (1981).

[SD76] M.J.L. Sangster and M. Dixon. Interionic potentials in alkali halides and their use in simulations of the molten salts. *Advances in Physics*, 25:247–342, (1976).

[SD03] Massimo Stefani and Christopher M. Dobson. Protein aggregation and aggregate toxicity: new insights into protein folding, misfolding diseases and biological evolution. *J.Mol.Med.*, 81:678–699, (2003).

[SD10] C.V. Sindelar and K.H. Downing. An atomic-level mechanism for activation of the kinesin molecular motors. *PNAS*, 107:4111–4116, (2010).

[SHT+99] W.R.P. Scott, P.H. Huenenberger, I.G. Tironi, A.E. Mark, S.R. Billeter, J. Fennen, A.E. Torda, T. Huber, P. Krueger, and W.F. van Gunsteren. The GROMOS biomolecular simulation program package. *J. Phys. Chem. A*, 103:3596–3607, (1999).

[SJ05] Thomas Sommerfeld and Kenneth D. Jordan. Quantum drude oscillator model for describing the interaction of excess electrons with water clusters: An application to $(H_2O)_{13}^-$. *J. Phys. Chem. A*, 109:11531, 2005.

[SK88] Michiel Sprik and Michael L. Klein. A polarizable model for water using distributed charge sites. *J.Chem.Phys.*, 89:7556, (1988).

[SKW90] Michiel Sprik, Michael L. Klein, and Kyoko Watanabe. Solvent polarization and hydration of the chlorine anion. *J.Chem.Phys.*, 94:64836488, (1990).

[SLW99] Daniela Stock, Andrew G. W. Leslie, and John E. Walker. Molecular architecture of the rotary motor in atp synthase. *Science*, 286:1700–1705, (1999).

[SM01] Philipp Schöffel and Martin H. Müser. Elastic constants of quantum solids by path integral simulations. *Phys. Rev. B*, 63:224108, (2001).

[Smi82] W. Smith. *CCP5 Info. Quart.*, 4, 1982.

[Smi98] W. Smith. *CCP5 Info. Quart.*, 46, 1998.

[SO96] Attila Szabo and Neil S. Ostlund. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*. Dover Publications Inc., (1996).

[SP88] A.J. Stone and S.L. Price. Some new ideas in the theory of intermolecular forces: anisotropic atom-atom potentials. *J.Phys.Chem.*, 92:33253335, (1988).

[SR74] F.H. Stillinger and A. Rahman. Improved simulation of liquid water by molecular dynamics. *J.Chem.Phys.*, 60:1545–1557, (1974).

[SSCW81] Kenneth S. Schweizer, Richard M. Stratt, David Chandler, and Peter G. Wolynes. Convenient and accurate discretised path integral methods for equilibrium quantum mechanical calculations. *J.Chem.Phys.*, 75:1347, (1981).

[Sto96] A.J. Stone. *The Theory of Intermolecular Forces*. Oxford University Press, (1996).

[Suz76] M. Suzuki. Generalized trotter's formula and systematic approximants of exponential operators and inner derivations with applications to many-body problems. *Commun.Math.Phys.*, 51:183–190, (1976).

[Suz86] M. Suzuki. Quantum statistical monte carlo methods and applications to spin systems. *J. Stat. Phys.*, 43:883–909, (1986).

[Suz94] M. Suzuki. Convergence of general decompositions of exponential operators. *Commun. Math. Phys.*, 163:491–508, (1994).

[TBB83] Devarajan Thirumalai, Eric J. Bruskin, and Bruce J. Berne. An iterative scheme for the evaluation of discretized path integrals. *J.Chem.Phys.*, 79:5063, (1983).

[TBJK93] Mark E. Tuckerman, Bruce J. Berne, Glenn J.Martyna, and Michael L. Klein. Efficient molecular dynamics and hybrid monte carlo algorithms for path integrals. *J. Chem. Phys.*, 99:2796, (1993).

[TBM92] M. Tuckerman, B. J. Berne, and G. J. Martyna. Reversible multiple time scale molecular dynamics. *J.Chem.Phys.*, 97:491–508, (1992).

[TG02] D.M. Taverna and R.A. Goldstein. Why are proteins marginally stable? *Proteins*, 46:105–109, (2002).

[Tia10] Pu Tian. Computational protein design, from single domain soluble proteins to membrane proteins. *Chem.Soc.Rev.*, 39:2071–2082, (2010).

[TNC76] K.T. Tang, J.M. Norbeck, and P.R. Certain. Upper and lower bounds of two- and three-body dipole, quadrupole, and octupole van der Waals coefficients for hydrogen, noble gas, and alkali atom interactions. *J.Chem.Phys.*, 64:3063, (1976).

[TP71] John C. Tully and Richard K. Preston. Trajectory surface hopping approach to nonadiabatic molecular collisions: The reaction of H^+ with D_2. *J.Chem.Phys*, 55:562–572, (1971).

[Tro59] H.F. Trotter. On the product of semi-groups of operators. *Proc.Am.Math.Soc*, 10:545–551, (1959).

[TS01] P. Tangney and S. Scandolo. How well do CarParrinello simulations reproduce the BornOppenheimer surface? Theory and examples. *J.Chem.Phys.*, 116:14, (2001).

[TT99] Glenn Toole and Susan Toole. *New Understanding Biology for Advanced Level*. Thomas Nelson, (1999).

[UF00] C.J. Umrigar and C. Filippi. Correlated sampling in quantum Monte Carlo: A route to forces. *Phys.Rev.B*, 61:R16291, (2000).

[UNR93] C. J. Umrigar, M. P. Nightingale, and K. J. Runge. A diffusion Monte Carlo algorithm with very small time-step errors. *J.Chem.Phys.*, 99:2865, (1993).

[Vin09] J.F.V. Vincent. Biomimetics a review. *Proc.Inst.Mech.Eng*, 223:919–939, (2009).

[WCCM96] Mark Wilson, Benedito J. Costa-Cabral, and Paul A. Madden. Quadrupole polarization in simulations of ionic systems: Application to agcl. *J.Phys.Chem.*, 100:12271237, 1996.

[WM06] Troy W. Whitfield and Glenn J. Martyna. A unified formalism for many-body polarization and dispersion: The quantum Drude model applied to fluid Xenon. *Chem. Phys. Lett.*, 424:409–413, (2006).

[WM07] Troy Whitfield and Glenn J. Martyna. Low variance energy estimators for systems of quantum Drude oscillators: Treating harmonic path integrals with large separations of time scales. *J. Chem. Phys.*, 126:074104, (2007).

[ZHMB04] Ruhong Zhou, Xuhui Huang, Claudio J. Margulis, and Bruce J. Berne. Hydrophobic collapse in multidomain protein folding. *Science*, 305:1605–1609, (2004).