Improving the accuracy of surface simulations

by

Ben Wood

Thesis submitted for the degree of
Doctor of Philosophy of the University of London
and the
Diploma of Imperial College

Condensed Matter Theory Group
The Blackett Laboratory
Imperial College
London

January 2005

Abstract

This thesis will address the problem of carrying out electronic structure calculations on quasi-2D systems, focusing on the use of quantum Monte Carlo methods and using the quasi-2D electron gas as a test system.

Two techniques aimed at improving the accuracy of these simulations will be proposed and tested:

- the use of a modified periodic Coulomb interaction to reduce finite-size errors. It will be shown that the modified interaction does not reduce the finite-size errors; this failure will be analysed and explained. However, it is significantly faster than the conventional interaction, and is therefore recommended.

- the application of classical plasmon theory to improve the quality of trial wave functions. A new, intuitive approach to the connection between plasmons and the electronic wave function based on physical arguments will be presented. The predictions for the general form of the wave function in inhomogeneous materials agree with previous results obtained by other methods. In addition, the new approach makes clear the role of classical plasmon normal modes; an analysis of these modes in the case of the quasi-2D electron gas leads directly to an improved wave function for this system.

Finally, a new calculation of the surface energy of the electron gas will be presented.

Acknowledgements

I am greatly indebted to the many people who have helped and supported me in various ways during the course of this degree.

My flatmates from Burgoyne Road, my friends from Oxford, my colleagues from the Condensed Matter Theory group, my teammates from Brixton Town and my fellow Beards have all contributed to keeping me happy over the last three years.

I have been very lucky to have a supervisor as enthusiastic, dedicated and supportive as Matthew Foulkes, who has taught me a great deal.

Finally, I am incredibly grateful to Debora, Emma and my parents for their wonderful love and support.

Contents

Table of Contents **4**

List of Figures **8**

List of Tables **15**

1 Introduction **17**

2 The simplification of many-electron quantum mechanics **20**

2.1 First steps 20

2.2 The single-electron approach 22

2.3 Hartree-Fock theory 23

2.4 Density-functional theory 25

3 Quantum Monte Carlo methods **30**

3.1 Random sampling 31

3.2 Variational Monte Carlo 33

3.2.1 Importance sampling in VMC 34

3.2.2 The Metropolis algorithm 34

3.2.3 Advantages and disadvantages of VMC 36

3.3 Diffusion Monte Carlo 36

3.3.1 Simulating classical diffusion 37

3.3.2 Application to quantum mechanics 38

3.3.3 The ground-state wave function 40

3.3.4 Importance-sampled diffusion Monte Carlo 41
3.3.5 Fermions 48
3.3.6 Estimators 49
3.4 Trial wave functions 52
3.4.1 The Jastrow factor 53
3.4.2 Optimisation 54

4 Errors in QMC simulations **57**
4.1 Finite-size errors 57
4.1.1 The independent-particle finite-size effect 58
4.1.2 Coulomb finite-size errors 61
4.2 Fixed-node errors 65
4.3 Statistical noise 66
4.4 Surface calculations 67
4.4.1 System geometry 67
4.4.2 The surface energy 69

5 The jellium slab **70**
5.1 Defining the system 70
5.1.1 Constraining the electrons 71
5.2 The surface energy of jellium 72
5.3 Preliminary investigations 76
5.4 The jellium slab in QMC 80

6 The modified periodic Coulomb interaction in quasi-2D systems **84**
6.1 The MPC interaction 84
6.2 Quasi-2D systems 88
6.3 Results 89
6.4 Analysis 93
6.5 The exchange-correlation hole 94

7 The electronic ground-state wave function from classical plasmon normal modes 106

7.1 Derivation of the classical plasmon Hamiltonian and Lagrangian . . . 107

7.2 Quantisation and diagonalisation of the Hamiltonian 111

7.3 The ground-state wave function 114

7.4 Normal modes 116

8 Applying the plasmon normal mode theory to slab systems 120

8.1 Classical plasmons in slabs 120

8.1.1 The metal conductivity 121

8.1.2 Solving Maxwell's equations 122

8.1.3 Boundary conditions 124

8.1.4 Surface plasmons 125

8.1.5 Bulk plasmons 130

8.1.6 Surface plasmons in electrostatic theory 134

8.2 The plasmon wave function for slab systems 137

8.2.1 Approximate analytic solution for an infinite slab 141

8.3 Creating a realistic Jastrow factor 143

8.3.1 Removing undesirable cusps 147

8.3.2 Applying the electron-electron cusps 148

8.4 Results 154

8.4.1 Unbounded slab 155

8.4.2 Bounded slab 160

9 A new calculation of the jellium surface energy 166

9.1 Improved orbitals 167

9.2 Alternative **k**-point sampling 169

9.3 Surface energy calculations 174

9.3.1 Results 175

9.3.2 Analysis 176

10 Conclusions **181**

A The quasi-2D Ewald sum **184**

A.1 Derivation 184

A.2 Expansion 187

B The cusp conditions **192**

C Integrating the cusp function **196**

D Reconstructing a probability density function **200**

D.1 Use of a weighting function 200

D.2 Projection onto basis functions 204

Bibliography **207**

List of Figures

4.1 Illustrating the lattice generated by the periodic repeat of the simulation cell. Note that any movement of the particle in the simulation cell is copied by all the images. 58

4.2 Shell-filling effects. Plotted here is the kinetic energy per electron of a non-interacting 3D electron gas obtained for different system sizes. (The density parameter is $r_s = 2$.) 59

4.3 The effective system when applying periodic boundary conditions to the simulation cell. The shaded areas represent the positive background; the electron density is similar, though with less sharply-defined edges. 68

5.1 Electron density profiles for a conventional jellium slab with $r_s = 2.07$ and $s = 17.64248$. The profiles are obtained from LDA calculations and demonstrate the effect on the electron density of using a cell of finite extent in the xy-plane. Because r_s and s are fixed, the cell size is determined by the number of electrons. 72

5.2 Electron density profiles for a bounded jellium slab, obtained in LDA. 73

5.3 The LDA energy per electron as a function of the number of grid points. Cells of size w ∼105 and w ∼210 are compared, both for finite systems containing 360 electrons and for systems infinite in the xy-direction. Note that the two vertical scales are offset. The slab width is 17.64248; the density parameter is $r_s = 2.07$. The length of the cell in the z-direction varies slightly as a function of the number of grid points; this is to ensure optimal sampling of the important slab region. 77

5.4 The dependence of the LDA energy per electron on the slab width; for this system, $r_s = 2.07$, the number of electrons is 360 and the cell size in the z-direction is around 105. 78

5.5 The surface energy as a function of slab width, calculated using different values of the bulk energy. The middle curve uses the Perdew-Wang parameterisation [69] of Ceperley and Alder's early QMC results [12], which was the functional employed in the slab calculation. The upper curve is from the Perdew-Zunger [70] paremeterisation of the same results; the lower curve is from the Perdew-Zunger parameterisation of Ortiz and Ballone's later QMC calculations [63]. 79

5.6 The components of the LDA surface energy as a function of the slab width. All the surface energies are given in mHa bohr^{-2}. 80

5.7 The distribution of configuration energies, before and after variance minimisation. 10000 configurations have been included; the Jastrow factor was of the form introduced in section 3.4.1 and described in detail in section 6.3. Optimisation of both u and χ terms was allowed; The minimisation was performed without reweighting. 81

6.1 The total VMC energy per electron as a function of system size, comparing the Ewald and MPC interactions. The exchange energy given by equation (6.25) is also plotted (note the different vertical scale). Energies, as usual, are in Ha. 90

6.2 The total energy per electron as a function of system size, obtained using fixed-node DMC. 91
6.3 The time taken to calculate the electron-electron interaction during an 8000-move VMC simulation, comparing the quasi-2D versions of the MPC and Ewald interactions. 92
6.4 The charge disturbance created by an electron-like test charge at the surface ($z_0 = 0$) and in the bulk ($z_0 = 10$). The plots show the value of $\delta n_{\mathrm{ind}}(\mathbf{r})$ at the points $(x, 0, z_0)$ (labelled x) and $(0, 0, z_0 + z)$ (labelled z). In the bulk, the x and z curves are identical. 99
6.5 The effect of cell size on the shape of the hole. Systems of size $L = 20$, $L = 40$ and $L = 80$ are compared at a distance of 10 from the centre of the imposed charge; this corresponds to the edge of the smallest cell, where any difference should be most clearly visible. The curves for the bulk hole at each of the three sizes are indistinguishable on this scale. 100
6.6 The Thomas-Fermi hole for different cell sizes in a more realistic model of the slab. The induced electron density δn_{ind} is plotted at the points $(x, 0, z_0)$, where z_0 corresponds to the slab centre (labelled 'middle') or the conventional slab edge (labelled 'surface'). 102
6.7 The exchange-correlation hole for a realistic slab. 105

8.1 The metal slab. 121
8.2 The surface charge density. 125
8.3 The dispersion relation for a slab with $s = 200$ and density parameter $r_s = 2.0$ (in Hartree atomic units). This density corresponds to a plasma frequency $\omega_p = 0.6124$. The crosses represent the results of applying a numerical root-finding algorithm to equation (8.39); the smooth lines correspond to the small-k approximation given in equation (8.42). In the limit of large k, ω tends to $\omega_p/\sqrt{2}$. 128

8.4 The x- and z-components of the longitudinal electric field associated with the surface plasmon oscillations. 131

8.5 The dispersion relation for surface plasmons in the electrostatic theory. The slab width is 20 au. For metallic densities, the Fermi wave vector is of order 1, whereas $k_p \sim 0.01$. The plasmon frequency very rapidly reaches the large-k limit of $\omega_p/\sqrt{2}$; this is the result obtained for a semi-infinite system, and indicates that the coupling between surface plasmon modes on the two sides of the slab is weaker than that obtained when using the full dynamical theory. 136

8.6 The plasmon contribution to the two-body function in the Jastrow factor. In this and subsequent graphs, one electron is fixed (here at $z' = 2.0$, inside the slab); the plot shows the dependency of $u_{\rm pl}^{\infty}$ on the parallel separation $\Delta r_{\parallel}$ and the z-coordinate of the other electron. 143

8.7 The fixed electron is at $z' = 1.0$: inside the slab but close to the surface. 144

8.8 The fixed electron is now just outside the slab at $z' = -1.0$; the bulk plasmons no longer contribute to $u_{\rm pl}^{\infty}$. 145

8.9 The fixed electron is now further outside the slab ($z' = -2.0$), and the strength of the correlation is reduced. 146

8.10 Removing the cusps from $u_{\rm pl}$ at the slab boundaries. In each plot, one electron is fixed while the other is scanned along a line in the z-direction; the x- and y-coordinates are chosen to be the same, so that $\Delta r_{\parallel} = 0$. The smoothed and unsmoothed curves are almost indistinguishable when the electrons are far from the slab edges; the most pronounced difference appears when one electron is just inside the slab edge (the black curves). This and the following plots were calculated for a cell containing 600 electrons, with $s = 17.64248$ and $r_s = 2.07$ (in Hartree atomic units). 149

8.11 Removing the cusps from $\chi_{\rm bulk}$ at the slab boundaries; $\chi_{\rm bulk}^{\rm s}$ is compared with $\chi_{\rm bulk}$. 150

8.12 The smoothed, cusp-corrected plasmon two-body function. The homogeneous curve is plotted for comparison. 151
8.13 The one-body term χ_{cusp}, for various values of the cut-off distance L_c, calculated using equation (8.119); this equation was derived using the approximation of constant electron density within the slab. This term is designed to combat the density-changing effects of u_{cusp}. As L_c becomes large, the curves tend to a limit, because the decay of u_{cusp} is then dominated by k_c rather than L_c. 154
8.14 The electron density profile, from LDA and VMC. The wave function used in VMC was a product of determinants made up of LDA orbitals, with no Jastrow factor; the densities are the same, to within the VMC noise. 156
8.15 The effect on the electron density of including a homogeneous two-body term of the form given in equation (8.113) in the wave function. The density becomes slightly more homogeneous, with more electrons being pushed into the vacuum regions outside the slab. 157
8.16 The electron density profile when the plasmon two-body term ($u_{\text{cusp}}+u^{\text{s}}_{\text{pl}}$) is included in the VMC wave function. The profile is completely different to the original curve, with almost all the electrons now at the slab edges. The change is much more pronounced than when the homogeneous (and short-ranged) u_{cusp} alone was used. 158
8.17 The density profile when the full plasmon Jastrow factor is included in the VMC wave function. The result is very close to the original density, although not perfect. 159
8.18 The electron density profile for a Jastrow factor consisting of only the short-range two-body function u_{cusp} and the corresponding one-body term χ_{cusp}. 160
8.19 Electron density profile for the bounded slab, in LDA and VMC (with no Jastrow factor). 161

8.20 Electron density profile for the bounded slab, with a Jastrow factor containing only the short-ranged two-body term. The disturbance to the density is small; this is reflected in the VMC energy, which is significantly lower than when using no Jastrow factor. 162

8.21 Electron density profile for the bounded slab, with the plasmon two-body term only. The long-ranged correlations alter the density drastically, as in the unbounded slab; electrons are pushed to the slab edges. 163

8.22 Electron density profile for the bounded slab, with the full plasmon Jastrow factor. The density is very close to the LDA form, and the energy is lower than when using only the short-ranged Jastrow. . . . 164

8.23 Electron density profile for the bounded slab, with the full short-ranged Jastrow factor. The LDA density is almost restored, and the energy is lower than when using only u_{cusp}, though not as low as that obtained with the full plasmon Jastrow factor. 165

9.1 The effect of including the image tail in V_{XC}. The simple LDA form (shown for comparison) is matched to the image potential outside the slab. The test system has density parameter $r_s = 2.07$ and slab width $s = 17.64248$; the simulation cell contains 600 electrons. In addition to the exchange-correlation potential V_{XC}, the Kohn-Sham potential $V_{\mathrm{KS}} = V_{\mathrm{Ha}} + V_{\mathrm{B}} + V_{\mathrm{XC}}$ is also plotted. The positions of the image planes are indicated by ticks on the z-axis. 168

9.2 The effect of the image tail on the highest occupied single-electron sub-band. The original orbital is shown, along with the difference between this and the modified orbitals (magnified one hundred times - note the two scales). 169

9.3 The electron density profile with periodic and antiperiodic boundary conditions for two different cell sizes. The infinite-cell limit is also shown (because this may be calculated in DFT). 171

9.4 The lowest sub-band for different boundary conditions and cell sizes. The infinite-cell form is included for reference. 172
9.5 Optimising the parameter α. The VMC energies are shown along with the quadratic function fitted to them. These results are for a slab width of 11.7783 with a cell containing 588 electrons; the minimum is at $\alpha = 2.26$. 176
9.6 The slab energy per electron, plotted against s/N, where s is the slab width and N is the number of electrons. The abscissa is proportional to $1/L^2$, and was chosen to demonstrate the form of the finite-size errors. 177
9.7 Comparing the MPC and Ewald interactions when $s = 18.4851$. . . . 178
9.8 The slab energy per electron, with $L^2 = 1150$, as a function of the slab width. 180

List of Tables

8.1 The function $F_{k_\parallel}(z, z')$. This contains all the z- and z'-dependence of the part of $u_{\rm pl}$ arising from the surface plasmon contribution. The constants $A_{k_\parallel}$, $B_{k_\parallel}$, $C_{k_\parallel}$ and $D_{k_\parallel}$ are defined in equation (8.91). 140

8.2 The energy per electron, calculated in VMC, for the unbounded jellium slab. Results for two different cell sizes and various forms of Jastrow factor are shown. The 'full plasmon Jastrow' refers to a wave function of the type described in equation (8.121); the 'short-ranged Jastrow' uses only $u_{\rm cusp}$ and $\chi_{\rm cusp}$. 159

8.3 The VMC energy per electron for the bounded jellium slab. 162

9.1 Comparison of energies calculated with different versions of the exchange-correlation potential. The DFT energies are identical for all potentials. The trial wave function for the VMC calculations contained no Jastrow factor; the simulations used a simulation cell containing 600 electrons. 170

9.2 The occupation numbers of each sub-band for two cell sizes and periodic or anti-periodic boundary conditions. 172

9.3 The fractional occupation of each sub-band. This is simply the occupation number divided by the total number of electrons, and allows the results of finite-cell calculations to be compared with those obtained in the infinite-cell limit. 173

9.4 Comparison of energies calculated with different versions of the exchange-correlation potential. All energies are in mHa. The trial wave function for the VMC calculations contained no Jastrow factor. 173

9.5 Slab energies per electron calculated in DMC, for various slab widths and system sizes. The surface energy estimates listed in the final column are included to demonstrate the errors introduced by combining the results of bulk and slab calculations without due care. 179

Chapter 1

Introduction

In recent decades, the use of computer simulation as a tool complementary to traditional experiment has become widely accepted. In many situations, 'real-life' experiments are currently beyond our capabilities; in others, the difficulty of the experiments means that the results are open to interpretation.

One area in which simulations have become ubiquitous is electronic structure. Here, the most widely-used computational tool is density-functional theory (DFT); an indication of the significance and impact of this method is the decision to award the 1998 Nobel Prize for Chemistry to its creator, Walter Kohn. Electronic structure is also the category into which this thesis falls, although the focus is not on DFT but on another technique: quantum Monte Carlo (QMC) simulation.

While DFT has become hugely popular, and has relatively low computational cost, it is not regarded as the most accurate tool available. For many systems, the benchmark against which other techniques are judged is QMC, and in particular, the version of QMC known as diffusion Monte Carlo (DMC). In fact, DMC results are used to model the function at the heart of DFT calculations: the exchange-correlation potential.

QMC calculations are computationally expensive, and despite efforts to devise faster, more efficient algorithms, will never be able to match DFT in this regard. Their relevance comes from the need for greater accuracy than can be provided by other techniques.

To attain this high level of accuracy, great care needs to be taken in order to eliminate known sources of error. One source of error is the fixed-node approximation, introduced to address the problem of mapping an antisymmetric (and hence non-positive-definite) fermionic wave function onto a density; although several alternatives have been proposed to avoid this uncontrolled approximation, they all have significant computational cost.

Another kind of error arises when the aim is to model an infinite system; because QMC is limited to a finite number of electrons, it is normal practice to apply periodic boundary conditions to a finite simulation cell. This procedure creates finite-size errors.

Controlling errors is of particular importance when surfaces are being studied. Evaluating the surface energy of a system involves taking a difference of energies; the result may be very small, and thus more sensitive to errors than a bulk calculation.

The simplest of surface systems is a slab containing an electron gas. This system has infinite extent in two spatial dimensions, but is finite in the third. Although useful as a primitive model of a metal film, the main function of such a slab is as a test system, and as such, it has been studied using many different techniques, including QMC. However, recent results suggest that the previous QMC simulations were inaccurate. To re-establish the viability of QMC as a tool for studying surfaces, it is necessary to examine these results and to locate any sources of error.

The effort to improve the accuracy of QMC surface simulations is the main subject of this work; the test system is the simple slab described above, which is described in detail in chapter 5. The thesis begins with a description of several computational techniques, including DFT and QMC, in chapters 2 and 3. Chapter 4 contains an analysis of errors in QMC simulations, including those particular to surface calculations.

A technique for reducing finite-size errors is described in chapter 6, along with the results of applying this technique to the test system. The technique is the extension to quasi-2D systems of a method which has proved successful in the 3D case; however, there are important differences which will be described and explained.

A successful QMC calculation relies on having a good approximation to the true wave function. Typically, QMC trial wave functions are of the Slater-Jastrow form, with the information about electron-electron correlations contained in the Jastrow factor. An alternative way of generating the Jastrow factor analytically will be outlined in chapter 7; this builds on similar work carried out in the past, but presents a new perspective. The application of this method to the test system is the subject of chapter 8.

Finally, a new calculation of the surface energy of the electron gas will be presented in chapter 9.

Chapter 2

The simplification of many-electron quantum mechanics

The observable properties of any material are determined by the quantum mechanics of all the particles contained within that material. If the goal of solid state physics is to understand these properties, all that is required is the solution of the relevant many-body Schrödinger equation:[1]

$$\hat{H}\Psi(\{\xi_j\},t) = i\frac{\partial\Psi(\{\xi_j\},t)}{\partial t}. \tag{2.1}$$

where $\hat{H}$ is the Hamiltonian operator and Ψ is the wave function, which depends on time t and the full set of particle coordinates $\{\xi_j\}$. The problem is therefore easily formulated; unfortunately, it is not easily solved.

2.1 First steps

To make any progress, it is necessary to make approximations and assumptions. The first of these is to neglect all interactions except for the Coulomb. Nuclei are treated

[1]For most of this thesis, the Hartree system of atomic units will be employed, in which $\hbar = e = 4\pi\epsilon_0 = m_e = 1$; omitting these constants produces equations which are easier to read. In chapters 7 and 8, the constants will be shown explicitly to avoid confusion.

as whole entities. Secondly, it is assumed that electrons and nuclei do not move at relativistic speeds.[2] At this point, the Hamiltonian operator may be written as

$$\begin{aligned}\hat{H}(\{\mathbf{r}_i\},\{\mathbf{d}_\alpha\}) = &-\frac{1}{2}\sum_i \nabla^2_{\mathbf{r}_i} + \frac{1}{2}\sum_{i\neq j}\frac{1}{|\mathbf{r}_i-\mathbf{r}_j|} - \sum_{i,\alpha}\frac{Z_\alpha}{|\mathbf{d}_\alpha-\mathbf{r}_i|} \\ &- \frac{1}{2}\sum_\alpha \frac{1}{M_\alpha}\nabla^2_{\mathbf{d}_\alpha} + \frac{1}{2}\sum_{\alpha\neq\beta}\frac{Z_\alpha Z_\beta}{|\mathbf{d}_\alpha-\mathbf{d}_\beta|}.\end{aligned} \tag{2.2}$$

Here, $\{\mathbf{r}_i\}$ and $\{\mathbf{d}_\alpha\}$ represent electronic and nuclear positions respectively, while $\{Z_\alpha\}$ are the nuclear charges.

The next step, the *Born-Oppenheimer* or *adiabatic* approximation, relies on the fact that the nuclear masses M_α are very large. The electrons are considered to react immediately to any change in the nuclear coordinates, and the relaxed electronic configuration then provides a potential for the heavy, slow-moving nuclei. The wave function is written as a product:

$$\Psi(\{\mathbf{x}_i\},\{\mathbf{d}_\alpha\},t) = \psi(\{\mathbf{x}_i\};\{\mathbf{d}_\alpha\})\chi(\{\mathbf{d}_\alpha\})e^{-iEt}. \tag{2.3}$$

The dependence of ψ on $\{\mathbf{d}_\alpha\}$ is assumed to be parametric: ψ must vary smoothly as a function of the nuclear coordinates. The $\mathbf{x}_i$ represent electron spins as well as positions.

This approximation reduces equation (2.1) to two separate equations involving the nuclear and electronic coordinates respectively:

$$\left(-\frac{1}{2}\sum_i \nabla^2_{\mathbf{r}_i} + \frac{1}{2}\sum_{i\neq j}\frac{1}{|\mathbf{r}_i-\mathbf{r}_j|} - \sum_{i,\alpha}\frac{Z_\alpha}{|\mathbf{d}_\alpha-\mathbf{r}_i|}\right)\psi_n = E_n(\{\mathbf{d}_\alpha\})\psi_n \tag{2.4}$$

$$\left(-\frac{1}{2}\sum_\alpha \frac{1}{M_\alpha}\nabla^2_{\mathbf{d}_\alpha} + \frac{1}{2}\sum_{\alpha\neq\beta}\frac{Z_\alpha Z_\beta}{|\mathbf{d}_\alpha-\mathbf{d}_\beta|} + E_n(\{\mathbf{d}_\alpha\})\right)\chi_{n\lambda} = E_{n\lambda}\chi_{n\lambda} \tag{2.5}$$

In equation (2.4), the nuclear coordinates $\{\mathbf{d}_\alpha\}$ are fixed. The electronic coordinates $\{\mathbf{x}_i\}$ do not enter into equation (2.5); the link is provided by the electronic energy eigenvalue $E_n(\{\mathbf{d}_\alpha\})$, which is equivalent to a potential for the nuclei.

[2] In fact, it is not strictly necessary [5, 38, 29] to exclude relativistic effects and magnetic interactions, which become important in certain situations. However, in many other cases, they do not play a significant rôle and may be neglected. They are not relevant to the topics discussed in this thesis.

The product wave function introduced in equation (2.3) is not an exact eigenfunction of $\hat{H}$. To see this, consider the following matrix element, which is obtained by combining equations (2.2), (2.3), (2.4), and (2.5):

$$\begin{aligned}\langle \Psi_{n'\lambda'} | \hat{H} | \Psi_{n\lambda} \rangle = E_{n\lambda}\delta_{nn'}\delta_{\lambda\lambda'} \\ - \frac{1}{2}\sum_{\alpha}\frac{1}{M_\alpha}\int\cdots\int \psi^*_{n'}\chi^*_{n'\lambda'}\left(\nabla^2_{\mathbf{d}_\alpha}\psi_n\right)\chi_{n\lambda}\,\left(\Pi_i\, d\mathbf{x}_i\right)\left(\Pi_\beta\, d\mathbf{d}_\beta\right) \\ - \sum_{\alpha}\frac{1}{M_\alpha}\int\cdots\int \psi^*_{n'}\chi^*_{n'\lambda'}\nabla_{\mathbf{d}_\alpha}\psi_n\cdot\nabla_{\mathbf{d}_\alpha}\chi_{n\lambda}\,\left(\Pi_i\, d\mathbf{x}_i\right)\left(\Pi_\beta\, d\mathbf{d}_\beta\right).\end{aligned} \tag{2.6}$$

If $\Psi_{n\lambda}$ and $\Psi_{n'\lambda'}$ were true eigenfunctions, this matrix element would be diagonal in n and λ. Equation (2.6) shows the part played by the nuclear masses M_α in keeping the off-diagonal elements small. It may be shown [81] that the ratio of off-diagonal to diagonal elements is $1/M_\alpha$ when $n = n'$ and $1/\sqrt{M_\alpha}$ when $n \neq n'$.

The rest of this thesis will be concerned with the attempted solution of the electronic equation, equation (2.4).

2.2 The single-electron approach

Two important methods for solving equation (2.4) are based on a mapping of the many-electron problem on to one involving independent electrons; both will be outlined here. Hartree-Fock theory will be used to introduce certain features of the problem, while density-functional theory is included because it will be used later to generate trial wave functions for use in quantum Monte Carlo simulations.

Consider a simple first attempt at a solution of equation (2.4), a product of one-electron wave functions:[3]

$$\Psi(\{\mathbf{x}_i\}) = \prod_i \phi_i(\mathbf{x}_i). \tag{2.7}$$

A way of optimising this solution is to employ the variational principle. In this case,

[3]From now on, Ψ will represent the electronic wave function.

the optimum solution satisfies the condition:

$$\frac{\delta}{\delta\phi_i^*(\mathbf{x})}\left(\langle\Psi|\hat{H}|\Psi\rangle - \sum_j \lambda_j\langle\phi_j|\phi_j\rangle\right) = 0. \tag{2.8}$$

The normalisation constraints on the individual ϕ_i are incorporated into this variational equation through the Lagrange multipliers, λ_j. It should be noted that $\hat{H}$ is now the electronic Hamiltonian:

$$\hat{H}(\{\mathbf{r}_i\}) = -\frac{1}{2}\sum_i \nabla^2_{\mathbf{r}_i} + \frac{1}{2}\sum_{i\neq j}\frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} + \sum_i v_{\text{ext}}(\mathbf{r}_i) \tag{2.9}$$

where the interaction of each electron with the nuclei has been condensed to the external potential

$$v_{\text{ext}}(\mathbf{r}) = -\sum_\alpha \frac{Z_\alpha}{|\mathbf{d}_\alpha - \mathbf{r}|}. \tag{2.10}$$

Carrying out the functional differentiation indicated in equation (2.8) gives an effective Schrödinger equation for each one-electron wave function:

$$\left(-\frac{1}{2}\nabla^2_{\mathbf{r}} + \sum_{j\neq i}\int \frac{|\phi_j(\mathbf{r}')|^2}{|\mathbf{r} - \mathbf{r}'|}\,d\mathbf{r}' + v_{\text{ext}}(\mathbf{r})\right)\phi_i(\mathbf{r}) = \lambda_i\phi_i(\mathbf{r}). \tag{2.11}$$

The Lagrange multipliers, λ_i, are the equivalent of energy eigenvalues for these effective Schrödinger equations. However, the total energy cannot be obtained simply by summing them:

$$E[\Psi] = \langle\Psi|\hat{H}|\Psi\rangle \neq \sum_i \lambda_i. \tag{2.12}$$

The use of a wave function of the form described in equation (2.3) is called the Hartree approximation [33, 34], the most obvious disadvantage of which is that the wave function is not antisymmetric. The construction of a wave function which is explicitly antisymmetric leads to Hartree-Fock theory.

2.3 Hartree-Fock theory

A many-electron wave function which is antisymmetric under interchange of electrons may be constructed from a set of orthonormal one-electron orbitals in the form

of a *Slater determinant*:

$$\Psi_{\mathrm{HF}}(\{\mathbf{x}_i\}) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \phi_1(\mathbf{x}_1) & \phi_1(\mathbf{x}_2) & \cdots & \phi_1(\mathbf{x}_n) \\ \phi_2(\mathbf{x}_1) & \phi_2(\mathbf{x}_2) & \cdots & \phi_2(\mathbf{x}_n) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_n(\mathbf{x}_1) & \phi_n(\mathbf{x}_2) & \cdots & \phi_n(\mathbf{x}_n) \end{vmatrix}. \tag{2.13}$$

With this wave function, the energy expectation value $E[\Psi_{\mathrm{HF}}]$ becomes

$$\langle \Psi_{\mathrm{HF}} | \hat{H} | \Psi_{\mathrm{HF}} \rangle = \sum_i H_i + \frac{1}{2} \sum_{i,j} (J_{ij} - K_{ij}) \tag{2.14}$$

where

$$H_i = \int \phi_i^*(\mathbf{x}) \left(-\frac{1}{2}\nabla^2 + v_{\mathrm{ext}} \right) \phi_i(\mathbf{x})\, d\mathbf{x} \tag{2.15}$$

$$J_{ij} = \iint \phi_i(\mathbf{x}) \phi_i^*(\mathbf{x}) \frac{1}{|\mathbf{r} - \mathbf{r}'|} \phi_j^*(\mathbf{x}') \phi_j(\mathbf{x}')\, d\mathbf{x}\, d\mathbf{x}' \tag{2.16}$$

$$K_{ij} = \iint \phi_i^*(\mathbf{x}) \phi_j(\mathbf{x}) \frac{1}{|\mathbf{r} - \mathbf{r}'|} \phi_j(\mathbf{x}') \phi_i^*(\mathbf{x}')\, d\mathbf{x}\, d\mathbf{x}'. \tag{2.17}$$

The J_{ij} are called Coulomb integrals and were already present in the Hartree approximation. However, the *exchange integrals* K_{ij} represent something new. Note that it is no longer necessary to exclude the $i = j$ term from the double summation in equation (2.14), because $J_{ii} = K_{ii}$.

The optimum set of single-electron orbitals is found by the same procedure as before: the energy is minimised, subject to the constraint that the ϕ_i must remain normalised (equation (2.8)). As with the Hartree wave function, carrying out the variation generates an equation for each ϕ_i:

$$\left(-\frac{1}{2}\nabla_{\mathbf{r}}^2 + \hat{\jmath} - \hat{k} + v_{\mathrm{ext}}(\mathbf{r}) \right) \phi_i(\mathbf{r}) = \lambda_i \phi_i(\mathbf{r}) \tag{2.18}$$

where

$$\hat{\jmath}\phi_i(\mathbf{x}) = \sum_j \phi_i(\mathbf{x}) \int \frac{|\phi_j(\mathbf{x}')|^2}{|\mathbf{r} - \mathbf{r}'|}\, d\mathbf{x}' \tag{2.19}$$

$$\hat{k}\phi_i(\mathbf{x}) = \sum_j \phi_j(\mathbf{x}) \int \frac{\phi_j^*(\mathbf{x}')\phi_i(\mathbf{x}')}{|\mathbf{r} - \mathbf{r}'|}\, d\mathbf{x}'. \tag{2.20}$$

The Coulomb operator $\hat{j}$ represents the interaction of each electron with the average charge density; the unrestricted sum means that its own contribution to this charge density is included. This would give rise to an unphysical self-interaction, were it not for the fact that a corresponding term exists in $\hat{k}$, and the two cancel.

An important point about the non-local exchange operator $\hat{k}$ may be highlighted by making explicit the dependence of ϕ_i on spin:

$$\phi_i(\mathbf{x}) = \phi_i(\mathbf{r}, \sigma) = \zeta_i(\mathbf{r})\delta_{\sigma\sigma_i}. \tag{2.21}$$

Substituting this expression into equation (2.20) gives

$$\hat{k}\phi_i(\mathbf{x}) = \sum_j \phi_j(\mathbf{x})\delta_{\sigma_i\sigma_j} \int \frac{\zeta_j^*(\mathbf{r}')\zeta_i(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}\, d\mathbf{r}' \tag{2.22}$$

showing that the exchange interaction only affects electrons with like spins. In Hartree-Fock theory, electrons of like spin are kept apart, and this lowers the energy expectation value. The *exchange energy* is the difference between the Hartree and Hartree-Fock values.

Unfortunately, the theory does nothing to keep electrons of opposite spin apart. Such electrons interact only via the average charge density appearing in the Coulomb operator; there is no pairwise interaction to make it unfavourable for electrons of opposite spin to come together. This means that the ground-state energy calculated in Hartree-Fock theory is always higher than the true ground-state energy. The *correlation energy* is defined as the difference between the exact energy of the system and that calculated in Hartree-Fock theory.

2.4 Density-functional theory

The theorem underlying density-functional theory, due to Hohenberg and Kohn [35], states that the ground state of an N-electron system is completely determined by the one-electron density

$$n(\mathbf{r}) = \sum_{i=1}^{N} \int |\Psi(\mathbf{x}_1, \ldots, \mathbf{x}_N)|^2 \delta(\mathbf{r} - \mathbf{r}_i)\, d\mathbf{x}_1 \cdots d\mathbf{x}_N. \tag{2.23}$$

It also states that for a given Hamiltonian, the energy functional

$$E[n] = \langle \Psi_0[n] | \hat{H} | \Psi_0[n] \rangle \tag{2.24}$$

is variational. Here $\Psi_0[n]$ is the ground state wave function associated with the density n.

The proof of this theorem relies on the one-to-one mapping of the set of ground-state densities to the set of physically distinct local one-particle potentials. A clear statement of the proof may be found in the book by Dreizler and Gross [18].

The Hamiltonian is as described in equation (2.9). It is helpful to separate the term involving the external potential:

$$\hat{H} = \hat{H}_0 + \sum_i v_{\text{ext}}(\mathbf{r}_i). \tag{2.25}$$

Then

$$E[n] = F[n] + \int n(\mathbf{r}) v_{\text{ext}}(\mathbf{r})\, d\mathbf{r} \tag{2.26}$$

where

$$F[n] = \langle \Psi_0[n] | \hat{H}_0 | \Psi_0[n] \rangle \tag{2.27}$$

is a universal functional of the density, independent of the particular system (defined by v_{ext}). Minimising E with respect to n should enable the ground-state density of any system to be found, and hence all other ground-state properties.

The problem with the original proof is that it only applies to densities which are *pure-state v-representable*: ground-state densities of the set of external potentials $\{v_{\text{ext}}\}$. Searching over this restricted class of densities is difficult; the following derivation due to Levy [51, 52, 54] shows the way to extend the domain of the search to include all *pure-state N-representable* densities.

Consider the ground-state energy of a system with the one-body external poten-

tial v_{ext}:

$$\begin{aligned}
\langle\Psi_0|\hat{H}|\Psi_0\rangle &= \min_{\Psi}\langle\Psi|\hat{H}|\Psi\rangle && (2.28)\\
&= \min_{n}\min_{\Psi\to n}\langle\Psi|\hat{H}|\Psi\rangle && (2.29)\\
&= \min_{n}\min_{\Psi\to n}\left(\langle\Psi|\hat{H}_0|\Psi\rangle + \int n(\mathbf{r})v_{\text{ext}}(\mathbf{r})\,d\mathbf{r}\right) && (2.30)\\
&= \min_{n}\left(\min_{\Psi\to n}\langle\Psi|\hat{H}_0|\Psi\rangle + \int n(\mathbf{r})v_{\text{ext}}(\mathbf{r})\,d\mathbf{r}\right) && (2.31)\\
&= \min_{n}\left(F[n] + \int n(\mathbf{r})v_{\text{ext}}(\mathbf{r})\,d\mathbf{r}\right) && (2.32)\\
&= \min_{n} E[n] && (2.33)
\end{aligned}$$

where F and E have been redefined:

$$\begin{aligned}
F[n] &= \min_{\Psi\to n}\langle\Psi|\hat{H}_0|\Psi\rangle && (2.34)\\
E[n] &= F[n] + \int n(\mathbf{r})v_{\text{ext}}(\mathbf{r})\,d\mathbf{r}. && (2.35)
\end{aligned}$$

The ground-state density of the system may now be found by minimising $E[n]$ subject to normalisation constraints. The search in equation (2.28) is over all anti-symmetric, normalised N-electron wave functions; consequently, the search over n in equation (2.29) is now only restricted to those densities which correspond to such wave functions, ie. those which are N-representable. It has been shown [32] that all normalised non-negative functions are N-representable.

The remaining problem is that the precise form of the universal functional $F[n]$ is unknown. Kohn and Sham [41] proposed a scheme to overcome this difficulty, thus making DFT a practical calculational tool.

In the Kohn-Sham scheme, the system under consideration is compared with a non-interacting system with the same density. First, the energy functional is rewritten as

$$E[n] = T_{\text{s}}[n] + \frac{1}{2}\iint \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\,d\mathbf{r}\,d\mathbf{r}' + \int n(\mathbf{r})v_{\text{ext}}(\mathbf{r})\,d\mathbf{r} + E_{\text{XC}}[n], \qquad (2.36)$$

where the *exchange-correlation energy* E_{XC} has been defined to be

$$E_{\text{XC}}[n] = F[n] - T_{\text{s}}[n] - \frac{1}{2}\iint \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\,d\mathbf{r}\,d\mathbf{r}' \qquad (2.37)$$

and now contains all the difficult parts of E. $T_\mathrm{s}[n]$ is the kinetic energy of a non-interacting system of density n, and is therefore easily calculable:

$$T_\mathrm{s}[n] = -\frac{1}{2}\sum_{i=1}^{N}\int \psi_i(\mathbf{r})\nabla^2\psi_i(\mathbf{r})\,d\mathbf{r}. \tag{2.38}$$

The ψ_i are the occupied single-electron orbitals of the non-interacting system. This system has been defined to have density n, giving

$$\sum_{i=1}^{N}\sum_{\sigma}|\psi_i(\mathbf{r})|^2 = n(\mathbf{r}). \tag{2.39}$$

$E[n]$ must be minimised subject to the constraint $\int n(\mathbf{r})\,d\mathbf{r} = N$:

$$\frac{\delta}{\delta n(\mathbf{r})}\left(E[n] - \mu\int n(\mathbf{r})\,d\mathbf{r}\right) = 0. \tag{2.40}$$

This gives

$$\mu = v_\mathrm{s}(\mathbf{r}) + \frac{\delta T_\mathrm{s}[n]}{\delta n(\mathbf{r})} \tag{2.41}$$

where

$$v_\mathrm{s}(\mathbf{r}) = v_\mathrm{ext}(\mathbf{r}) + \int \frac{n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\,d\mathbf{r}' + v_\mathrm{XC}(\mathbf{r}) \tag{2.42}$$

with the *exchange-correlation potential*

$$v_\mathrm{XC}(\mathbf{r}) = \frac{\delta E_\mathrm{XC}[n]}{\delta n(\mathbf{r})}. \tag{2.43}$$

This is exactly equivalent to the result which would be obtained for a non-interacting system with the external potential v_s. The required density may therefore be calculated by solving the one-electron equations

$$\left(-\frac{1}{2}\nabla^2 + v_\mathrm{s}(\mathbf{r})\right)\psi_i(\mathbf{r}) = \epsilon_i\psi_i(\mathbf{r}) \tag{2.44}$$

and using equation (2.39). Because v_s itself depends on n, equations (2.42), (2.39), and (2.44) must be solved self-consistently. Having obtained n, the energy $E[n]$ may be calculated immediately, as long as E_XC is known.

This brief outline of the theory has not addressed the important question of establishing the existence of a non-interacting system whose density is exactly equal

to that of the interacting system being studied. This and other issues too detailed to belong here are described in the books by Dreizler and Gross [18] and Parr and Yang [64].

Unfortunately, E_{XC} is not known, although various approximations and parameterisations exist. One possibility is to apply the local density approximation (LDA), in which the exchange-correlation energy functional is taken to be

$$E_{\mathrm{XC}}^{\mathrm{LDA}}[n] = \int n(\mathbf{r})\epsilon_{\mathrm{XC}}(n(\mathbf{r}))\, d\mathbf{r} \tag{2.45}$$

where $\epsilon_{\mathrm{XC}}(n)$ is the exchange-correlation energy per electron of a uniform electron gas with density n. This approximation would be expected to give reasonable results for slowly-varying densities; however, it also works for a surprising number of strongly-inhomogeneous systems. The next logical step is to allow ϵ_{XC} to depend on ∇n as well as n: this is the generalised gradient approximation (GGA) [50, 8, 67]. Higher derivatives may also be included, giving the meta-GGA [68].

In fact, even $\epsilon_{\mathrm{XC}}(n)$ is not known analytically. The parameterised versions being used today are based on theoretical predictions combined with results from quantum Monte Carlo calculations [70].

Chapter 3

Quantum Monte Carlo methods

The term 'Quantum Monte Carlo' embraces several different techniques for solving the quantum many-body problem, all of which employ random sampling. This chapter will (briefly) describe the theory behind two of these techniques: variational and fixed-node diffusion QMC.[1] Both are zero-temperature methods, and both are routinely used in the calculation of properties of real solids. An important aspect of the application of this theory — the general form of the trial wave function — will also be described.

The aim of this chapter is not to present all the technical details associated with implementing a modern QMC algorithm; rather, the fundamental ideas which underlie the technique will be described, along with some of the more important aspects of the implementation. All the QMC calculations contained in this thesis were carried out using the CASINO program developed in Cambridge [61].

[1] These versions of QMC are discussed in detail in the review by Foulkes and coworkers [24], and, along with various others, in the book by Hammond et al. [30].

3.1 Random sampling

To calculate expectation values of observables in quantum mechanics, we need to evaluate expressions of the form:

$$\langle \hat{O} \rangle = \frac{\int \Psi^*(\mathbf{X}) \hat{O} \Psi(\mathbf{X}) \, d\mathbf{X}}{\int \Psi^*(\mathbf{X}) \Psi(\mathbf{X}) \, d\mathbf{X}}, \tag{3.1}$$

where $\hat{O}$ is a Hermitian operator and $\mathbf{X}$ is a vector representing all the electron coordinates.

For large systems, the dimension of the integrals is correspondingly large. Computation of such integrals by conventional quadrature methods becomes impractical, because the error in these calculations scales poorly in high dimensions. For example, the error in a Simpson's Rule calculation scales as $M^{-4/d}$, where M is the number of grid points and d is the dimensionality of the integral. As d increases, improving the accuracy of the calculation becomes prohibitively expensive.

In contrast, the Monte Carlo method of integration generates a statistical error in the value of the integral which scales as $M^{-1/2}$, independent of the dimension. While this is worse than grid-based techniques when d is small, it is far superior for large d.

One way of evaluating the integral

$$I = \int_\Omega g(\mathbf{r}) \, \mathrm{d}\mathbf{r} \tag{3.2}$$

is to sample a set of M random vectors $\{\mathbf{r}_i\}$ from a distribution which is uniform over the region of integration Ω. Each vector is assigned a *score* $g(\mathbf{r}_i)$ and the integral is estimated as

$$I \approx \frac{\Omega}{M} \sum_{i=1}^{M} g(\mathbf{r}_i). \tag{3.3}$$

This is the Monte Carlo method of integration, and in this form, it is not very efficient. It can be much improved by using *importance sampling*. The integral is first rewritten as

$$I = \int f(\mathbf{r}) P(\mathbf{r}) \, d\mathbf{r}, \tag{3.4}$$

where the integrand, $g(\mathbf{r})$, has been split into a *score function*, $f(\mathbf{r})$, and an *importance function*, $P(\mathbf{r})$. The importance function is restricted to the form of a probability density: it must be non-negative and normalised to unity over the region of integration. In regions where g is non-zero, P must also be non-zero.

A set of random vectors $\{\mathbf{r}_i\}$ is then drawn from the distribution $P(\mathbf{r})$. The integral is estimated as

$$I \approx \frac{1}{M} \sum_{i=1}^{M} f(\mathbf{r}_i). \tag{3.5}$$

In the limit $M \to \infty$, this expression for I is exact. For finite M, the standard error in the estimate is

$$\sigma_I = \frac{\sigma_f}{\sqrt{M}}, \tag{3.6}$$

where the variance of f is

$$\sigma_f^2 = \int \left[f(\mathbf{r}) - \mu_f\right]^2 P(\mathbf{r})\, d\mathbf{r} \tag{3.7}$$

and μ_f is the mean value; equation (3.4) shows that $\mu_f = I$. From equation (3.6), it can be seen that the error scales as $M^{-1/2}$, and also that the choice of score and importance functions is very important. The optimum selection is the one that minimises the variance; this is the solution of the equation

$$\frac{\delta}{\delta P(\mathbf{r})} \left(\sigma_f^2[P(\mathbf{r})] - \lambda \int P(\mathbf{r})\, d\mathbf{r} \right) = 0. \tag{3.8}$$

Here, λ is the Lagrange multiplier associated with the normalisation constraint $\int P(\mathbf{r})\, d\mathbf{r} = 1$. Writing σ_f^2 in terms of g and P and carrying out the functional differentiation gives

$$\lambda = -\frac{g^2(\mathbf{r})}{P^2(\mathbf{r})}, \tag{3.9}$$

which leads to the following result for the ideal form of $P(\mathbf{r})$:

$$P(\mathbf{r}) = \frac{|g(\mathbf{r})|}{\int |g(\mathbf{r}')|\, d\mathbf{r}'}. \tag{3.10}$$

The corresponding score function is

$$f(\mathbf{r}) = \frac{g(\mathbf{r})}{|g(\mathbf{r})|} \int |g(\mathbf{r}')|\, d\mathbf{r}'. \tag{3.11}$$

This ideal form unfortunately cannot be achieved because it requires the evaluation of the difficult integral $\int |g(\mathbf{r}')|\, d\mathbf{r}'$. However, choosing a good importance function, close to the optimum, reduces the statistical error in the estimate of I significantly.

3.2 Variational Monte Carlo

The various forms of quantum Monte Carlo implement the techniques outlined above in different ways. Variational Monte Carlo is very direct: Monte Carlo methods are used to evaluate expressions like

$$E_T = \frac{\int \Psi_T^*(\mathbf{R}) \hat{H} \Psi_T(\mathbf{R})\, d\mathbf{R}}{\int \Psi_T^*(\mathbf{R}) \Psi_T(\mathbf{R})\, d\mathbf{R}}. \tag{3.12}$$

E_T is an energy expectation value for a system with Hamiltonian operator $\hat{H}$; $\Psi_T(\mathbf{R})$ is a trial wave function depending on all the spatial coordinates,[2] which are represented here by the vector $\mathbf{R}$. The expectation value is variational: $E_T \geq E_0$, where E_0 is the exact ground-state energy. As the quality of the trial wave function improves, E_T becomes closer to E_0. The two become equal when $\Psi_T \propto \Psi_0$, where Ψ_0 is the ground-state wave function.

Consider a trial wave function close to the exact ground-state eigenfunction:

$$\Psi_T = \Psi_0 + \sum_{i>0} \epsilon_i \Psi_i \tag{3.13}$$

so that the coefficients $\{\epsilon_i\}$ are small. The energy expectation value is then

$$E_T = E_0 + \sum_{i>0} |\epsilon_i|^2 (E_i - E_0) + \mathcal{O}[\epsilon_i^4]. \tag{3.14}$$

The error in the energy is of order ϵ_i^2; this demonstrates that the energy expectation value for a given trial wave function is more accurate than the wave function itself.

The quality of the trial wave function is clearly very important. The subject of the form and optimisation of the trial wave function will be discussed in more detail later in this thesis.

[2] In many cases, it is convenient to think of the electron spins as fixed; the spin-dependence of the wave function will be suppressed from now on.

3.2.1 Importance sampling in VMC

To implement the procedure described in section 3.1, the integrand of equation (3.12) must be written as the product of score and importance functions. The prescription of equation (3.10) for the ideal importance function is

$$P(\mathbf{R}) = \frac{|\Psi_T^*(\mathbf{R})\hat{H}\Psi_T(\mathbf{R})|}{\int |\Psi_T^*(\mathbf{R}')\hat{H}\Psi_T(\mathbf{R}')|\, d\mathbf{R}'}, \tag{3.15}$$

which is as difficult to calculate as E_T itself. However, consider what happens when Ψ_T is equal to Ψ_0. Then

$$\hat{H}\Psi_0 = E_0\Psi_0 \tag{3.16}$$

and the optimal importance function is

$$P(\mathbf{R}) = \frac{|\Psi_T(\mathbf{R})|^2}{\int |\Psi_T(\mathbf{R}')|^2\, d\mathbf{R}'}. \tag{3.17}$$

So, for a trial wave function of good quality, the most efficient choice of importance function is almost identical to the natural one: $P(\mathbf{R}) \propto |\Psi_T(\mathbf{R})|^2$.

The score function corresponding to this choice is

$$f(\mathbf{R}) = \frac{\hat{H}\Psi_T(\mathbf{R})}{\Psi_T(\mathbf{R})}. \tag{3.18}$$

This function is known as the *local energy* and is denoted $E_L(\mathbf{R})$.

The VMC estimate of the ground-state energy is then

$$E_{\mathrm{VMC}} = \frac{1}{M}\sum_{i=1}^{M} E_L(\mathbf{R}_i), \tag{3.19}$$

where the $\{\mathbf{R}_i\}$ are drawn from a probability distribution proportional to $|\Psi_T(\mathbf{R})|^2$. The next challenge is therefore to generate a set of points distributed in this way.

3.2.2 The Metropolis algorithm

A simple and robust way of sampling an arbitrary many-dimensional probability distribution is provided by the Metropolis algorithm [58]. The algorithm prescribes a set of uncomplicated rules for moving a *walker* through configuration space. The

set of vectors $\{\mathbf{R}_i\}$ corresponding to the points in configuration space visited by the walker will be distributed according to the required probability density. To sample a distribution $P(\mathbf{R})$:

1. Start the walker at a random point $\mathbf{R}$.
2. Propose a trial move to a point $\mathbf{R}'$, chosen from some simple probability density function $T(\mathbf{R}' \leftarrow \mathbf{R})$.
3. Accept the move with probability $A(\mathbf{R}' \leftarrow \mathbf{R}) = \min\left(1, \frac{T(\mathbf{R} \leftarrow \mathbf{R}')P(\mathbf{R}')}{T(\mathbf{R}' \leftarrow \mathbf{R})P(\mathbf{R})}\right)$.
4. Record the walker position, whether or not it has changed.
5. If more points are required, return to step 2 and repeat.

Not all the points should be recorded: there is an equilibration period at the start of the walk, during which the points generated depend on the initial position. In addition, neighbouring points on the walk are usually correlated; this means that several moves must be made for each independent sample.

Although in principle any reasonable[3] probability density function can be used for $T(\mathbf{R}' \leftarrow \mathbf{R})$, the choice of function does affect the efficiency of the algorithm. A function which proposes too many large moves produces too many rejections, which is inefficient; in contrast, one which proposes only small moves takes longer to generate uncorrelated points, which is also inefficient. In between these extremes is a function which maximises the sampling efficiency; a commonly-used rule of thumb is to aim for an acceptance rate of 50%.

To gain some insight into the way the algorithm works, consider a population of walkers, with density $n(\mathbf{R})$. Assume that the walkers have reached a steady state, so that all information about starting positions has been lost; also assume that the detailed balance condition is obeyed:

$$n(\mathbf{R})A(\mathbf{R}' \leftarrow \mathbf{R})T(\mathbf{R}' \leftarrow \mathbf{R}) = n(\mathbf{R}')A(\mathbf{R} \leftarrow \mathbf{R}')T(\mathbf{R} \leftarrow \mathbf{R}'). \tag{3.20}$$

[3] $T(\mathbf{R}' \leftarrow \mathbf{R})$ must be ergodic; if $T(\mathbf{R}' \leftarrow \mathbf{R})$ is non-zero then $T(\mathbf{R} \leftarrow \mathbf{R}')$ must also be.

Substituting the expressions for the acceptance probabilities,

$$\frac{n(\mathbf{R}')}{n(\mathbf{R})} = \frac{T(\mathbf{R}' \leftarrow \mathbf{R}) \min\left(1, \frac{T(\mathbf{R} \leftarrow \mathbf{R}')P(\mathbf{R}')}{T(\mathbf{R}' \leftarrow \mathbf{R})P(\mathbf{R})}\right)}{T(\mathbf{R} \leftarrow \mathbf{R}') \min\left(1, \frac{T(\mathbf{R}' \leftarrow \mathbf{R})P(\mathbf{R})}{T(\mathbf{R} \leftarrow \mathbf{R}')P(\mathbf{R}')}\right)} \tag{3.21}$$

$$= \frac{P(\mathbf{R}')}{P(\mathbf{R})}. \tag{3.22}$$

This illustrates that the equilibrium walker density is proportional to the required probability density; the generated points are indeed sampled from this distribution.

3.2.3 Advantages and disadvantages of VMC

The degree of success of any VMC calculation is heavily dependent on the quality of the trial wave function: a bad wave function will lead to a bad estimate of the energy. The variational nature of the calculation makes it useful for establishing an upper bound to the ground-state energy; the expectation values of other operators (besides $\hat{H}$) may also be calculated using this method, but these values are not variational. The sampling and accuracy are better for operators $\hat{O}$ for which the commutator $[\hat{H}, \hat{O}] = 0$, since the ground-state wave function is then also an eigenfunction of $\hat{O}$ (if the ground state is not degenerate). The arguments of the preceding sections regarding optimal sampling (leading to equation (3.17)) and the error in the estimated energy (leading to equation (3.14)) may then be duplicated, replacing $\hat{H}$ with $\hat{O}$.

The main disadvantage of VMC, the total reliance on the trial wave function, does not apply to diffusion Monte Carlo.

3.3 Diffusion Monte Carlo

The Diffusion Monte Carlo method provides a means of improving an estimated ground-state wave function. It is a projector method: in theory, the ground-state component of any trial wave function is projected out, giving the exact ground state (subject to statistical errors). The factors which render this ideal projection unachievable will be described in later chapters.

The method is based on the similarity between the imaginary-time Schrödinger equation and the equation which describes classical diffusion.

3.3.1 Simulating classical diffusion

The classical diffusion equation is

$$\frac{\partial f(\mathbf{x},t)}{\partial t} = D\nabla^2 f(\mathbf{x},t). \tag{3.23}$$

It describes the time evolution of a gas of particles with density f. With no external potential, the particles spread out over time. The Green's function for this equation is

$$G(\mathbf{y},\mathbf{x};t) = \frac{1}{(4\pi Dt)^{d/2}} e^{-(\mathbf{y}-\mathbf{x})^2/4Dt}, \tag{3.24}$$

where d is the dimension of the physical space. Substitution confirms that this is a solution of equation (3.23), with the additional property that

$$G(\mathbf{y},\mathbf{x};0) = \delta(\mathbf{y}-\mathbf{x}). \tag{3.25}$$

The Green's function encapsulates all the details of the propagation of the density in time: if the initial density is $f(\mathbf{x},0)$, then the density at time t is

$$f(\mathbf{y},t) = \int G(\mathbf{y},\mathbf{x};t) f(\mathbf{x},0)\, d\mathbf{x}. \tag{3.26}$$

One interpretation of the equation is in terms of probabilities: the probability that a particle initially at $\mathbf{x}$ moves to $\mathbf{y}$ after a time t is $G(\mathbf{y},\mathbf{x};t)$. This is important because it suggests a way to simulate the diffusion process.

A set of particles ('walkers') is used, with initial positions $\mathbf{x}_i(0)$ sampled from the probability distribution proportional to $f(\mathbf{x},0)$. A finite time step Δt is chosen; the new positions $\mathbf{x}_i(\Delta t)$ are then calculated as follows:

$$\mathbf{x}_i(\Delta t) = \mathbf{x}_i(0) + \boldsymbol{\xi} \tag{3.27}$$

where $\boldsymbol{\xi}$ is a random variable taken from a Gaussian distribution with zero mean and variance $4D\Delta t$. This distribution is chosen so that the probability density function

for $\boldsymbol{\xi}$ is $G(\boldsymbol{\xi}, \mathbf{0}; \Delta t)$ (compare equation (3.24)). In other words, the probability that a particle moves from $\mathbf{x}$ to $\mathbf{y}$ in the time Δt is $G(\mathbf{y}, \mathbf{x}; \Delta t)$. Thus, equations (3.27) and (3.26) are consistent.

Because G is homogeneous — it depends only on the size of the time step Δt, and not on the absolute time — the same rule for moving walkers may be repeated many times. Since G is known exactly, n steps of size Δt are equivalent to a single step of size $n\Delta t$.

Equation (3.27) is a Langevin equation: it describes the rules for updating a configuration over a time step. The walkers are not real particles: they do not collide with each other or interact in any way. However, after n time steps, the position of each walker samples the probability distribution proportional to $f(\mathbf{x}, n\Delta t)$.

This example is trivial, because equation (3.23) can be solved analytically; this is how the Green's function was obtained. There are many other cases where the Green's function cannot be calculated precisely, but has some small-t approximation, and in these cases, the technique outlined above is very useful for obtaining information about the evolution of f.

3.3.2 Application to quantum mechanics

The diffusion equation, equation (3.23), is similar in form to the imaginary-time Schrödinger equation:

$$\frac{\partial \Psi(\mathbf{R}, \tau)}{\partial \tau} = \left[\frac{1}{2} \nabla^2_{\mathbf{R}} - V(\mathbf{R}) \right] \Psi(\mathbf{R}, \tau). \tag{3.28}$$

Here, the potential V includes the electron-electron interactions, as well as any contribution from the external potential.

Without V, this equation is identical to equation (3.23) (with $D = 1/2$), and can be simulated in the same way, by having walkers move around according to the rules set out in equation (3.27). For this, the wave function Ψ must be real; however, in systems for which the Hamiltonian has time-reversal symmetry, this condition can always be satisfied [57], and all wave functions will be assumed to be real from now on. Note that each walker represents an entire configuration of the system, not just

a single particle; the diffusion takes place in Nd dimensions, where N is the number of particles and d is the dimension of the physical space.

The challenge is to incorporate the potential into the simulation. This can be achieved by assigning a weight w_i to each walker; the weights are allowed to vary as the walker explores configuration space, and ultimately the density of the walker weights (rather than the walkers themselves) is used to represent Ψ. Neglecting the diffusion term in equation (3.28) leaves the following rate equation:

$$\frac{\partial \Psi(\mathbf{R},\tau)}{\partial \tau} = -V(\mathbf{R})\Psi(\mathbf{R},\tau) \tag{3.29}$$

which has the solution

$$\Psi(\mathbf{R},\tau) = \Psi(\mathbf{R},0)e^{-\tau V(\mathbf{R})}. \tag{3.30}$$

The potential has the effect of increasing the wave function where V is negative, and decreasing it where V is positive. When a walker arrives at the position $\mathbf{R}$, its weight should be multiplied by $e^{-\tau V(\mathbf{R})}$, in accordance with equation (3.30). After n steps of size $\Delta\tau$, the walker's weight becomes

$$w_i(n\Delta\tau) = \exp\left[-\Delta\tau \sum_{m=1}^{n} V\big(\mathbf{R}_i(m\Delta\tau)\big)\right]. \tag{3.31}$$

In fact, the weight of a walker should reflect the whole path in configuration space along which it has travelled; however, in a simulation, the complete path of the walker is not specified because walkers must move in discrete steps. This is a very important point: the procedure described here is valid only for small time steps, when the potential V can be assumed to remain constant over the course of a move. In the limit of infinitely small time steps, the walker weight becomes

$$\lim_{\substack{\Delta\tau\to 0 \\ n\Delta\tau=\tau}} w_i(\tau) = \exp\left[-\int_0^{\tau} V\big(\mathbf{R}_i(\tau')\big)\, d\tau'\right], \tag{3.32}$$

as it should.

The difference between the simple diffusion equation and the imaginary-time Schrödinger equation is that the Green's function for the latter is not known precisely. The simulation technique outlined here is equivalent to using an approximation to the Green's function which is valid for short time steps.

3.3.3 The ground-state wave function

In the previous section, a technique for simulating the development of a wave function in imaginary time was established. To see why this is useful, consider the formal solution of the Schrödinger equation in imaginary time:

$$|\Psi(\tau)\rangle = e^{-\tau\hat{H}}|\Psi(0)\rangle. \tag{3.33}$$

In terms of the eigenfunctions of $\hat{H}$, the initial state may be written

$$\Psi(\mathbf{R}, 0) = \sum_{i=0}^{\infty} c_i \Phi_i(\mathbf{R}), \tag{3.34}$$

which, on substitution into equation (3.33), gives

$$\Psi(\mathbf{R}, \tau) = \sum_{i=0}^{\infty} c_i e^{-\tau E_i} \Phi_i(\mathbf{R}). \tag{3.35}$$

This shows that as τ increases, the eigenstate with the lowest energy provides the dominant contribution to $\Psi(\mathbf{R}, \tau)$. In the limit $\tau \to \infty$, only the ground-state solution Φ_0 remains (as long as $c_0 \neq 0$). Any wave function which has a non-vanishing overlap with Φ_0 therefore evolves to the ground state in the limit of large imaginary time. The value of the imaginary-time simulation described in the previous section is that it allows the ground state wave function to be determined, thus reducing the difficult quantum many-body problem of finding the ground-state wave function to a much simpler problem in classical particle dynamics.

However, since each time step is required to be short, many steps are required before the large-τ limit is reached. This is not the only problem with this method:

- the wave function has been tacitly assumed to be positive, like the particle density in section 3.3.1, and there is no mechanism for the weight of a walker to change sign;
- the walkers are free to diffuse around: this is at best inefficient, because time is wasted sampling unimportant regions of configuration space, and at worst disastrous, as in a finite system, where the walkers are effectively able to leave the system;

- the weights are liable to vary wildly, since V is not a well-behaved function; sampling is dominated by one or two walkers with much larger weight than the others.

Fortunately, these problems can be addressed.

3.3.4 Importance-sampled diffusion Monte Carlo

The simple DMC algorithm can be dramatically improved by the use of importance sampling, which was introduced in section 3.1 and applied to VMC in section 3.2.1.

The first step is to multiply equation (3.28) on the left with a trial wave function $\Psi_T(\mathbf{R})$:

$$\Psi_T(\mathbf{R})\frac{\partial\Psi(\mathbf{R},\tau)}{\partial\tau} = -\Psi_T(\mathbf{R})(\hat{H} - E_T)\Psi(\mathbf{R},\tau). \tag{3.36}$$

An energy shift E_T has also been introduced; this will be used later to control the walker weights. Without the energy shift, equation (3.35) shows that the wave function ultimately decays (or grows) exponentially as $e^{-\tau E_0}$; introducing E_T and setting $E_T \approx E_0$ allows this to be avoided.

The new analogue of the particle density is the product of the trial wave function with the solution of the Schrödinger equation:

$$f(\mathbf{R},\tau) = \Psi_T(\mathbf{R})\Psi(\mathbf{R},\tau). \tag{3.37}$$

Expressing equation (3.36) in terms of f is a matter of algebra, which eventually gives

$$\frac{\partial f}{\partial\tau} = \frac{1}{2}\nabla^2 f - \nabla\cdot(\mathbf{v}f) + (E_T - E_L)f \tag{3.38}$$

where a new quantity has been introduced: the *drift velocity*

$$\mathbf{v}(\mathbf{R}) = \frac{1}{2}\nabla \ln\left(|\Psi_T(\mathbf{R})|^2\right). \tag{3.39}$$

The local energy

$$E_L(\mathbf{R}) = \frac{\hat{H}\Psi_T(\mathbf{R})}{\Psi_T(\mathbf{R})} \tag{3.40}$$

was introduced in section 3.2.1, and represents the energy of a configuration $\mathbf{R}$ calculated with respect to the trial wave function. Because the trial wave function

is not time-dependent, the dependence of f on τ is entirely contained in Ψ, which has been shown to converge to the ground-state wave function Ψ_0 in the limit of large τ. It follows that the form of f in this limit is

$$f(\mathbf{R}, \tau \to \infty) = c_0 e^{-\tau(E_0 - E_T)} \Psi_T(\mathbf{R}) \Phi_0(\mathbf{R}). \tag{3.41}$$

Equation (3.38) may be used to construct a practical, efficient simulation algorithm.

Unlike the original imaginary-time Schrödinger equation, which resembles a model of diffusion, the importance-sampled version is very similar to the equation of a drift-diffusion process. In this comparison, $f(\mathbf{R}, \tau)$ represents a particle density, whose evolution in (real) time τ is given by equation (3.38). The first term on the right-hand side of this equation is the familiar diffusion term; the diffusion constant D again has value $1/2$. The second term is new. In the diffusion picture, it corresponds to a drift of particles, where the drift velocity is $\mathbf{v}$: the particles no longer diffuse freely, but are now guided in their motion. The remaining term, $(E_T - E_L)f$, has the same form as the potential term $V\Psi$ in the non-importance-sampled model: it is a rate term, which acts to increase or decrease the density exponentially, depending on the sign of $(E_T - E_L)$. One of the advantages of using importance sampling is now clear: the rate term depends on the local energy E_L, which is generally a much better-behaved function than the potential V.

The only difference in form between equations (3.38) and (3.28) is the drift term. This suggests that in order to sample the evolution of the new function f, almost the same techniques as before may be used: a set of classical particles is allowed to diffuse around, each carrying a weight which increases or decreases depending on the local energy; this time, however, the particles are also subject to drift.

To put this procedure on a more mathematical footing, it is helpful to investigate the Green's function for the problem. One way of defining this function is via the propagation in time of f:

$$f(\mathbf{R}, \tau) = \int \tilde{G}(\mathbf{R}, \mathbf{R}'; \tau) f(\mathbf{R}', 0)\, d\mathbf{R}'. \tag{3.42}$$

The Green's function for the original problem, the imaginary-time Schrödinger equa-

tion, may be defined in the same way:

$$\Psi(\mathbf{R},\tau) = \int G(\mathbf{R},\mathbf{R}';\tau)\Psi(\mathbf{R}',0)\,d\mathbf{R}'. \tag{3.43}$$

Comparing these two equations and using the definition of f gives a simple relationship between the original and modified Green's functions:

$$\tilde{G}(\mathbf{R},\mathbf{R}';\tau) = \Psi_T(\mathbf{R})G(\mathbf{R},\mathbf{R}';\tau)\frac{1}{\Psi_T(\mathbf{R}')}. \tag{3.44}$$

This demonstrates that $\tilde{G}$ is not symmetric (since G is): this lack of symmetry is a consequence of the fact that the operator

$$\hat{F} = \frac{1}{2}\nabla^2 - \nabla\cdot\mathbf{v} - \mathbf{v}\cdot\nabla + E_T - E_L, \tag{3.45}$$

which appears on the right-hand side of equation (3.38), is not Hermitian.

Of course, the exact analytical form of $\tilde{G}$ is not known. However, as before, an approximation which is valid for small τ can be used. A formal expression for $\tilde{G}$ is the following:

$$\tilde{G}(\mathbf{R}',\mathbf{R};\tau) = \left\langle \mathbf{R}' \middle| e^{\tau\hat{F}} \middle| \mathbf{R} \right\rangle = \left\langle \mathbf{R}' \middle| e^{-\tau(\hat{\mathcal{T}}+\hat{\mathcal{V}})} \middle| \mathbf{R} \right\rangle \tag{3.46}$$

where two new operators have been introduced for convenience:

$$\hat{\mathcal{T}} = -\frac{1}{2}\nabla^2 + (\nabla\cdot\mathbf{v}) + \mathbf{v}\cdot\nabla \tag{3.47}$$

$$\hat{\mathcal{V}} = E_L - E_T. \tag{3.48}$$

The operators $\hat{\mathcal{T}}$ and $\hat{\mathcal{V}}$ represent modified kinetic and potential energies respectively. The kinetic energy operator is associated with the drift-diffusion process, while the potential energy operator comes from the rate equation. Using the Trotter-Suzuki formula,[4] the Green's function can be approximately factorised:

$$\begin{aligned}\tilde{G}(\mathbf{R}',\mathbf{R};\tau) &= \left\langle \mathbf{R}' \middle| e^{-\frac{1}{2}\tau\hat{\mathcal{V}}} e^{-\tau\hat{\mathcal{T}}} e^{-\frac{1}{2}\tau\hat{\mathcal{V}}} \middle| \mathbf{R} \right\rangle + \mathcal{O}\left[\tau^3\right] \\ &= e^{-\frac{1}{2}\tau\mathcal{V}(\mathbf{R}')} \left\langle \mathbf{R}' \middle| e^{-\tau\hat{\mathcal{T}}} \middle| \mathbf{R} \right\rangle e^{-\frac{1}{2}\tau\mathcal{V}(\mathbf{R})} + \mathcal{O}\left[\tau^3\right] \\ &= W(\mathbf{R}',\mathbf{R};\tau)G_D(\mathbf{R}',\mathbf{R};\tau) + \mathcal{O}\left[\tau^3\right]\end{aligned} \tag{3.49}$$

[4]The formula gives an approximate form for the exponential of a sum of two operators: $e^{-\tau(\hat{A}+\hat{B})} = e^{-\tau\hat{B}/2}e^{-\tau\hat{A}}e^{\tau\hat{B/}2} + \mathcal{O}\left[\tau^3\right]$.

where the contribution to $\tilde{G}$ from the drift-diffusion process is

$$G_D(\mathbf{R}', \mathbf{R}; \tau) = \left\langle \mathbf{R}' \middle| e^{-\tau \hat{\mathcal{T}}} \middle| \mathbf{R} \right\rangle \tag{3.50}$$

and the part arising from the rate term is

$$W(\mathbf{R}', \mathbf{R}; \tau) = e^{-\frac{1}{2}\tau(\mathcal{V}(\mathbf{R}')+\mathcal{V}(\mathbf{R}))}. \tag{3.51}$$

The use of the letter W anticipates the association of this term with the walker weight.

Inspection of equation (3.50) shows that G_D is the Green's function for the drift-diffusion equation. Once again, it cannot be evaluated analytically, but it may be estimated by making the further approximation that the drift velocity is constant over the time τ. The result is

$$G_D(\mathbf{R}', \mathbf{R}; \tau) = (2\pi\tau)^{-d/2} \exp\left(-\frac{(\mathbf{R}' - \mathbf{R} - \tau\mathbf{v}(\mathbf{R}))^2}{2\tau}\right) + \mathcal{O}\left[\tau^2\right]. \tag{3.52}$$

Here d is the dimension of configuration space. This approximation for G_D allows a Langevin equation to be written down for the drift-diffusion process. No weights are required, because drift and diffusion conserve the total density; the probability that a walker moves from $\mathbf{R}$ to $\mathbf{R}'$ in time $\Delta\tau$ is given by $G_D(\mathbf{R}, \mathbf{R}'; \Delta\tau)$. The Langevin equation is therefore

$$\mathbf{R}_i(\tau + \Delta\tau) = \mathbf{R}_i(\tau) + \mathbf{v}\Delta\tau + \boldsymbol{\xi} \tag{3.53}$$

where $\boldsymbol{\xi}$ is a random variable taken from an Nd-dimensional Gaussian distribution with zero mean and variance $2\Delta\tau$.

A population of walkers whose dynamics are specified by equation (3.53) samples the density f_D, where

$$\frac{\partial f_D}{\partial \tau} = -\hat{\mathcal{T}} f_D. \tag{3.54}$$

The steady-state solution of this equation is

$$f_D \propto \Psi_T^2, \tag{3.55}$$

which can be verified by substitution; the operator $\hat{\mathcal{T}}$ guides the walkers so that their distribution follows the square of the trial wave function. This is importance sampling: walkers are concentrated in important regions (where the magnitude of the wave function is large). Allowing walkers to drift and diffuse in this manner is therefore a sampling technique in its own right, and can be used as an alternative to the Metropolis method in VMC.

However, the aim here is not simply to sample Ψ_T^2, but to sample $\Psi_T\Phi_0$. This is achieved by introducing weights. Initially, the walker weights are set to unity; then, when a walker moves from $\mathbf{R}$ to $\mathbf{R}'$ during the time step τ, its weight is multiplied by $W(\mathbf{R}, \mathbf{R}'; \Delta\tau)$. To see that this works, consider the density function for the walker weights at $\mathbf{R}$ after a single move:

$$\begin{aligned} f_{\text{sim}}(\mathbf{R}, \Delta\tau) &= \int W(\mathbf{R}, \mathbf{R}'; \Delta\tau) G_D(\mathbf{R}, \mathbf{R}'; \Delta\tau) f(\mathbf{R}, 0)\, d\mathbf{R}' \\ &\approx \int \tilde{G}(\mathbf{R}, \mathbf{R}'; \Delta\tau) f(\mathbf{R}, 0)\, d\mathbf{R}' \\ &\approx f(\mathbf{R}, \Delta\tau). \end{aligned} \tag{3.56}$$

The initial distribution of walkers is $f(\mathbf{R}, 0)$; the probability that a walker moves from $\mathbf{R}'$ to $\mathbf{R}$ during the time step $\Delta\tau$ is $G_D(\mathbf{R}, \mathbf{R}'; \Delta\tau)$, and the weight associated with such a move is $W(\mathbf{R}, \mathbf{R}', \Delta\tau)$. The combination of weights and drift-diffusion therefore accurately simulates the evolution of f, as long as the time step $\Delta\tau$ is small.

In order to reach the desired limit of large τ, many small steps of duration $\Delta\tau$ must be carried out. If, during this time, a walker spends a lot of time in a region of space where $E_L < E_T$, its weight continues to increase, and may become very large. In contrast, if a walker spends a lot of time in a region where $E_L > E_T$, its weight may become very small, and it contributes little to the ultimate sampling of f, which is dominated by walkers with large weights. This is one of the problems alluded to in section 3.3.3, and makes the sampling inefficient.

The solution to this problem is to allow walkers to multiply or die out according to their weight. Either when the weight of a walker becomes too large or too small, or at regular intervals, the walker is replaced by a certain number of 'descendants'

of unit weight. The number of descendants depends on the weight of the walker: walkers with more weight should produce more descendants. The correct distribution of weights is preserved by using the following formula to calculate the number of descendants:

$$n_i = \mathrm{INT}(w_i + \eta), \tag{3.57}$$

where w_i is the weight of the original walker and η is a random variable sampled from $U[0,1]$. The function INT returns the truncated integer part of its argument, so that a walker of initial weight 2.6 has a 60% probability of generating three descendants and a 40% probability of generating two. It is possible for walkers with low weight ($w_i < 1$) to produce no descendants; this happens with probability $100(1 - w_i)\%$. These walkers are simply removed from the simulation.

This technique is known as *branching*, and greatly improves the sampling efficiency. Walkers proliferate in regions where $E_L < E_T$, and die out in regions where $E_L > E_T$, but the distribution of weights is now much narrower than before. No time is wasted moving walkers which ultimately contribute almost nothing to the result; instead, the computational effort is concentrated in more useful areas.

There is still a problem: the total number of walkers depends on the average value of $E_L - E_T$. If $E_L \neq E_T$, the population of walkers either grows or decays exponentially in time. The reason for introducing the energy shift (or 'trial energy') E_T is now clear: by adjusting the value of E_T during a simulation, the population can be controlled. One way of achieving this [79] is to modify E_T according to

$$E_T(\tau) = E_{\mathrm{est}}(\tau) - \frac{1}{\tau_g} \ln\left(\frac{M(\tau)}{M_0}\right) \tag{3.58}$$

where $E_{\mathrm{est}}(\tau)$ is a current estimate of the ground-state energy, τ_g and M_0 are constants, and $M(\tau)$ is the total weight:[5]

$$M(\tau) = \sum_i w_i(\tau). \tag{3.59}$$

The quantity M_0 may be viewed as a target population; the procedure aims to bring $M(\tau)$ back to this value over a time-scale set by τ_g. Unless Ψ_T is an exact

[5]If branching is carried out at every time step, M is simply the total number of walkers.

eigenfunction of $\hat{H}$, the local energy is subject to fluctuations; as a result, the walker population also fluctuates. Using equation (3.58) allows the population fluctuations to be kept under control; simply setting $E_T = E_0$ would not achieve this, even though the average population would remain constant. In addition, of course, the precise value of E_0 is not known — calculating E_0 is usually one of the aims of the simulation.

Modifying the reference energy introduces a bias into the sampling, because the equation being simulated is no longer (3.38), and therefore has different eigenfunctions. The bias can be reduced by making τ_g as large as possible, although this leads to greater population fluctuations.

Even when τ_g is large and the population-control bias is negligible, there remain sampling errors, caused by the use of an approximate Green's function. The true Green's function satisfies a form of detailed balance; using equation (3.44), and noting that $G(\mathbf{R}, \mathbf{R}'; \Delta\tau)$ is symmetric in $\mathbf{R}$ and $\mathbf{R}'$ gives

$$\Psi_T^2(\mathbf{R}')\tilde{G}(\mathbf{R}, \mathbf{R}'; \Delta\tau) = \Psi_T^2(\mathbf{R})\tilde{G}(\mathbf{R}', \mathbf{R}; \Delta\tau). \tag{3.60}$$

This is not the usual version of detailed balance, because $\tilde{G}$ does not have the form of a probability; specifically, the total density is not conserved between moves, and may increase or decrease.

The approximate Green's function does not satisfy equation (3.60). This can be traced back to the expression for G_D, equation (3.52), which is only correct to $\mathcal{O}[\Delta\tau]$. The sampling error can be reduced by making the time step $\Delta\tau$ smaller; however, this is inefficient, because many more steps are required between uncorrelated configurations. A better way of reducing the time-step error is to enforce the detailed balance condition by introducing a Metropolis rejection step, as described in section 3.2.2, with an acceptance probability now given by

$$A(\mathbf{R}' \leftarrow \mathbf{R}) = \min\left(1, \frac{\tilde{G}(\mathbf{R}, \mathbf{R}'; \Delta\tau)\Psi_T^2(\mathbf{R}')}{\tilde{G}(\mathbf{R}', \mathbf{R}; \Delta\tau)\Psi_T^2(\mathbf{R})}\right). \tag{3.61}$$

For small time steps, the rejection probability tends to zero, since the approximate drift-diffusion Green's function becomes more exact. However, for non-zero time

steps, a small number of moves will be rejected; consequently, the effective duration of a walker move is reduced. An effective time step, smaller than $\Delta\tau$, must then be used in the calculation of the branching factor.

A more serious issue is encountered when simulating fermionic systems.

3.3.5 Fermions

The ground-state wave function of a fermionic system is required to be antisymmetric. However, this information is not built into the Hamiltonian, and in general, the lowest energy eigenvalue corresponds to the completely symmetric bosonic ground-state wave function Φ_0^{B}. Any wave function with a non-zero projection on Φ_0^{B}, when allowed to evolve in imaginary time according to the Schrödinger equation (3.28), eventually tends to Φ_0^{B}; this is a problem which must be overcome in order to simulate fermionic systems.

The most commonly-employed solution follows naturally from the importance-sampling transformation. The sampled function f is forced to be non-negative everywhere, because the walker weights are initially positive and have no mechanism for changing sign; then, because $f = \Psi_T\Psi$, the Schrödinger wave function Ψ must have the same sign as the trial wave function Ψ_T. Regions of (Nd-dimensional) space in which $\Psi_T > 0$ are separated from those in which $\Psi_T < 0$ by the *nodal surface*. Since Ψ and Ψ_T have the same sign everywhere, they must also share the same nodal surface. If a walker crosses the nodal surface, the trial wave function changes sign; by testing for this, and rejecting such moves,[6] the walkers are prevented from leaving the nodal pocket in which they currently sit. This constitutes the *fixed-node approximation*: the nodes of Ψ are forced to be the same as those of Ψ_T.

In this approximation, the simulation proceeds independently in the various occupied nodal pockets. The large-imaginary-time limit of f is then proportional to

[6]The drift velocity diverges at the nodes; if the true Green's function were used, then no walker would ever attempt to cross a node. The fact that walkers do attempt to cross the nodes is a consequence of the small-τ approximation to G_D.

$\Psi_T \Phi_0^{\mathrm{FN}}$, where Φ_0^{FN} is known as the fixed-node ground state. It may be shown [75, 60] that the fixed-node energy is variational: that is, $E_0^{\mathrm{FN}} \geq E_0$, where E_0 is the energy of the true fermionic ground state. Typically, for the trial wave functions used in QMC, all the nodal pockets are equivalent [23], so that the calculated energy does not depend on which pockets are populated. The errors associated with the fixed-node approximation will be mentioned in chapter 4, along with some of the techniques which aim to go beyond it.

The nodes of the wave function cause other problems, because both the drift velocity and the local energy diverge here. The approximation for the drift-diffusion Green's function G_D uses the fact that the potential energy of the system does not change much during the course of a move. However, near the nodal surface, the move size can become large (because $\mathbf{v}$ diverges) and the energy can change rapidly; the result is that the approximation is no longer a good one. A better approximation can be obtained by limiting both the drift velocity and the local energy [79].

3.3.6 Estimators

In the preceding sections, the fixed-node diffusion Monte Carlo method has been described; the result of applying this technique is a set of walkers with weights distributed according to $\Psi_T \Phi_0^{\mathrm{FN}}$, where the fixed-node ground state Φ_0^{FN} is usually a good approximation to the true ground state Φ_0. For the method to be useful, these walkers and weights must provide a way of estimating operator expectation values; this is the link between simulation and measurable reality.

Two estimators of the ground-state energy have in fact been described already: E_T, the trial energy, and $\langle E_L \rangle$, the average local energy. The expectation value of the local energy in the limit of large imaginary time is

$$\lim_{\tau\to\infty} \left\langle \sum_i E_L\big(\mathbf{R}_i(\tau)\big) w_i(\tau) \right\rangle = \lim_{\tau\to\infty} \int f(\mathbf{R},\tau) E_L(\mathbf{R})\, d\mathbf{R}. \tag{3.62}$$

Substituting for f and E_L, and using the fact that $\hat{H}$ is Hermitian then gives

$$\begin{aligned}\lim_{\tau\to\infty}\left\langle \sum_i E_L\big(\mathbf{R}_i(\tau)\big)w_i(\tau)\right\rangle &\approx \int c_0 e^{-\tau(E_0-E_T(\tau))}\Psi_T(\mathbf{R})\Phi_0(\mathbf{R})\frac{\hat{H}\Psi_T(\mathbf{R})}{\Psi_T(\mathbf{R})}\,d\mathbf{R} \\ &= c_0 e^{-\tau(E_0-E_T(\tau))}\int \Phi_0(\mathbf{R})\hat{H}\Psi_T(\mathbf{R})\,d\mathbf{R} \\ &= E_0 c_0 e^{-\tau(E_0-E_T(\tau))}\int \Phi_0(\mathbf{R})\Psi_T(\mathbf{R})\,d\mathbf{R}.\end{aligned} \tag{3.63}$$

The aim is to calculate E_0; the other factors on the right-hand side of this equation come from the integration of f, which is approximated by the total weight:

$$\begin{aligned}\lim_{\tau\to\infty}\big\langle M(\tau)\big\rangle &= \lim_{\tau\to\infty}\int f(\mathbf{R},\tau)\,d\mathbf{R} \\ &\approx c_0 e^{-\tau(E_0-E_T(\tau))}\int \Phi_0(\mathbf{R})\Psi_T(\mathbf{R})\,d\mathbf{R}).\end{aligned} \tag{3.64}$$

This implies that the quantity

$$E_{\text{mixed}}(\tau) = \frac{\sum_i E_L\big(\mathbf{R}_i(\tau)\big)w_i(\tau)}{M(\tau)} \tag{3.65}$$

is an estimator for the ground-state energy when τ is large. It is known as the *mixed* estimator, and can be extended to apply to any local operator which commutes with the Hamiltonian (and therefore has Φ_0 as an eigenfunction). The general mixed estimator for an operator $\hat{O}$, where $\big[\hat{H},\hat{O}\big]=0$, is

$$O_{\text{mixed}}(\tau) = \frac{\sum_i O_L\big(\mathbf{R}_i(\tau)\big)w_i(\tau)}{M(\tau)}, \tag{3.66}$$

where

$$O_L(\mathbf{R}) = \frac{\hat{O}\Psi_T(\mathbf{R})}{\Psi_T(\mathbf{R})}. \tag{3.67}$$

Equation (3.64) shows that the total weight grows as $e^{-\tau(E_0-E_T)}$; adjusting E_T so that the population remains roughly constant (using equation (3.58)) ensures that $E_T \approx E_0$, which means that E_T is an alternative estimator for the ground-state energy.

A third estimator for E_0 can also be obtained by considering the evolution of the total walker weight. During a single time step, the expectation value of this quantity changes from $\big\langle M(\tau)\big\rangle$ to $\big\langle M(\tau+\Delta\tau)\big\rangle$, where

$$\big\langle M(\tau+\Delta\tau)\big\rangle = \big\langle M(\tau)\big\rangle e^{-\Delta\tau(E_0-E_T(\tau))}, \tag{3.68}$$

as long as τ is large. Rearranging this formula gives

$$E_0 = E_T(\tau) - \frac{1}{\Delta\tau} \ln \frac{\langle M(\tau + \Delta\tau) \rangle}{\langle M(\tau) \rangle}. \tag{3.69}$$

This is the motivation for the *growth* estimator

$$E_{\text{growth}}(\tau) = E_T(\tau) - \frac{1}{\Delta\tau} \ln \frac{M(\tau + \Delta\tau)}{M(\tau)}. \tag{3.70}$$

More than one time step can be included in the growth estimator; however, E_T is updated after every step, and this effect must then be unravelled. The same 'unravelling' procedure can also be applied to the mixed estimator to account for changes in E_T [79].

When $[\hat{H}, \hat{O}] \neq 0$, the mixed estimator is no longer equal to the ground-state expectation value O_0. However, although now more difficult, it is still possible to construct an estimator for this quantity by combining the mixed estimate with the VMC result, O_{VMC}. The *extrapolated* estimator is

$$O_{\text{ext}} = 2O_{\text{mixed}} - O_{\text{VMC}}. \tag{3.71}$$

To see why this estimator works, consider the integrals which the VMC and mixed estimators attempt to solve:

$$O_{\text{VMC}} \to \frac{\int \Psi_T(\mathbf{R}) \hat{O} \Psi_T(\mathbf{R})\, d\mathbf{R}}{\int \Psi_T^2(\mathbf{R})\, d\mathbf{R}} \tag{3.72}$$

$$O_{\text{mixed}} \to \frac{\int \Phi_0(\mathbf{R}) \hat{O} \Psi_T(\mathbf{R})\, d\mathbf{R}}{\int \Phi_0(\mathbf{R}) \Psi_T(\mathbf{R})\, d\mathbf{R}}. \tag{3.73}$$

Suppose that the trial wave function differs from the true wave function only slightly, so that

$$\Psi_T = \Phi_0 + \Delta\Phi. \tag{3.74}$$

Then substitution shows that the extrapolated estimator samples

$$\begin{aligned} 2\left(\frac{\int \Phi_0(\mathbf{R}) \hat{O} \Psi_T(\mathbf{R})\, d\mathbf{R}}{\int \Phi_0(\mathbf{R}) \Psi_T(\mathbf{R})\, d\mathbf{R}}\right) - \frac{\int \Psi_T(\mathbf{R}) \hat{O} \Psi_T(\mathbf{R})\, d\mathbf{R}}{\int \Psi_T^2(\mathbf{R})\, d\mathbf{R}} &= \frac{\int \Phi_0(\mathbf{R}) \hat{O} \Phi_0(\mathbf{R})\, d\mathbf{R}}{\int \Phi_0^2(\mathbf{R})\, d\mathbf{R}} \\ &\quad + \mathcal{O}\left[(\Delta\Phi)^2\right] \\ &= O_0 + \mathcal{O}\left[(\Delta\Phi)^2\right], \end{aligned} \tag{3.75}$$

so that the error in the extrapolated estimate is of order $(\Delta\Phi)^2$. However, in order for this method to be useful, the trial wave function must be of very high quality.

3.4 Trial wave functions

In VMC, the quality of the trial wave function sets a limit on the accuracy of the calculation. In fixed-node DMC, this is also true to some extent: the nodes are determined by the trial wave function, which therefore also determines the fixed-node error. However, the statistical efficiency of a DMC calculation strongly depends on the quality of the trial wave function, as this determines the importance sampling. In addition, from the computational point of view, evaluating the trial wave function normally constitutes the major part of the calculation. The form of the trial wave function is therefore very important; it must be both accurate and easy to evaluate.

The wave functions normally used in QMC simulations are of the Slater-Jastrow type, consisting of a Slater determinant multiplied by an exponential Jastrow factor:

$$\Psi(\mathbf{X}) = e^{J(\mathbf{X})} D(\mathbf{X}), \tag{3.76}$$

where $D(\mathbf{X})$ is a determinant of one-electron orbitals, exactly as in equation (2.13).

In fact, the computation is made significantly faster by the use of wave functions of the form

$$\Psi(\mathbf{X}) = e^{J(\mathbf{X})} D^{\uparrow}(\mathbf{R}^{\uparrow}) D^{\downarrow}(\mathbf{R}^{\downarrow}). \tag{3.77}$$

The spin-dependence of the one-electron orbitals in the determinants has been removed, and the evaluation of the two smaller determinants $D^{\uparrow}$ and $D^{\downarrow}$ is more efficient than that of the large determinant D. This function is no longer antisymmetric on exchange of electrons with opposite spins; however, the expectation value of any spin-independent operator is unaffected by this alteration.

The single-electron orbitals may be obtained from density-functional theory or Hartree-Fock calculations. The optimal orbitals in these two mean-field schemes are usually very similar. The nodal surface of the resulting trial wave function is completely defined by these orbitals, since the Jastrow factor is never zero; for DMC calculations, this means the expectation value of the energy is not affected by including the Jastrow factor. However, the variance of the energy is affected, and the Jastrow factor is an important part of the trial wave function.

3.4.1 The Jastrow factor

In section 2.3 it was shown that a single-determinant wave function takes account of exchange but not of correlation; the Jastrow factor allows correlation effects to be incorporated.

The most important correlations are those involving pairs of electrons. These are included by having a term of the form

$$-\sum_{i>j} u_{\sigma_i,\sigma_j}(|\mathbf{r}_i - \mathbf{r}_j|) \tag{3.78}$$

in the Jastrow exponent $J(\mathbf{X})$. Recall that the single-determinant wave function does nothing to prevent electrons of opposite spin from coming together; this term keeps these electrons apart, resulting in a significant lowering of energy. Electrons of like spin are also kept apart more than before, although this affects the energy less dramatically.

The two-body term of equation (3.78) does not simply keep electrons apart. Both the long- and short-range behaviour of u are constrained by theoretical arguments.

When two electrons approach each other, the Coulomb energy diverges; for a wave function to be an eigenstate of $\hat{H}$, this divergence must be cancelled by a corresponding divergence in the kinetic energy. Such a divergence is produced by cusps in the wave function: discontinuities in the first derivative with respect to the distance between the electrons. A full discussion of the cusp conditions is given in appendix B.

The long-range behaviour of u may be determined by arguments based on the random phase approximation of Bohm and Pines [9], and is the subject of chapter 7. A connection is made between the long-range electron-electron correlations and the long-wavelength density fluctuations known as plasmons; for a homogeneous system, the resulting u function has the form $1/\omega_p|\mathbf{r}_i - \mathbf{r}_j|$ in the limit $|\mathbf{r}_i - \mathbf{r}_j| \to \infty$, where $\omega_p = \sqrt{4\pi n}$ is the plasma frequency.

A function which combines the required short- and long-range behaviour is

$$u_{\sigma_i\sigma_j}(|\mathbf{r}_i - \mathbf{r}_j|) = \frac{1}{\omega_p|\mathbf{r}_i - \mathbf{r}_j|}\Big(1 - e^{-|\mathbf{r}_i - \mathbf{r}_j|/F_{\sigma_i\sigma_j}}\Big), \tag{3.79}$$

where the cusp conditions require that $F_{\sigma_i\sigma_j} = \sqrt{(1+\delta_{\sigma_i\sigma_j})/\omega_p}$.

The introduction of the u function inevitably modifies the density (except in a homogeneous system). However, it is often the case that the original density is very close to the true value, especially if it is derived from density-functional theory calculations. It is therefore desirable to restore the original density profile, and this motivates the introduction of a one-body term,

$$J_1(\mathbf{X}) = \sum_i \chi(\mathbf{r}_i). \tag{3.80}$$

Many and various forms of the function χ are in use, but the primary aim is always to restore the desired one-electron density. With this in mind, a useful estimate [21] of the optimal function is

$$\chi(\mathbf{r}) \propto \ln\left(\frac{\rho_0}{\rho_u}\right)^{1/2}. \tag{3.81}$$

Here ρ_0 is the original density, obtained before the introduction of the Jastrow factor; ρ_u is the density obtained after the introduction of the two-body term, but before the introduction of χ.

3.4.2 Optimisation

In practice, both one- and two-body terms include variational parameters which are optimised to generate the best possible Jastrow factor.[7] It would seem natural to optimise the Jastrow factor by minimising the variational energy produced by a VMC simulation,

$$E_V(\alpha) = \frac{\int \Psi^2(\mathbf{X};\alpha) E_L(\mathbf{X};\alpha)\, d\mathbf{X}}{\int \Psi^2(\mathbf{X};\alpha)\, d\mathbf{X}}, \tag{3.82}$$

where α represents the set of variational parameters and E_L is the local energy defined in section 3.3.4. However, it is more common to minimise the variance of the local energy:

$$\sigma^2_{E_L}(\alpha) = \frac{\int \Psi^2(\mathbf{X};\alpha)(E_L(\mathbf{X};\alpha) - E_V(\alpha))^2\, d\mathbf{X}}{\int \Psi^2(\mathbf{X};\alpha)\, d\mathbf{X}}. \tag{3.83}$$

[7] Three-body terms are also often included [36].

When the trial wave function is equal to an eigenfunction of $\hat{H}$, this variance is reduced to zero. Minimising the variance, rather than the energy, has several advantages, the most important being that it is much more numerically stable.

Variance minimisation first appeared in the 1930s [82, 7] and was used in numerical optimization of trial wave functions by Conroy in the 1960s [16, 14, 15]. It was originally applied to VMC by Coldwell [13], but became popular largely thanks to the efforts of Umrigar and coworkers [80]. Several modifications of the original scheme exist and are in use [40].

The usual way to carry out the minimisation involves correlated sampling. A set of configurations is generated by a VMC calculation, using an initial set of parameter values α_0. The variance for a new set of parameter values α (close to the original set) is then

$$\sigma^2_{E_L}(\alpha) = \frac{\int \Psi^2(\mathbf{X};\alpha_0) w(\alpha,\alpha_0)(E_L(\mathbf{X};\alpha) - E_V(\alpha))^2 \, d\mathbf{X}}{\int \Psi^2(\mathbf{X};\alpha_0) w(\alpha,\alpha_0) \, d\mathbf{X}} \tag{3.84}$$

where the variational energy is now calculated as

$$E_V(\alpha) = \frac{\int \Psi^2(\mathbf{X};\alpha_0) w(\alpha,\alpha_0) E_L(\mathbf{X};\alpha) \, d\mathbf{X}}{\int \Psi^2(\mathbf{X};\alpha_0) w(\alpha,\alpha_0) \, d\mathbf{X}} \tag{3.85}$$

and a weighting factor has been introduced:

$$w(\alpha,\alpha_0) = \frac{\Psi^2(\mathbf{X};\alpha)}{\Psi^2(\mathbf{X};\alpha_0)}. \tag{3.86}$$

The number of configurations is, of course, finite, and the integrals indicated here are approximated by finite sums.

The advantage of the correlated sampling approach is that, in theory, only one set of configurations needs to be generated. In practice, it is almost always necessary to generate more than one set of configurations. This is because the minimisation process may become numerically unstable; this instability is characterised by a few configurations acquiring very large weights, and leads to incorrect results [40]. Thus, once the variance of the weights reaches a certain level, it is normal to regenerate the configurations (and therefore reset all weights to unity) [22]. It is often preferable (particularly in large systems) simply to set all the weights equal to unity [76, 83],

and regenerate the configurations once the minimum has been found. Despite the difficulties involved in these schemes, they are far more efficient than regenerating configurations for each set of parameters.

Variance minimisation is not infallible, and several alternatives have been proposed; these include different optimisation functionals [2, 11] or the method of stochastic gradients [31].

Chapter 4

Errors in QMC simulations

This chapter will describe the most significant errors inherent in extended-system QMC simulations. Two of these — statistical noise and the fixed-node error — have already been mentioned in the previous chapter; when attempting to simulate extended systems, finite-size errors also become important.[1]

4.1 Finite-size errors

The QMC methods described in Chapter 3 apply to finite systems. To investigate the properties of materials which have infinite extent in at least one spatial dimension, some kind of scheme for extrapolation is required.

In practice, such materials are studied by placing a finite set of particles in a simulation cell, which is then subjected to periodic boundary conditions. This defines a lattice (see figure 4.1). To complicate things, the particles must also undergo interactions with their periodically-repeated images; the best way to take account of this is not obvious. In this report, the particles will be exclusively electrons, interacting via the Coulomb force, the long-ranged nature of which is a key part of the problem.

[1]The use of pseudopotentials introduces additional errors in the simulation of real atoms, but these will not be discussed in this work.

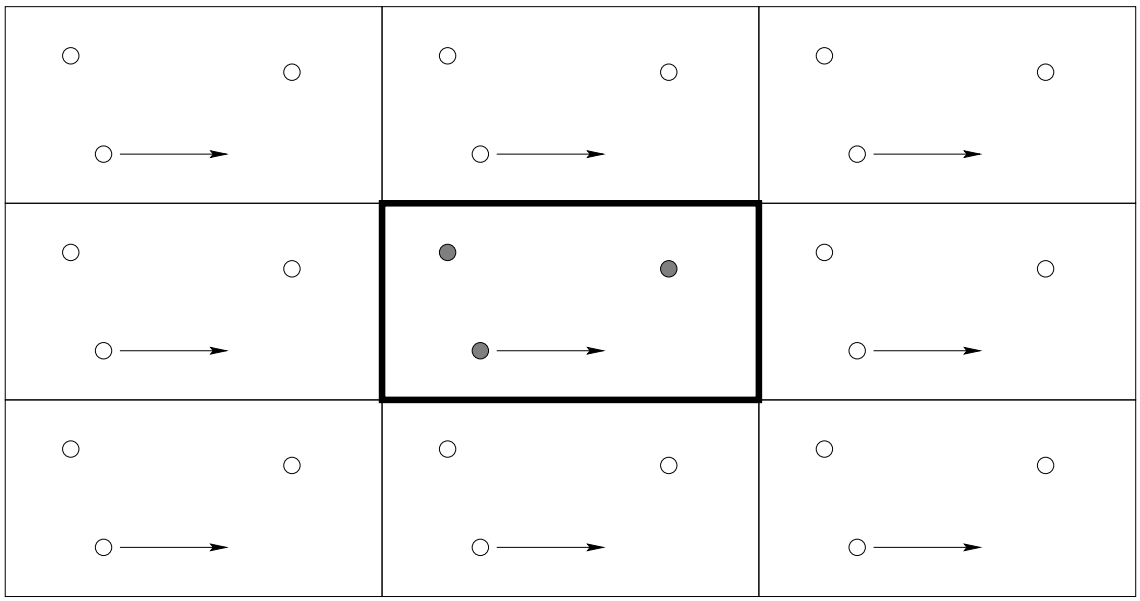


Figure 4.1: Illustrating the lattice generated by the periodic repeat of the simulation cell. Note that any movement of the particle in the simulation cell is copied by all the images.

The periodic repeat of the simulation cell leads to two distinct kinds of error, known as the *independent-particle* and *Coulomb* finite-size errors.

4.1.1 The independent-particle finite-size effect

The independent-particle finite-size effect is a result of replacing the smooth density of states of an infinite material with a set of discrete energy levels, as is inevitable when moving to a finite system. Figure 4.2 illustrates this point; it shows the kinetic energy per electron of a non-interacting electron gas as a function of the size of the box in which it is contained. The oscillations evident in figure 4.2 are a result of shell filling. Each effective one-electron wave function is associated with a wave vector; the wave vectors are grouped in shells of equivalent magnitude (and correspondingly equivalent one-particle kinetic energy).

This effect may be described more completely in terms of **k**-point sampling. In this context, two cases will be discussed: the periodic boundary conditions may be imposed on the wave function or (more generally) on the Hamiltonian operator. The more general case will be described first. In real solids, there is usually an under-

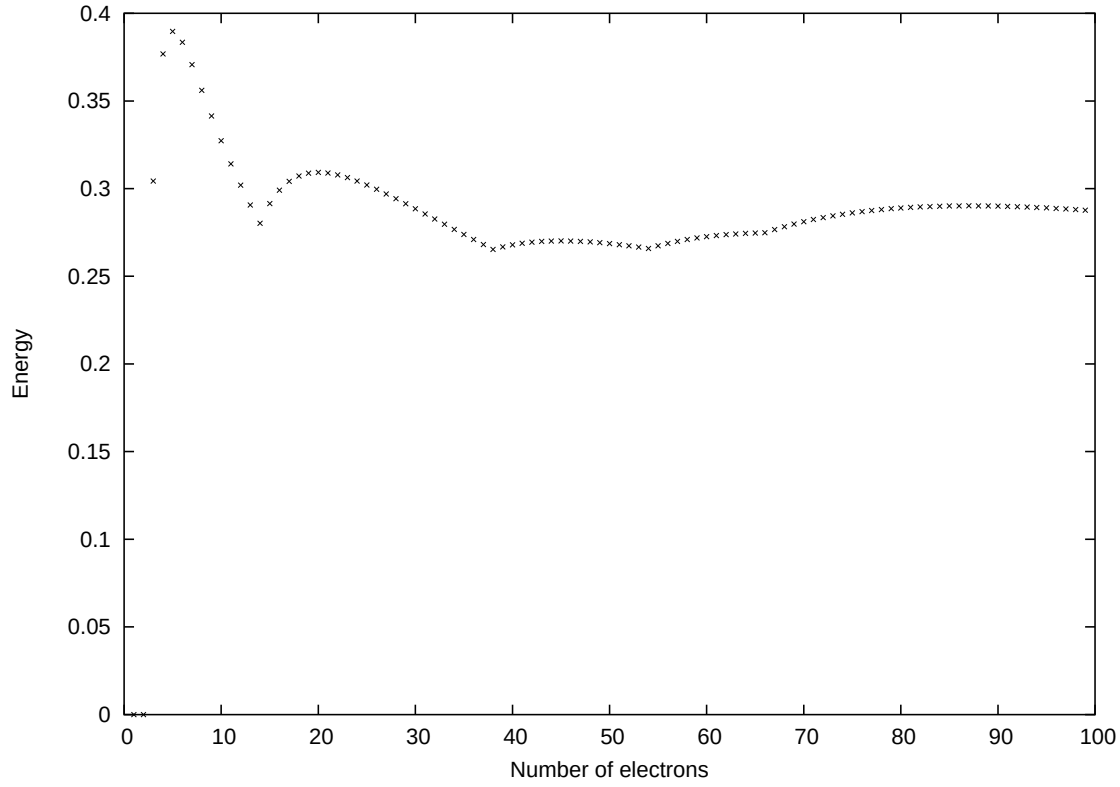


Figure 4.2: Shell-filling effects. Plotted here is the kinetic energy per electron of a non-interacting 3D electron gas obtained for different system sizes. (The density parameter is $r_s = 2$.)

lying physical periodicity (in addition to the artificial one created by the boundary conditions); however, in the discussion which follows, all references to periodicity and lattice vectors apply only to the artificial lattice defined by the simulation cell.

The Hamiltonian operator $\hat{H}$ is invariant under the translation of any electron by a lattice vector; such an operation therefore commutes with $\hat{H}$, as well as with other translations:

$$\left[\hat{H}, \hat{T}_{j\mathbf{R}}\right] = \left[\hat{T}_{j\mathbf{R}}, \hat{T}_{j'\mathbf{R}'}\right] = 0. \tag{4.1}$$

Here $\hat{T}_{j\mathbf{R}}$ denotes the operator corresponding to translation of electron j by the lattice vector $\mathbf{R}$.[2] Thus it is possible to choose a wave function to which is simulta-

[2]In this chapter, $\mathbf{R}$ will be used solely to refer to a simulation cell lattice vector; there should be no confusion with the notation of chapter 3, in which it represented the vector containing all the electron coordinates.

neously an eigenfunction of $\hat{H}$ and of the set of all such translations:

$$\hat{H}\Psi = E\Psi \tag{4.2}$$

$$\hat{T}_{j\mathbf{R}}\Psi = T_{\mathbf{R}}\Psi \quad \text{for all } \mathbf{R},\, j. \tag{4.3}$$

The eigenvalue $T_{\mathbf{R}}$ does not depend on j because of the antisymmetry of the wave function.

Two consecutive translations correspond to another single translation; it follows that

$$T_{\mathbf{R}}T_{\mathbf{R}'} = T_{(\mathbf{R}+\mathbf{R}')}. \tag{4.4}$$

The eigenvalues therefore have the exponential form

$$T_{\mathbf{R}} = e^{i\mathbf{k}\cdot\mathbf{R}}, \tag{4.5}$$

where at this stage the components of $\mathbf{k}$ are arbitrary complex numbers.[3] The unitarity of the translation operator means that the eigenvalues must have modulus 1, which in turn implies that $\mathbf{k}$ is in fact real. The result is just Bloch's theorem:

$$\hat{T}_{j\mathbf{R}}\Psi = e^{i\mathbf{k}\cdot\mathbf{R}}\Psi. \tag{4.6}$$

Thus, when the Hamiltonian operator is periodic in real space, the wave function must have the form

$$\Psi(\{\mathbf{r}_i\}) = \exp\left(i\mathbf{k}\cdot\sum_{i=1}^{N}\mathbf{r}_i\right)U(\{\mathbf{r}_i\}) \tag{4.7}$$

where $\hat{T}_{j\mathbf{R}}U = U$. The wave vector $\mathbf{k}$ may be reduced into the first Brillouin zone of the simulation cell: any remaining factors of

$$\exp\left(i\mathbf{K}\cdot\sum_{i=1}^{N}\mathbf{r}_i\right), \tag{4.8}$$

where $\mathbf{K}$ is a simulation cell reciprocal lattice vector, are periodic and may be included in U.

[3]More detail can be found in the book by Ashcroft and Mermin [4]; this section follows their proof of Bloch's theorem.

There is an infinite number of permissible **k**-vectors in the first Brillouin zone; imposing strictly periodic boundary conditions on the wave function, rather than on $\hat{H}$, corresponds to taking $\mathbf{k} = \mathbf{0}$.

However, there is no *a priori* reason why this particular **k**-point should be better than any other. In a non-interacting system, the single-electron wave functions are plane waves, and the infinite-system limit is obtained by integrating over all **k**-points. In an interacting system, this may also be the ideal approach [55]. The aim is always to choose the **k**-point(s) which best reproduce(s) the infinite-system result; several studies have investigated ways to achieve this goal in the context of QMC simulations [74, 39]. The use of alternative **k**-points is also discussed in chapter 9.

Another way to correct this kind of finite-size error in QMC results is to apply a correction of the form $(E_\infty^{\mathrm{DFT}} - E_N^{\mathrm{DFT}})$, where E_N^{DFT} and E_∞^{DFT} are the DFT results for finite and infinite cells respectively. This is a valid procedure because finite-cell DFT calculations suffer from the same **k**-point sampling errors, and it is usually easier to extrapolate the results of DFT calculations to the infinite-cell limit, thus obtaining E_∞^{DFT}. However, DFT calculations do not suffer from the second type of finite-size error: that related to the Coulomb interaction.

4.1.2 Coulomb finite-size errors

A requirement of any QMC simulation is the ability to determine the Coulomb energy of a given configuration. In a finite system, this is trivial; in a finite system which is designed to model an infinite system, it is not.

Referring to figure 4.1, each 'real' electron in the simulation cell must interact with all the others; there must also be some interaction between these electrons and the 'imaginary' electrons outside the simulation cell.

The conventional approach to this problem is to solve Poisson's equation for the charges in the cell with periodic boundary conditions.[4] The result is known as the

[4]Note that the simulation cell as a whole must have zero net charge; all the electron charges must be cancelled by an equivalent amount of positive charge (provided by ions in real materials,

Ewald sum; in a system with full three-dimensional periodicity, the potential at a position $\mathbf{r}$ due to a unit charge at the origin is given by

$$v_{\mathrm{E}}(\mathbf{r}) = \frac{1}{\Omega}\sum_{\mathbf{K}\neq 0}\frac{4\pi}{K^2}\exp\left(-\frac{\sigma^2K^2}{4}+i\mathbf{K}\cdot\mathbf{r}\right)+\sum_{\mathbf{R}}\frac{\mathrm{erfc}(|\mathbf{r}+\mathbf{R}|/\sigma)}{|\mathbf{r}+\mathbf{R}|}-\frac{\pi\sigma^2}{\Omega} \qquad (4.9)$$

where $\mathbf{R}$ and $\mathbf{K}$ are vectors of the real and reciprocal space lattices. The Ewald sum is evaluated efficiently because the original charge distribution has been divided into a smoothly-varying contribution, evaluated in reciprocal space, and a sum of localised quickly-decaying terms which can be evaluated in real space [78]. The parameter σ determines this separation, and can be chosen so that both real and reciprocal space sums converge quickly.

Using the Ewald interaction in place of the Coulomb interaction gives a well-defined result, the unique (up to an arbitrary constant) periodic solution of Poisson's equation. It also tends to the correct $1/r$ limit as the cell size tends to infinity. However, for finite cell sizes, it has been shown [25] that this interaction introduces an additional contribution to the electric field, and therefore to the potential. This unwanted contribution is known as the Coulomb finite-size error.

With the new interaction, the total electrostatic potential[5] at $\mathbf{r}$ is

$$\phi(\mathbf{r}) = \sum_i q_i v_{\mathrm{E}}(\mathbf{r}-\mathbf{r}_i), \qquad (4.10)$$

where the set $\{q_i\}$ represents all the charges in the system. This expression must be modified slightly to give the potential experienced by the charge q_i:

$$\bar{\phi}(\mathbf{r}_i) = \sum_{j\neq i} q_j v_{\mathrm{E}}(\mathbf{r}_i-\mathbf{r}_j) + q_i\xi. \qquad (4.11)$$

The constant ξ is the self-interaction potential. It appears because each charge must interact with its own periodically-repeated images:

$$\xi = \lim_{\mathbf{r}\to\mathbf{0}}\left(v_{\mathrm{E}}(\mathbf{r})-\frac{1}{r}\right). \qquad (4.12)$$

or a positive background charge density in the case of the electron gas). If this were not the case, the potential would be infinite.

[5]The notation used here follows that of Fraser and coworkers [25].

The operator which gives the total electrostatic energy of the system is therefore

$$\begin{aligned}\hat{E}_{\mathrm{es}}^{\mathrm{EW}} &= \frac{1}{2}\sum_i q_i \bar{\phi}(\hat{\mathbf{r}}_i) \\ &= \frac{1}{2}\sum_i q_i \left(\sum_{j\neq i} q_j v_{\mathrm{E}}(\hat{\mathbf{r}}_i - \hat{\mathbf{r}}_j) + q_i \xi \right) \\ &= \frac{1}{2}\sum_{i\neq j} q_i q_j \Big(v_{\mathrm{E}}(\hat{\mathbf{r}}_i - \hat{\mathbf{r}}_j) - \xi \Big)\end{aligned} \tag{4.13}$$

where the last step follows because the cell is charge-neutral:

$$\sum_i q_i = 0. \tag{4.14}$$

Equation (4.13) shows that any constant may be added to v_{E} without affecting the total energy, because the same constant is also (by definition) contained in ξ.

In simulations, the Coulomb energy is conventionally divided into contributions from electron-electron, electron-ion, and ion-ion interactions:

$$\hat{E}_{\mathrm{es}}^{\mathrm{EW}} = \hat{E}_{\mathrm{e-e}}^{\mathrm{EW}} + \hat{E}_{\mathrm{e-i}}^{\mathrm{EW}} + \hat{E}_{\mathrm{i-i}}^{\mathrm{EW}} \tag{4.15}$$

where

$$\hat{E}_{\mathrm{e-e}}^{\mathrm{EW}} = \frac{1}{2}\sum_{i=1}^{N}\sum_{\substack{j=1 \\ j\neq i}}^{N} v_{\mathrm{E}}(\hat{\mathbf{r}}_i - \hat{\mathbf{r}}_j) + \frac{1}{2}N\xi \tag{4.16}$$

$$\hat{E}_{\mathrm{e-i}}^{\mathrm{EW}} = -\frac{1}{2}\sum_{\alpha=1}^{M}\sum_{i=1}^{N} Z_\alpha v_{\mathrm{E}}(\hat{\mathbf{d}}_\alpha - \hat{\mathbf{r}}_i) \tag{4.17}$$

$$\hat{E}_{\mathrm{i-i}}^{\mathrm{EW}} = \frac{1}{2}\sum_{\alpha=1}^{M}\sum_{\substack{\alpha=1 \\ \alpha\neq\beta}}^{M} Z_\alpha Z_\beta v_{\mathrm{E}}(\hat{\mathbf{d}}_\alpha - \hat{\mathbf{d}}_\beta) + \frac{1}{2}\sum_{\alpha=1}^{M} Z_\alpha^2 \xi. \tag{4.18}$$

Here, there are N electrons and M ions; the sets $\{\mathbf{d}_\alpha\}$ and $\{Z_\alpha\}$ denote the positions and charges of the ions. Although the total energy is independent of the constant term in the Ewald interaction (and therefore of ξ), the individual contributions are not.

The expectation value of the electron-electron part of the Coulomb energy is

$$E_{\mathrm{e-e}}^{\mathrm{EW}} = \int_{\mathrm{cell}} |\Psi(\mathbf{X})|^2 \sum_{i>j} v_{\mathrm{E}}(\mathbf{r}_i - \mathbf{r}_j)\, d\mathbf{X} + \frac{1}{2}N\xi. \tag{4.19}$$

To understand the origin of the finite-size error, it is helpful to separate $E_{\mathrm{e-e}}^{\mathrm{EW}}$ into *Hartree* and *exchange-correlation* terms. For this analysis, the two-electron density is required:

$$n(\mathbf{r},\mathbf{r}') = \int_{\mathrm{cell}} |\Psi(\mathbf{X})|^2 \sum_{i\neq j} \delta(\mathbf{r}-\mathbf{r}_i)\delta(\mathbf{r}'-\mathbf{r}_j)\, d\mathbf{X}. \tag{4.20}$$

With this definition, the electron-electron energy becomes

$$E_{\mathrm{e-e}}^{\mathrm{EW}} = \frac{1}{2}\iint_{\mathrm{cell}} n(\mathbf{r},\mathbf{r}')v_{\mathrm{E}}(\mathbf{r}-\mathbf{r}')\, d\mathbf{r}\, d\mathbf{r}' + \frac{1}{2}N\xi. \tag{4.21}$$

If the electrons were completely uncorrelated, then the two-electron density $n(\mathbf{r},\mathbf{r}')$ would simply be the product of the one-electron densities $n(\mathbf{r})$ and $n(\mathbf{r}')$. However, the electrons are not uncorrelated; the relationship between one- and two-electron densities defines the *exchange-correlation hole*:

$$n(\mathbf{r},\mathbf{r}') = n(\mathbf{r})n(\mathbf{r}') + n(\mathbf{r})n_{\mathrm{XC}}(\mathbf{r},\mathbf{r}'). \tag{4.22}$$

The exchange-correlation hole describes the way that electrons avoid each other; the reduction in the electron density at $\mathbf{r}$ when one electron is fixed at $\mathbf{r}'$ is described by $n_{\mathrm{XC}}(\mathbf{r},\mathbf{r}')$. Integrating equation (4.22) with respect to $\mathbf{r}'$ reveals the important property

$$\int_{\mathrm{cell}} n_{\mathrm{XC}}(\mathbf{r},\mathbf{r}')\, d\mathbf{r}' = -1. \tag{4.23}$$

This is a manifestation of the fact that the hole consists of the absence of a single electron from the overall density.

Applying this to equation (4.21) gives

$$\begin{aligned} E_{\mathrm{e-e}}^{\mathrm{EW}} &= \frac{1}{2}\iint_{\mathrm{cell}} n(\mathbf{r})n(\mathbf{r}')v_{\mathrm{E}}(\mathbf{r}-\mathbf{r}')\, d\mathbf{r}\, d\mathbf{r}' + \frac{1}{2}\iint_{\mathrm{cell}} n(\mathbf{r})n_{\mathrm{XC}}(\mathbf{r},\mathbf{r}')[v_{\mathrm{E}}(\mathbf{r}-\mathbf{r}')-\xi]\, d\mathbf{r}\, d\mathbf{r}' \\ &= U_{\mathrm{Ha}} + U_{\mathrm{XC}}^{\mathrm{EW}}. \end{aligned} \tag{4.24}$$

The first term, the Hartree energy, is the classical self-interaction energy per simulation cell of a static periodic charge density $n(\mathbf{r})$ (compare equation (2.11)). The second term is the interaction energy of the electron with the exchange-correlation hole; this dynamical correction appears because the electron motions are correlated.

The Ewald interaction v_{E} is the correct interaction to use for the Hartree energy,[6] but not for the exchange-correlation energy. The reason for this is that the entire exchange-correlation hole (a total charge deficit of one electron) is contained within the simulation cell; there should be no images of the exchange-correlation hole. The Ewald interaction, which includes the effects of image charges, is therefore inappropriate for the calculation of U_{XC}.

4.2 Fixed-node errors

After many time steps, the walkers in a conventional DMC simulation are distributed according to the fixed-node density $\Psi_T \Psi_0^{\mathrm{FN}}$ rather than the desired density $\Psi_T \Psi_0$.

Any fixed-node estimate of the ground-state energy must be variational: $E_0^{\mathrm{FN}} \geq E_0$. The equality holds only when the nodes are exact; while this may be achievable for a one-electron system, it is almost impossible in many-electron calculations. An indication of the difficulty in correctly guessing the nodal surface is the high dimensionality of that surface: $(Nd - 1)$, for a system of N electrons moving in d dimensions. It is not possible to deduce the nodal surface from the condition that the wave function be zero when two electrons coincide: this defines a surface of only $(N - 1)d$ dimensions.

The fixed-node approximation is uncontrolled; the size of the error it introduces cannot be calculated analytically. It is not surprising that a great deal of time and effort has been devoted to overcoming this problem, with limited success.

The release-node algorithm of Ceperley and Alder [12] uses separate populations of positive and negative walkers which are allowed to cross the nodes of the trial wave function. The problem with this method is that both walker populations grow geometrically in time, leading to exponentially-increasing statistical fluctuations; the method becomes a race to obtain convergence to the ground state before the

[6] Although the Hartree energy defined by equation (4.24) depends on the value of ξ, the total energy does not, because of the corresponding terms in the ion-ion energy. The exchange-correlation energy defined by the same equation does not depend on ξ.

fluctuations become too large. The problem worsens for large systems, because the fluctuations increase with the system size.

An alternative approach is to include backflow correlations in the trial wave function [45, 46]. The orbitals which make up the Slater determinant contain parameters which are to be optimised; the nodes are therefore no longer fixed. Of course, optimising these parameters also incurs an extra computational overhead.

4.3 Statistical noise

The entire philosophy of Monte Carlo methods is based on random sampling. As such, the result of any QMC calculation is subject to statistical fluctuations, unless the trial wave function is an exact eigenfunction of the operator being measured. One of the advantages of QMC cited in chapter 3 is the scaling of the expected statistical error, which is proportional to $1/\sqrt{N}$, where N is the number of sampled configurations. While this is certainly better than grid-based methods for evaluating high-dimensional integrals, it does not represent rapid convergence.

Assuming that the wave function has not been guessed correctly, estimating the statistical error in the result is important. If the recorded values are $\{O_i : i = 1, \ldots, N\}$, with mean $\bar{O}$, then the sample variance is

$$s_{N-1}^2 = \frac{1}{N-1} \sum_{i=1}^{N} (O_i - \bar{O})^2. \tag{4.25}$$

If the samples are independent, this is a good estimator for the true variance σ^2; an estimate for the standard deviation of the mean value (which is an indicator of the expected size of the error) is then $s_{N-1}/\sqrt{N}$.

However, successive sampled points in a QMC calculation are usually correlated. The calculation of the expected error is then more complicated. One way around the problem is to group the data into blocks:

$$\bar{O}_j = \frac{1}{M} \sum_{i=(j-1)M+1}^{jM} O_i. \tag{4.26}$$

Here M is the block length. If M is large enough, then successive values of $\bar{O}_j$ will be uncorrelated; this occurs if M is greater than the correlation length of the data. The new error estimate is given by equation (4.25) with O_i replaced by $\bar{O}_i$ and N replaced by the number of blocks.

If M is not sufficiently large then the blocking method gives an error estimate which is too small. Thus a way of determining the correct block length is to increase M until the calculated error reaches a plateau; this is then the best error estimate. This technique may not be possible if the correlation length is very long or there are insufficient sample points.

In both VMC and DMC, the correlation length depends on the time step. If the time step is too small, then it takes many steps before a configuration changes significantly; however, if it is too large, then too many steps are rejected, and the time required to generate the next uncorrelated configuration is also large.

4.4 Surface calculations

There are several additional challenges which must must be met when dealing with surfaces, which will be detailed in the rest of this chapter.

4.4.1 System geometry

In order to study surfaces, it is usual to simulate slab systems (with two surfaces); using only a single surface creates problems with the boundary conditions.

The aim is to model a slab with infinite extent in two dimensions, but the simulation cell must be finite; it is therefore normal to apply periodic boundary conditions to the cell. However, rather than applying these conditions in two dimensions, most surface calculations use fully three-dimensional periodicity. In addition, it is conventional to use the 3D version of the Ewald interaction; thus the system actually being simulated is a stack of slabs, as shown in figure 4.3.

In density-functional and other mean-field calculations, this is not important.

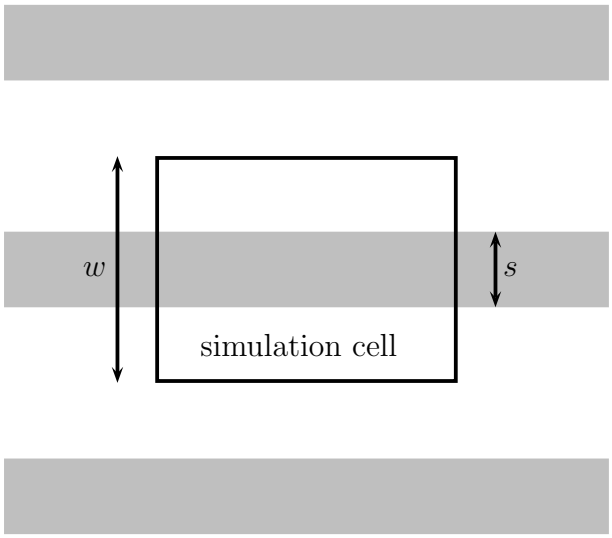


Figure 4.3: The effective system when applying periodic boundary conditions to the simulation cell. The shaded areas represent the positive background; the electron density is similar, though with less sharply-defined edges.

The electronic wave functions decay exponentially outside the slab, and any interaction between slabs is negligible.

However, in QMC calculations, this is no longer true. Rather than a smooth charge density, there is now a collection of individual electrons. Using the 3D Ewald interaction (which is equivalent to applying periodic boundary conditions in all three dimensions) introduces Coulomb finite-size errors; these were described in section 4.1.2. The form of the errors is not quite the same as in bulk systems: it is closer to that created when attempting to model a defect in an infinite system by using periodic boundary conditions (and thus repeating the defect).

It is not strictly necessary to use the 3D Ewald interaction, or to apply periodic boundary conditions in the direction perpendicular to the slab; the quasi-2D version of the Ewald interaction will be discussed later.

4.4.2 The surface energy

The energy per electron of a finite slab system is

$$\epsilon_{\text{slab}} = \epsilon_{\text{bulk}} + \frac{2A\sigma}{N} \tag{4.27}$$

where A is the area, σ is the surface energy, N is the number of electrons and ϵ_{bulk} is the energy per electron in the bulk material. The factor of 2 appears because there are two surfaces in a slab. In fact, this formula strictly only holds in the limit of an infinitely wide slab; for slabs of finite width, the energy per electron may also display oscillations [72]. In order to calculate the surface energy, one must therefore consider a slab which is sufficiently wide to render these oscillations unimportant; unfortunately, the wider the slab, the closer ϵ_{slab} becomes to ϵ_{bulk}, and the harder is the calculation of the surface energy:

$$\sigma = \frac{N}{2A}(\epsilon_{\text{slab}} - \epsilon_{\text{bulk}}). \tag{4.28}$$

This is one of the major problems in surface energy calculations: extremely high accuracy is often required, because the difference of two very similar numbers must be taken.

The situation is helped if the systematic errors in ϵ_{bulk} and ϵ_{slab} are the same, and cancel; however, the fundamental differences between slab and bulk calculations mean that this cannot generally be relied on.

The difficulties of surface energy calculations are exemplified by the simplest of surface systems — a slab of electron gas — which is described in the next chapter.

Chapter 5

The jellium slab

Historically, the electron gas (jellium) has been the proving-ground for electronic structure methods [12, 27];[1] without the additional complication of real ions, it provides the first basic test. The jellium slab, or quasi-2D electron gas, is the simplest surface system available, and is used as the subject of or test system for most of the work contained in this thesis.

5.1 Defining the system

A homogeneous electron gas is specified by the single parameter r_s. This is the radius of a sphere of size equal to the average volume of space per electron, which implies that the electron density is given by the relation

$$n = \frac{3}{4\pi r_s^3}. \tag{5.1}$$

The electron gas may be used as a crude model for a metal. Typical metallic densities correspond to the range $1 < r_s < 4$; in this work, the density parameter appropriate to aluminium ($r_s = 2.07$) has been used.

[1] Accurate studies of the electron gas are important in their own right, because (among other reasons) they provide the information on which exchange-correlation functionals (and hence density-functional theory calculations) are based.

The jellium slab is an electron gas with finite extent in one of the three spatial dimensions. There are different ways to constrain the electrons to a slab: two will be described here.

5.1.1 Constraining the electrons

The first and more usual way to define the slab is to set up the following background charge density:

$$\rho_b(z) = \begin{cases} 3/(4\pi r_s^3) & 0 < z < s \\ 0 & \text{otherwise.} \end{cases} \tag{5.2}$$

This fixes the slab width s. In a QMC simulation, a finite number of electrons must be used. The number of electrons then determines the size of the simulation cell in the xy-direction; the cell extends from $-\infty$ to ∞ in the z-direction. The integral of the electron density over z is equal to that of the positive background density; this is another way of saying that the system is charge-neutral.

The electrons are constrained by the attractive potential of the positive background, and pushed apart by their own mutual repulsion and kinetic energy. A typical electron density profile for this form of the jellium slab is shown in figure 5.1. Some electrons spill out of the slab into the vacuum region; standing wave oscillations caused by reflection of electron waves from the confining potential are evident, decaying from the edge of the slab towards the centre [47]. The oscillations are more pronounced when a small number of electrons is used.

It is also possible to constrain the electrons further, by imposing infinite barriers at $z = 0$ and at $z = s$. Electrons are no longer allowed to spill out into the vacuum region; the resultant density is shown in figure 5.2. This is the *infinite barrier model*; the term *jellium slab* is usually reserved for the unbounded system, and this convention will be applied here.

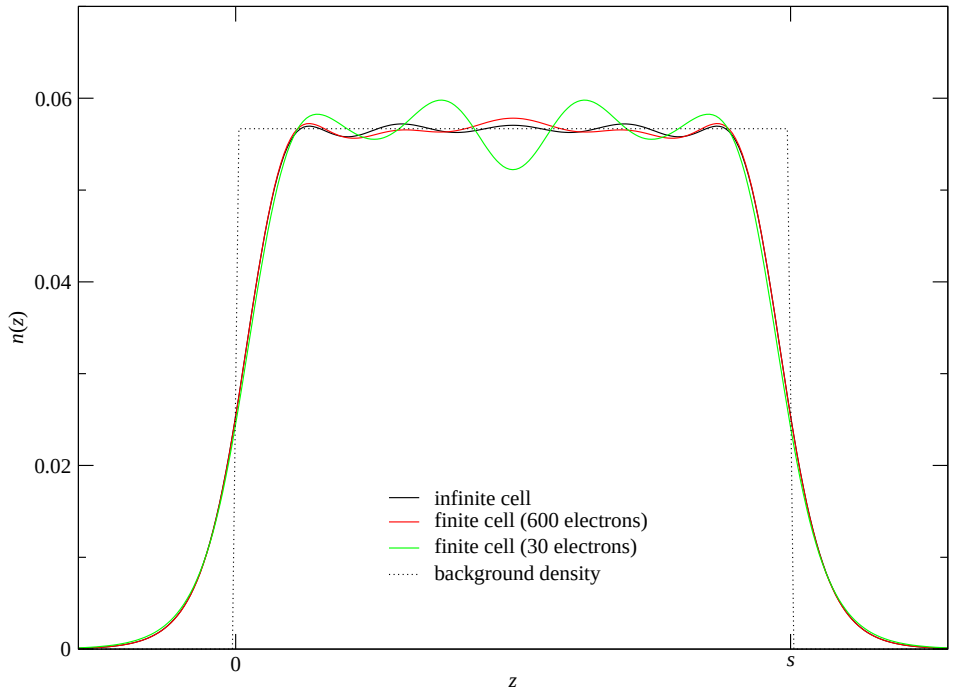


Figure 5.1: Electron density profiles for a conventional jellium slab with $r_s = 2.07$ and $s = 17.64248$. The profiles are obtained from LDA calculations and demonstrate the effect on the electron density of using a cell of finite extent in the xy-plane. Because r_s and s are fixed, the cell size is determined by the number of electrons.

5.2 The surface energy of jellium

One way in which the jellium slab differs from real materials is that the positive charge is fixed arbitrarily, and does not need to be in mechanical equilibrium; a consequence of this is that the surface energy of jellium can be negative for certain densities. This means that the accuracy problem discussed in section 4.4.2 is magnified.

The first DFT study of the jellium surface was carried out by Lang and Kohn [47] using the LDA; since then, several other investigations have been performed, using various exchange-correlation functionals of greater sophistication [49, 68]. A summary of the exchange-correlation contribution to the surface energy for the different functionals is contained in the paper by Yan and co-workers [85]. Surface

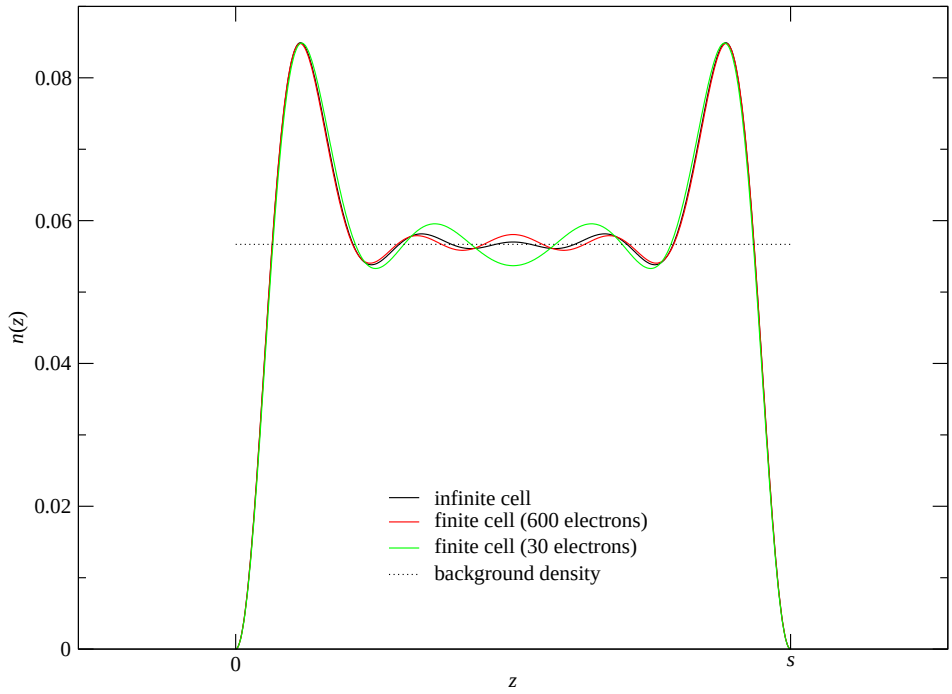


Figure 5.2: Electron density profiles for a bounded jellium slab, obtained in LDA.

energies are usually quoted in erg cm^{-2} or eV Å^{-2}; the conversion factors are

$$1 \text{ mHa bohr}^{-2} = 1556.8928 \text{ erg cm}^{-2} = 0.097173615 \text{ eV Å}^{-2}. \qquad (5.3)$$

Yan's summary also includes results obtained from the Fermi-hypernetted-chain method[2] and from fixed-node DMC.

The fixed-node DMC simulations were performed initially by Li et al. [53]; their results were later extended by Acioli and Ceperley [1]. Both groups attempted to model an infinite slab by applying periodic boundary conditions (as described in chapter 4). Although this is the usual way to calculate the jellium surface energy in QMC, it is not the only one. Sottile and Ballone [77] simulated jellium spheres using fixed-node DMC; for these finite systems, there are no finite-size errors. DFT simulations of jellium spheres were later carried out by Almeida, Perdew and Fiolhais [3].

[2]The Fermi-hypernetted-chain technique is a variational wave-function based method [42, 43]; it will not be described in detail in this work.

Yan and his co-authors demonstrate that the various DFT methods all give broadly the same values for the surface energy. More interestingly, these results also agree with the DMC results obtained by Sottile and Ballone using finite jellium spheres, but not with the extended-slab calculations of Acioli, Ceperley, Li et al.: the extended-slab DMC surface energies appear too large. In a recent paper, Pitarke [71] points out that Acioli and Ceperley incorrectly compared fixed-node slab energies with release-node bulk energies, and therefore overestimated the surface energy.[3] He argues that using the fixed-node bulk energy brings the DMC results closer to those obtained using DFT; however, they remain in disagreement, and the same correction does not apply to the earlier work of Li et al. It seems increasingly likely that the extended-slab DMC calculations were inaccurate.

Focusing on one density ($r_s = 2.07$) which is very often studied, some different values calculated for the surface energy are:

- -420 ± 80 erg cm^{-2} (Acioli and Ceperley [1], fixed-node DMC), corrected to -554 ± 80 erg cm^{-2} (Pitarke [71]);
- -465 ± 50 erg cm^{-2} (Li et al. [53], fixed-node DMC);
- -610 erg cm^{-2} (Yan et al. [85], LDA);
- -533 erg cm^{-2} (Yan et al. [85], wave vector interpolation based on the GGA);
- -690 erg cm^{-2} (Perdew et al. [68], GGA);
- -567 erg cm^{-2} (Perdew et al. [68], meta-GGA);
- -553 erg cm^{-2} (Kurth and Perdew [44], combination of random phase approximation and LDA);
- -587 erg cm^{-2} (Kurth and Perdew [44], combination of random phase approximation and GGA).

[3] The surface energy is negative; the 'overestimate' referred to here is a result which is insufficiently negative, and therefore smaller in magnitude than the true value.

The DMC results clearly lie above those obtained in DFT. More evidence is provided by Almeida et al. [3], who studied the jellium sphere system, and were able to match their DFT surface energies with Sottile and Ballone's finite-system DMC results over a range of densities (although the precise value $r_s = 2.07$ was not included in their calculations). With the correction suggested by Pitarke, the calculation of Acioli and Ceperley is brought closer to the DFT results, but this correction cannot be applied to the calculations of Li et al.

In addition, both groups (Li et al. and Acioli and Ceperley) applied only the independent-particle finite-size correction[4] to the slab energies, and not the Coulomb correction. The Coulomb finite-size correction would increase the slab energies; because the corresponding bulk energies were fully finite-size corrected, making the correction would raise the DMC surface energy still further, increasing the difference between the DMC and DFT results. Almeida believes the combined DFT-RPA calculations to be currently the most accurate, putting the surface energy between -550 and -590 erg cm^{-2}.

At this density, the separate contributions to the surface energy are individually sizeable, but cancel as a whole; Pitarke and Eguiluz [72] give the following breakdown:

- $\sigma_s = -4643$ erg cm^{-2} (kinetic);
- $\sigma_{es} = 1072$ erg cm^{-2} (electrostatic);
- $\sigma_{xc} = 3007$ erg cm^{-2} (exchange-correlation).

This is the opposite of the ideal situation, and is partly why jellium surface energy calculations are particularly difficult.

A relative error of 10% in the surface energy σ corresponds to around 50 erg cm^{-2} or 0.03 mHa bohr^{-2}. Using equation (4.28) to calculate σ, and assuming that there is no error in the bulk energy gives

$$\Delta\sigma = \frac{N}{2A}\Delta\epsilon_{\text{slab}} = \frac{3s}{8\pi r_s^3}\Delta\epsilon_{\text{slab}}. \tag{5.4}$$

[4]These finite-size errors are discussed in chapter 4.

Here $\Delta\epsilon_{\text{slab}}$ is the error in the calculated energy per electron of the slab and $\Delta\sigma$ is the resultant error in the jellium surface energy. Equation (5.1) has been used to relate the number of electrons per unit area N/A to the density parameter r_s. For a slab width of 20.0, setting $\Delta\sigma = 50$ erg cm^{-2} gives $\Delta\epsilon_{\text{slab}} \approx 0.1$ mHa (or 3 meV) per electron. This gives an idea of the accuracy needed in the Monte Carlo simulations.

5.3 Preliminary investigations

The surface energy of jellium is defined in the limit of an infinitely-wide slab with infinite extent in the xy-direction. In DFT, it is possible to achieve one of these limits. Because the external (background) potential does not depend on x or y, the system is homogeneous in the xy-plane; the non-interacting Kohn-Sham orbitals have the simple form

$$\phi_{n\mathbf{k}_\parallel}(\mathbf{r}) = u_n(z)e^{i\mathbf{k}_\parallel\cdot\mathbf{r}_\parallel}. \tag{5.5}$$

The density of states in $\mathbf{k}_\parallel$-space is therefore constant, and extrapolation to a system with infinite xy-extent is trivial. This is how the infinite-system density profiles displayed in figures 5.1 and 5.2 were obtained.

Unfortunately, this simple extrapolation is not possible in QMC simulations, which must use a finite number of electrons; for this reason, it is useful to study finite cells in DFT.

The DFT simulations are carried out on a grid in the z-direction which extends for some distance outside the slab, using a code supplied by Pablo Garcia-Gonzalez [26]. It is important to ensure that the results are converged with respect to both the number of grid points used and the spacing of these points (or equivalently, the effective cell size in the z-direction, which will be denoted as w).

Figure 5.3 shows the results of convergence testing, for both the finite and infinite horizontal cells. Cells of two different lengths in the z-direction have been compared: even the smaller of these is around five times larger than the slab width, and the figure demonstrates that in this regime the cell size is unimportant. What matters is the sampling of the slab region: this is why the results for a cell size of 105 with 512

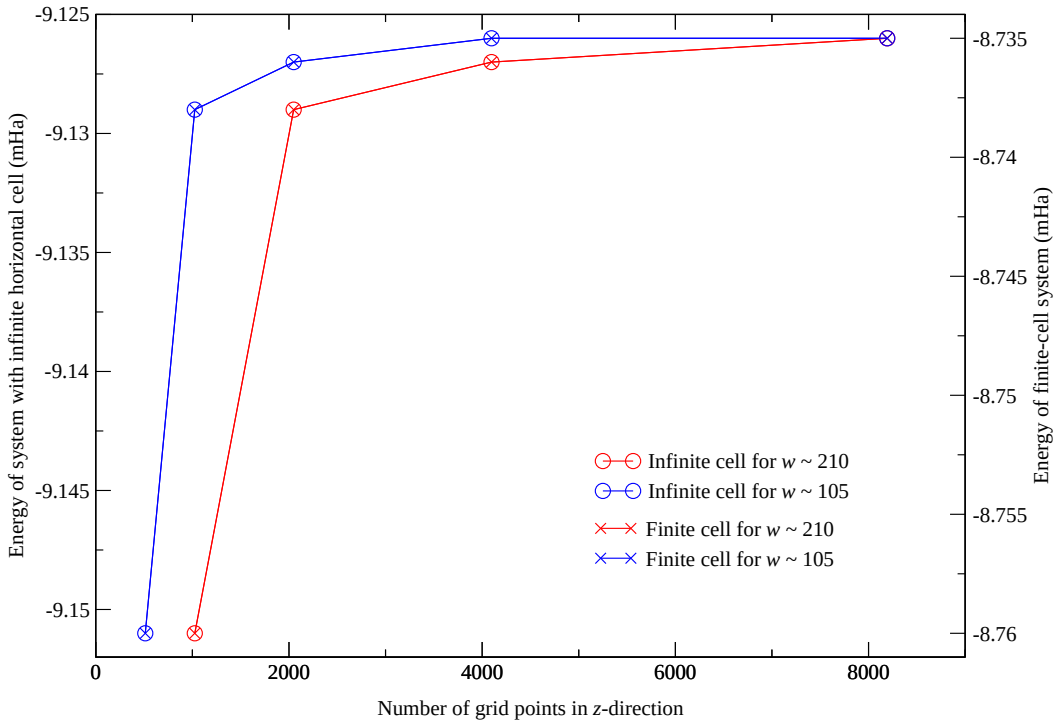


Figure 5.3: The LDA energy per electron as a function of the number of grid points. Cells of size $w \sim$105 and $w \sim$210 are compared, both for finite systems containing 360 electrons and for systems infinite in the xy-direction. Note that the two vertical scales are offset. The slab width is 17.64248; the density parameter is $r_s = 2.07$. The length of the cell in the z-direction varies slightly as a function of the number of grid points; this is to ensure optimal sampling of the important slab region.

grid points are the same as those for a cell size of 210 using 1024 points. In addition, the error caused by using an insufficient number of grid points is approximately the same for both the finite and infinite systems: the two vertical scales are the same, distinguished only by a constant offset. For a system of size $w \sim 105$, using 1024 grid points ensures that the finite-size error is less than 0.01 mHa.

The energy per electron in the slab system shows a more interesting (and more relevant, from the point of view of a surface energy calculation) dependence on the slab width. This is shown in figure 5.4. Rearranging equation (4.27) slightly gives the slab energy per electron for a particular density as a function of the slab width:

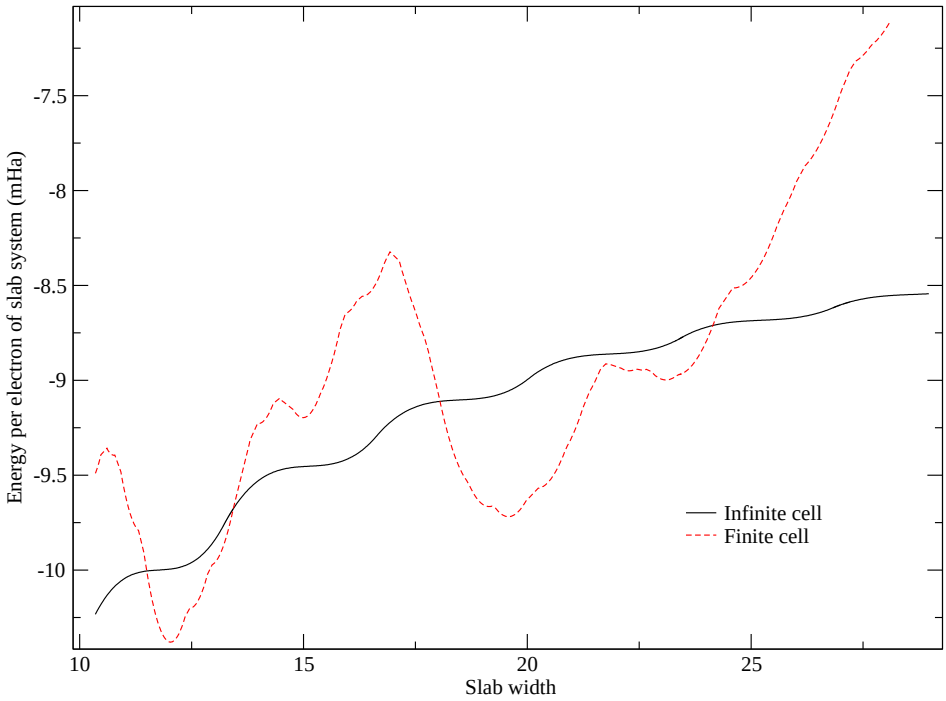


Figure 5.4: The dependence of the LDA energy per electron on the slab width; for this system, $r_s = 2.07$, the number of electrons is 360 and the cell size in the z-direction is around 105.

$$\epsilon_{\text{slab}} = \epsilon_{\text{bulk}} + \frac{8\pi r_s^3 \sigma}{3s}. \tag{5.6}$$

When $r_s = 2.07$, the surface energy is negative, which means that ϵ_{slab} should tend to the bulk value from below as the slab width s is increased; this behaviour is exhibited by the infinite-system curve in figure 5.4. The figure also shows that the errors introduced by using a finite simulation cell can be significant (~ 1 mHa).

The small oscillations in the infinite-system curve are caused by sub-bands: as the slab gets wider, more sub-bands begin to be filled. They decrease in magnitude as the slab becomes wider; the true surface energy is defined in the limit of infinite slab width, when they disappear altogether.

To calculate the surface energy within DFT using equation (4.28), the energy per electron in the bulk system (the homogeneous electron gas) is required. The correct approach is to use the same exchange-correlation energy functional in the

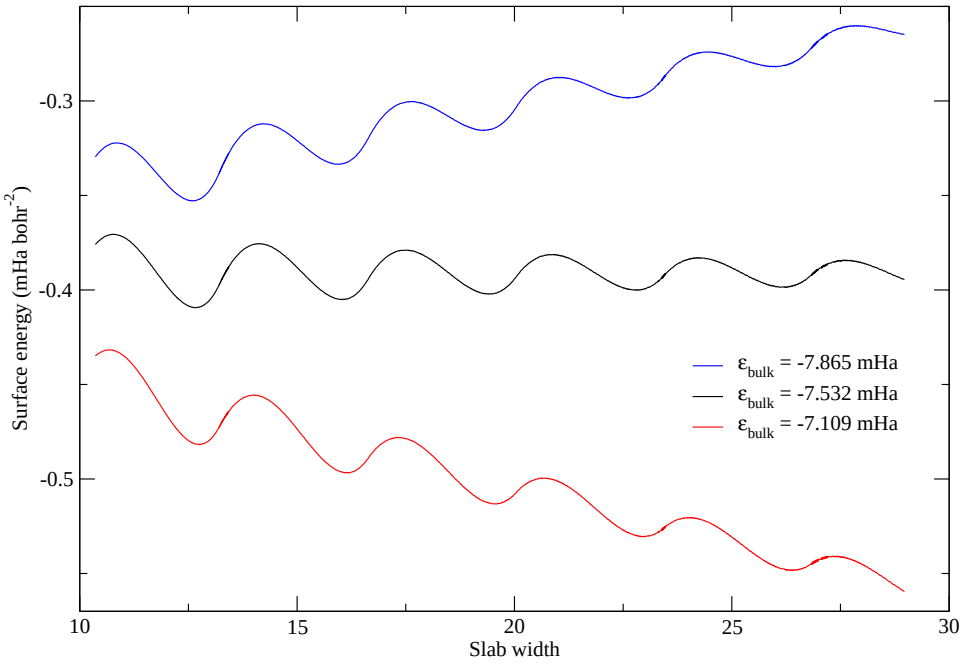


Figure 5.5: The surface energy as a function of slab width, calculated using different values of the bulk energy. The middle curve uses the Perdew-Wang parameterisation [69] of Ceperley and Alder's early QMC results [12], which was the functional employed in the slab calculation. The upper curve is from the Perdew-Zunger [70] paremeterisation of the same results; the lower curve is from the Perdew-Zunger parameterisation of Ortiz and Ballone's later QMC calculations [63].

slab and bulk calculations; figure 5.5 shows the effect of using the wrong bulk energy. This figure, though simple, emphasises the important requirement that the bulk and slab calculations must be carried out consistently. The error in the surface energy induced by using an incorrect value of the bulk energy increases with the slab width. If it is not possible to carry out consistent slab and bulk calculations, it is better to use slab results only and to deduce the surface energy by fitting ϵ_{slab} against $1/s$.

Pitarke and Eguiluz [72] analysed the oscillations in the infinite-system curve using the infinite barrier model, and showed that they should have a wavelength of $\lambda_{\text{F}}/2$, where the Fermi wavelength is

$$\lambda_{\text{F}} = 2\pi r_s \left(\frac{4}{9\pi}\right)^{1/3}. \tag{5.7}$$

When $r_s = 2.07$, the Fermi wavelength is 6.78. The wavelength of the oscillations

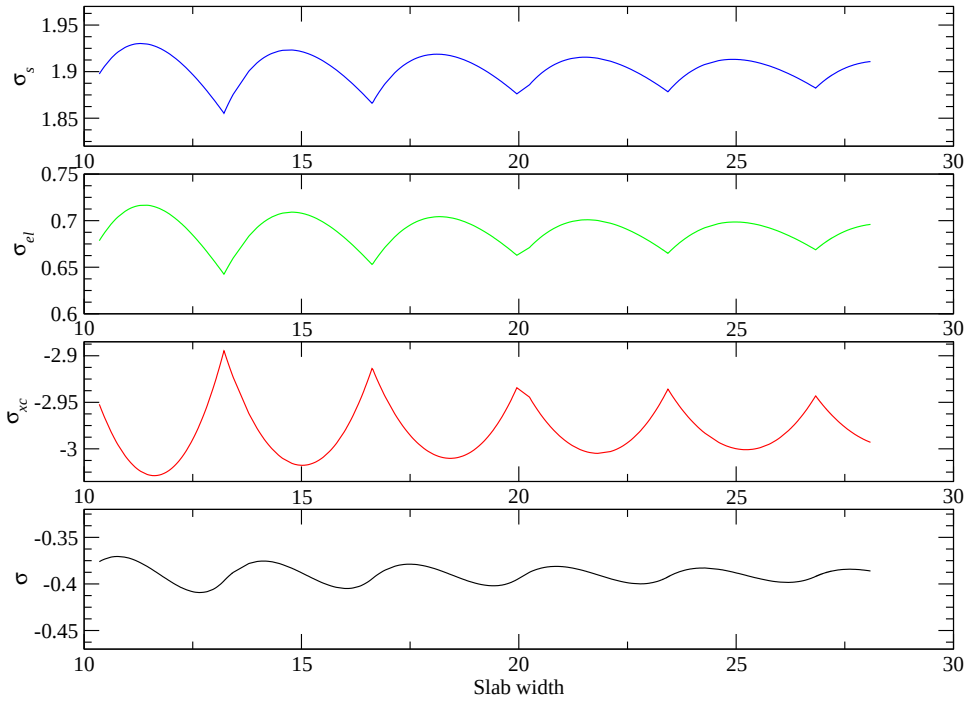


Figure 5.6: The components of the LDA surface energy as a function of the slab width. All the surface energies are given in mHa bohr^{-2}.

visible in figure 5.5 is around 3.88, which corresponds approximately to $\lambda_F/2$. This may be seen more clearly in figure 5.6, which shows the different components of the surface energy. The oscillations in the kinetic, electrostatic and exchange-correlation contributions to the surface energy largely cancel each other out. In particular, the cusps which are present in all the individual components do not appear in the total surface energy. The magnitude of the oscillations in the total surface energy is around 0.03 mHa bohr^{-2}, or 10%, for a slab width of 10.

5.4 The jellium slab in QMC

The Kohn-Sham orbitals obtained from the DFT calculations are the foundation for the QMC trial wave function (equation (3.77)); they are the components of the Slater determinant. The usual procedure is then to improve on the determinantal wave function by adding a Jastrow factor with a successively increasing number of

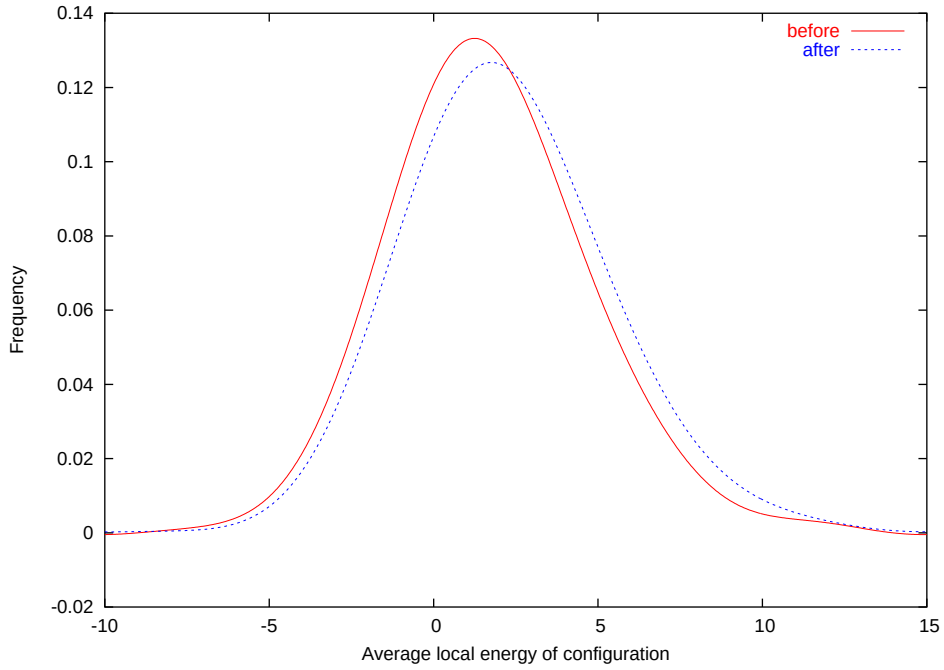


Figure 5.7: The distribution of configuration energies, before and after variance minimisation. 10000 configurations have been included; the Jastrow factor was of the form introduced in section 3.4.1 and described in detail in section 6.3. Optimisation of both u and χ terms was allowed; The minimisation was performed without reweighting.

terms. Each term includes parameters which are to be optimised using the variance minimisation technique described in section 3.4.2.

This procedure was found to be unsuccessful for the jellium slab system. The possible reasons for the failure of variance minimisation will be discussed in this section. Typically, the variance minimisation proceeds initially as expected: the variance of the local energy of the first set of configurations is reduced, along with the mean value. However, when a new set of configurations is generated with the modified parameters, it is subsequently found that the mean and variance of the local energy of the new configurations have both increased. This is illustrated in figure 5.7.

The new configurations are physically more spread out than the old: the electron density outside the slab increases, while that inside decreases. This has the effect of

reducing the average kinetic energy at the expense of increasing the potential energy by a larger amount.

The problem occurs both when reweighting of configurations is used and when it is not. This suggests that the cause is not outliers (configurations with energy far from the mean). Figure 5.7 shows that the energy distributions are close to being Gaussian, with only slightly 'fat' upper tails.

The configurations sampled initially do not include any with electrons far outside the slab; the electron density decays sharply to zero here, as can be seen in figure 5.1. After the first stage of variance minimisation, the Jastrow factor is altered so that the electron density extends much further outside the slab; therefore the next set of configurations which are generated will sample a different region of configuration space to the old set. The effect of pushing many electrons into the vacuum region, and the consequential increase in the potential energy, cannot be 'known' in advance by the variance minimisation routine: no such configurations have been sampled. This suggests that the process may be corrected by modifying the initial sampling distribution to include these configurations. Unfortunately, this approach did not work; the procedure remains unstable, with the mean and variance of the local energy usually worsening after each iteration.

Without reweighting, the object of the variance minimisation step is to minimise the following quantity:

$$O(\{\mathbf{X}_i;\alpha\}) = \sum_i \left[E_L(\mathbf{X}_i;\alpha) - \frac{1}{N}\sum_j E_L(\mathbf{X}_j;\alpha) \right]^2. \tag{5.8}$$

As before, α denotes the optimiseable parameter, $\{\mathbf{X}_i\}$ is the set of configurations and E_L is the local energy. This may be split into kinetic and potential terms; the potential energy does not depend on α:

$$O(\{\mathbf{X}_i;\alpha\}) = \sum_i \left[T_i(\alpha) - \bar{T}(\alpha) + V_i - \bar{V} \right]^2. \tag{5.9}$$

At the minimum,[5] this quantity must be stationary, giving

$$\frac{\partial O}{\partial \alpha} = 2 \sum_i \left[T_i(\alpha) - \bar{T}(\alpha) + V_i - \bar{V} \right] \left[\frac{\partial T_i}{\partial \alpha} - \frac{1}{N} \sum_j \frac{\partial T_j}{\partial \alpha} \right] = 0. \tag{5.10}$$

As well as the desirable minimum, when E_L is constant, there is at least one other stationary point, when $\partial T_i / \partial \alpha$ is constant; it is not trivial to determine whether or not this is a (local) minimum. It may be the case that there is a false minimum for the jellium slab in which the variance minimisation procedure becomes stuck.

The question of the correct form for the Jastrow factor is also a difficult one. If the form is wrong, giving too much or too little flexibility, then the variance minimisation procedure cannot be expected to work. Several different forms for both the one- and two-body terms were tested and found to be equally unsuccessful under variance minimisation.

The result is that in order to perform realistic VMC or efficient DMC simulations, manual optimisation is required. One parameter is adjusted at a time: VMC simulations for different values of the parameter are performed, and the results are fitted to a quadratic form.

The traditional u function described by equation (3.79), which has been successful in simulations of the homogeneous electron gas, was used in the calculations of chapter 6. Including this two-body term causes electrons to spill out of the slab; the one-body term must restore the correct electron density, as discussed in section 3.4.1.

Chapters 7 and 8 are devoted to improving the Jastrow factor; the new version is used in the calculations of chapter 9.

[5] For an exact eigenstate, the true minimum has $O = 0$; however, this point is not generally accessible.

Chapter 6

The modified periodic Coulomb interaction in quasi-2D systems

Finite-size errors related to the implementation of the Coulomb interaction in periodic systems were described in chapter 4. In systems with 3D periodicity, it has been shown [25] that using the Ewald interaction introduces an error which decays as the reciprocal of the number of electrons in the system. An effective cure for this finite-size error has been found: the model periodic Coulomb (MPC) interaction [84], which has the additional benefit of being significantly less computationally expensive than the Ewald sum. In this chapter, the theory of the MPC interaction will be outlined, and the extension of the interaction to quasi-2D systems will be investigated.

6.1 The MPC interaction

The analysis of the Coulomb finite-size error in section 4.1.2 is not specific to systems with 3D periodicity. There, the problem was shown to be caused by the use of the Ewald interaction in the exchange-correlation energy:

$$U_{\mathrm{XC}}^{\mathrm{EW}} = \frac{1}{2} \iint_{\mathrm{cell}} n(\mathbf{r}) n_{\mathrm{XC}}(\mathbf{r}, \mathbf{r}') [v_{\mathrm{E}}(\mathbf{r} - \mathbf{r}') - \xi] \, d\mathbf{r} \, d\mathbf{r}'. \tag{6.1}$$

In fact, since the exchange-correlation hole is contained entirely within the simulation cell and converges to the shape it would have in an infinite system rapidly as the system size increases, the correct interaction should have the usual Coulomb $1/|\mathbf{r}-\mathbf{r}'|$ form, with the proviso that the vector $(\mathbf{r}-\mathbf{r}')$ is first reduced into the Wigner-Seitz cell of the simulation lattice. This reduction of $(\mathbf{r}-\mathbf{r}')$ is known as the *minimum image* convention, and ensures that the interaction remains periodic (as it must). The correct interaction will be written $f(\mathbf{r}-\mathbf{r}')$, where

$$f(\mathbf{r}-\mathbf{r}') = \frac{1}{|(\mathbf{r}-\mathbf{r}')_{\mathrm{MI}}|}. \tag{6.2}$$

The solution to the problem is therefore simply to replace $[v_{\mathrm{E}}(\mathbf{r}-\mathbf{r}')-\xi]$ with $f(\mathbf{r}-\mathbf{r}')$ in equation (6.1):

$$U_{\mathrm{XC}}^{\mathrm{MPC}} = \frac{1}{2}\iint_{\mathrm{cell}} n(\mathbf{r})n_{\mathrm{XC}}(\mathbf{r},\mathbf{r}')f(\mathbf{r}-\mathbf{r}')\,d\mathbf{r}\,d\mathbf{r}'. \tag{6.3}$$

The total electron-electron interaction energy is then

$$\begin{aligned} E_{\mathrm{e-e}}^{\mathrm{MPC}} &= U_{\mathrm{Ha}} + U_{\mathrm{XC}}^{\mathrm{MPC}} \\ &= \frac{1}{2}\iint_{\mathrm{cell}} n(\mathbf{r},\mathbf{r}')f(\mathbf{r}-\mathbf{r}')\,d\mathbf{r}\,d\mathbf{r}' \\ &\quad + \frac{1}{2}\iint_{\mathrm{cell}} n(\mathbf{r})n(\mathbf{r}')[v_{\mathrm{E}}(\mathbf{r}-\mathbf{r}')-f(\mathbf{r}-\mathbf{r}')]\,d\mathbf{r}\,d\mathbf{r}'. \end{aligned} \tag{6.4}$$

The ground-state wave function Ψ associated with this expression for the electron-electron energy must minimise the total energy

$$E[\Psi] = \langle\Psi|\hat{T}+\hat{V}_{\mathrm{ext}}|\Psi\rangle + E_{\mathrm{e-e}}^{\mathrm{MPC}}[\Psi], \tag{6.5}$$

where $\hat{T}$ and $\hat{V}_{\mathrm{ext}}$ are operators for the kinetic energy and the external potential. At the same time, normalisation of Ψ must be preserved; this yields the variational principle

$$\delta\big(E[\Psi]-\lambda\langle\Psi|\Psi\rangle\big) = 0 \tag{6.6}$$

where λ is a Lagrange multiplier, and thus

$$\left(\hat{T} + \hat{V}_{\text{ext}} + \sum_i \int n(\mathbf{r}')[v_{\text{E}}(\mathbf{r}_i, \mathbf{r}') - f(\mathbf{r}_i, \mathbf{r}')]\, d\mathbf{r}' + \sum_{i>j} f(\mathbf{r}_i, \mathbf{r}_j) - \lambda\right)\Psi(\mathbf{r}_1 \ldots \mathbf{r}_N) = 0. \quad (6.7)$$

This is an eigenvalue equation for Ψ:

$$\big(\hat{T} + \hat{V}_{\text{ext}} + \hat{H}^{\text{MPC}}_{\text{e-e}}\big)\Psi = \lambda\Psi, \quad (6.8)$$

where the required term in the Hamiltonian is

$$\hat{H}^{\text{MPC}}_{\text{e-e}} = \sum_i \int n(\mathbf{r}')[v_{\text{E}}(\mathbf{r}_i, \mathbf{r}') - f(\mathbf{r}_i, \mathbf{r}')]\, d\mathbf{r}' + \sum_{i>j} f(\mathbf{r}_i, \mathbf{r}_j). \quad (6.9)$$

However, the eigenvalue λ does *not* correspond to the energy:

$$\begin{aligned} \lambda &= \langle\Psi|\hat{T} + \hat{V}_{\text{ext}} + \hat{H}^{\text{MPC}}_{\text{e-e}}|\Psi\rangle \quad (6.10)\\ &\neq E[\Psi]. \end{aligned}$$

The relationship between $\hat{H}^{\text{MPC}}_{\text{e-e}}$ and $E^{\text{MPC}}_{\text{e-e}}$ is the following:

$$E^{\text{MPC}}_{\text{e-e}} = \big\langle \hat{H}^{\text{MPC}}_{\text{e-e}} \big\rangle - \frac{1}{2}\iint_{\text{cell}} n(\mathbf{r})n(\mathbf{r}')[v_{\text{E}}(\mathbf{r} - \mathbf{r}') - f(\mathbf{r} - \mathbf{r}')]\, d\mathbf{r}\, d\mathbf{r}'. \quad (6.11)$$

In a DMC simulation, the modified Hamiltonian term $\hat{H}^{\text{MPC}}_{\text{e-e}}$ should be used to calculate the drift vector and the branching probability; this will ensure that the distribution converges to the correct wave function. However, when the goal is to estimate the ground-state energy, equation (6.11) should be used.

To evaluate $\hat{H}^{\text{MPC}}_{\text{e-e}}$ or $E^{\text{MPC}}_{\text{e-e}}$ during a simulation requires a knowledge of $n(\mathbf{r})$, the electron density. In general, this is not known exactly before the simulation begins. However, a good approximation may be obtained from the independent-particle calculation, which is already required for generating the orbitals in the trial wave function.

When using this approximation, it is possible for the resultant QMC density to differ from the approximate density used to calculate the electron-electron interaction energy during the simulation; the calculation is then not self-consistent.

However, as long as the density and wave function are close to their ground-state forms, this lack of self-consistency is not important; the energy calculated in QMC is not sensitive to small errors in the approximated density, as the following argument shows.

The estimated density will be denoted $\tilde{n}$; the QMC energy now depends on both $\tilde{n}$ and Ψ:

$$E[\Psi;\tilde{n}] = \left\langle \Psi \middle| \hat{T} + \hat{V}_{\text{ext}} + \hat{H}^{\text{MPC}}_{\text{e-e}}[\tilde{n}] \middle| \Psi \right\rangle - \frac{1}{2} \iint_{\text{cell}} \tilde{n}(\mathbf{r})\tilde{n}(\mathbf{r}')[v_{\text{E}}(\mathbf{r},\mathbf{r}') - f(\mathbf{r},\mathbf{r}')]\, d\mathbf{r}\, d\mathbf{r}'. \tag{6.12}$$

When the ground-state density is estimated correctly ($\tilde{n} = n_0$), the energy is minimised by the true ground-state wave function:

$$\left(\frac{\delta}{\delta \Psi} E[\Psi; n_0] \right)_{\Psi=\Phi_0} = 0. \tag{6.13}$$

This is just a restatement of the variational principle. A similar condition applies to the estimated density:

$$\frac{\delta}{\delta \tilde{n}(\mathbf{r})} E[\Psi;\tilde{n}] = \int_{\text{cell}} [n(\mathbf{r}') - \tilde{n}(\mathbf{r}')][v_{\text{E}}(\mathbf{r}-\mathbf{r}') - f(\mathbf{r}-\mathbf{r}')]\, d\mathbf{r}', \tag{6.14}$$

where n is the QMC density corresponding to Ψ. It follows that when the calculation is self-consistent

$$\left(\frac{\delta}{\delta \tilde{n}} E[\Psi;\tilde{n}] \right)_{\tilde{n}=n} = 0. \tag{6.15}$$

The implication of equations (6.13) and (6.15) is that when $\Psi = \Phi_0$ and $\tilde{n} = n_0$, the energy is stationary with respect to both Ψ and $\tilde{n}$; therefore, when both these functions are close to their ground-state forms, the error in the calculated QMC energy is second-order in $(\Psi - \Phi_0)$ and $(\tilde{n} - n_0)$.

Equation (6.9) illustrates the reason for the improvement in speed achieved by the MPC interaction. Two-body interactions require $\mathcal{O}[N^2]$ operations, while one-body interactions require only $\mathcal{O}[N]$; the only two-body term in the MPC interaction is f, which is a much simpler function to evaluate than the costly v_{E}. The remaining term in equation (6.9) is effectively a one-body potential.

6.2 Quasi-2D systems

In the past, slab calculations have often been performed using the 3D Ewald interaction. As described in section 4.4, this means that the system being simulated is actually a stack of slabs, and leads to finite-size errors in QMC simulations.

The quasi-2D generalisation of the Ewald potential was first published by Parry [65, 66], although the original derivation is unconvincing. An alternative derivation is presented in appendix A. As in the 3D case, the potential is the periodic solution of Poisson's equation, although the periodicity is now only in two dimensions. The resulting formula,

$$\begin{aligned} v_{\mathrm{E}}(\mathbf{r}) = \sum_{\mathbf{R}} \frac{1}{|\mathbf{r}-\mathbf{R}|} \operatorname{erfc}\left(\frac{|\mathbf{r}-\mathbf{R}|}{\sigma}\right) - \frac{2\pi}{A}\left[z \operatorname{erf}\left(\frac{z}{\sigma}\right) + \frac{\sigma}{\sqrt{\pi}} e^{-z^2/\sigma^2}\right] \\ + \sum_{\mathbf{k}} \frac{\pi}{kA}\left[e^{-kz}\operatorname{erfc}\left(\frac{\sigma k}{2} - \frac{z}{\sigma}\right) + e^{kz}\operatorname{erfc}\left(\frac{\sigma k}{2} + \frac{z}{\sigma}\right)\right] e^{i\mathbf{k}\cdot\mathbf{r}_\parallel}, \end{aligned} \tag{6.16}$$

is more awkward than the 3D version. As in the 3D sum, σ is the convergence parameter, which is chosen to ensure that both the real- and reciprocal-space sums converge quickly; $\mathbf{R}$ and $\mathbf{k}$ represent real- and reciprocal-space vectors of the 2D lattice defined by the simulation cell, which has area A.

The discussion presented in the previous section applies equally to 3D and quasi-2D systems; the MPC interaction should reduce finite-size errors in both.

There are some subtleties in the implementation of the alternative interaction in quasi-2D systems, however; these relate to the calculation of the first term on the right-hand side of equation (6.9), the one-body potential. This term is expressed using the 3D Fourier transforms

$$n_{\mathbf{k}} = \frac{1}{\Omega}\int_{\Omega} n(\mathbf{r}) e^{i\mathbf{k}\cdot\mathbf{r}}\, d\mathbf{r} \tag{6.17}$$

$$g_{\mathbf{k}} = \frac{1}{\Omega}\int_{\Omega} [v_{\mathrm{E}}(\mathbf{r}) - f(\mathbf{r})] e^{i\mathbf{k}\cdot\mathbf{r}}\, d\mathbf{r}, \tag{6.18}$$

so that it becomes

$$\Omega \sum_i \sum_{\mathbf{k}} n_{\mathbf{k}}^* g_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{r}_i}. \tag{6.19}$$

Since the transforms can be pre-calculated, the one-body term is not computationally costly. The use of 3D transforms means that the simulation cell may no longer have infinite extent in the non-periodic direction. To avoid overlapping, the electron density must be restricted to a range w, where the size of the simulation cell in this direction is at least $2w$.[1] The requirement of finite extent is not unreasonable for quasi-2D systems: the electron density usually tends exponentially to zero beyond a certain point.

6.3 Results

In this section, the quasi-2D versions of the MPC and Ewald interactions will be compared; the test system is the jellium slab described in chapter 5. The chosen density parameter is 2.07, which corresponds to aluminium; this density is very frequently studied. The simulation cell is square in the xy-plane, with the size determined by the number of electrons being used:

$$L = \sqrt{\frac{4\pi N r_s^3}{3s}}. \tag{6.20}$$

The number of electrons is N, while s is the slab width, chosen to be 18.63 in these investigations. At this width, LDA calculations reveal that six sub-bands are occupied.

The trial wave function is defined by the following set of equations:

$$\Psi(\mathbf{X}) = e^{J(\mathbf{X})} D^{\uparrow}(\mathbf{R}^{\uparrow}) D^{\downarrow}(\mathbf{R}^{\downarrow}) \tag{6.21}$$

$$J(\mathbf{X}) = -\sum_{i>j} u_{\sigma_i\sigma_j}(r_{ij}) + \sum_i \chi(z_i) \tag{6.22}$$

$$u_{\sigma_i\sigma_j}(r_{ij}) = \frac{A}{r_{ij}}\left[1 - \exp\left(-\frac{r_{ij}}{F_{\sigma_i\sigma_j}}\right)\right] \exp\left(-\frac{r_{ij}^2}{L_c^2}\right) \tag{6.23}$$

$$\chi(z_i) = \sum_k c_k \sin k z_i. \tag{6.24}$$

The motivation for choosing this form for the trial wave function was detailed in section 3.4.1. The variational parameters are A and c_k; $F_{\sigma_i\sigma_j}$ is related to A by the

[1] This enforced zero-padding is standard when calculating a convolution [62].

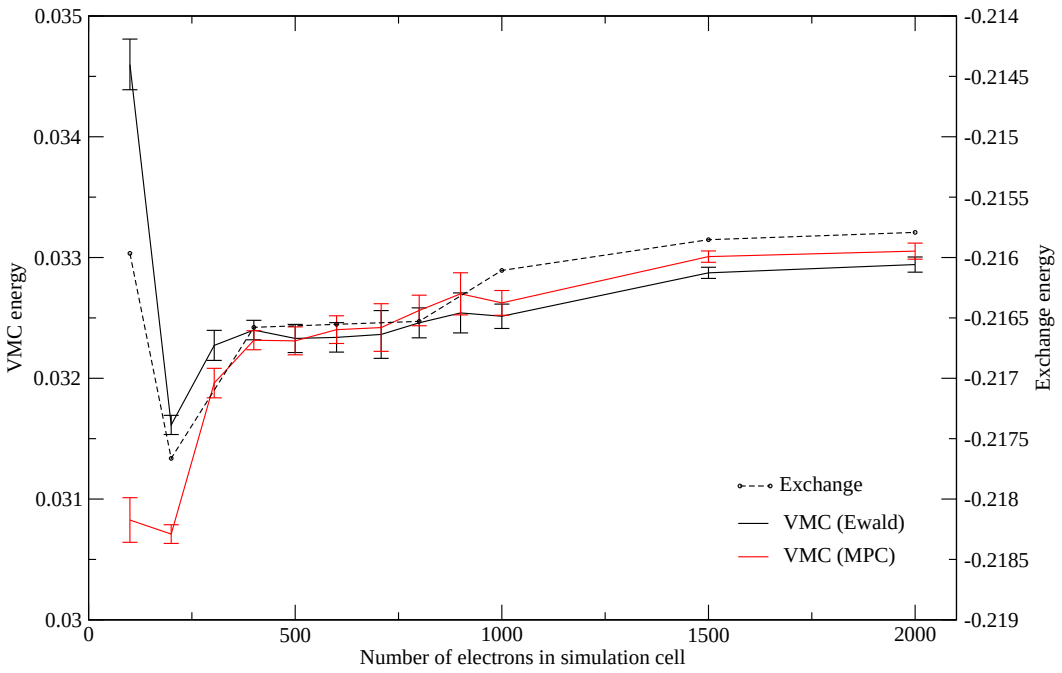


Figure 6.1: The total VMC energy per electron as a function of system size, comparing the Ewald and MPC interactions. The exchange energy given by equation (6.25) is also plotted (note the different vertical scale). Energies, as usual, are in Ha.

cusp conditions (see section 3.4.1 and appendix B). Note that the geometry is such that $\chi(z_i)$ is symmetric about the centre of the slab. The value of the parameter L_c is chosen to ensure that $u(r_{ij})$ decays to zero before r_{ij} approaches the size of the simulation cell; this avoids introducing unwanted gradient discontinuities into the wave function. The single-electron orbitals which make up the determinants $D^{\uparrow}$ and $D^{\downarrow}$ are taken from LDA calculations.

Both VMC and DMC simulations were performed; the results are shown in figures 6.1 and 6.2. Because of the difficulty of optimising the variational parameters in the Jastrow factor, the VMC simulations were carried out with no Jastrow factor in order to avoid introducing a bias caused by optimisation of varying quality. For

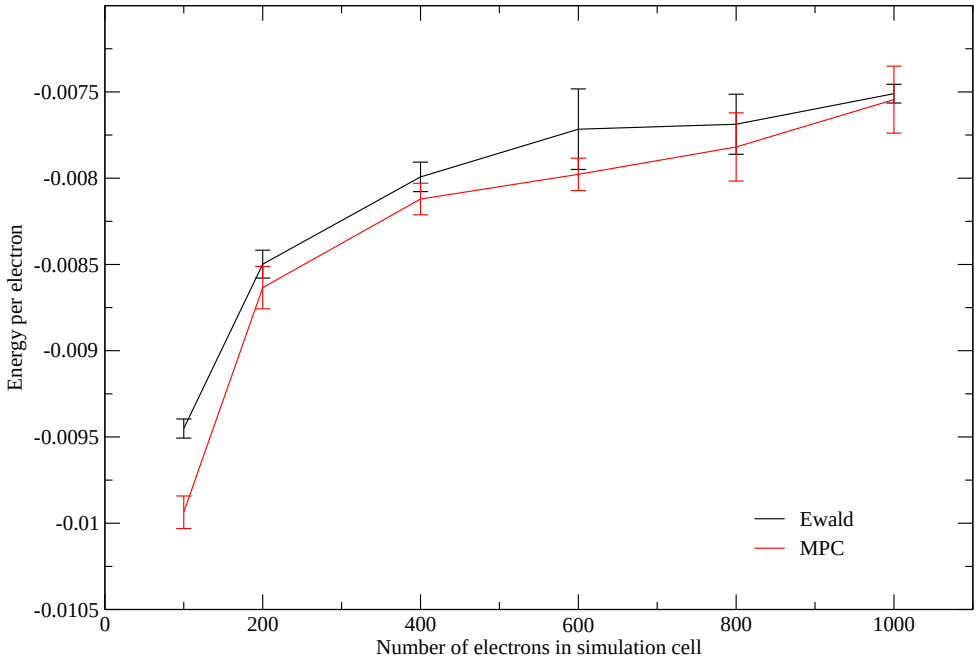


Figure 6.2: The total energy per electron as a function of system size, obtained using fixed-node DMC.

this reason, the exchange energy[2] is also plotted in figure 6.1:

$$E_{\mathrm{X}} = -\sum_{n}\sum_{n'}\iint_{\mathrm{cell}} \left[v_{\mathrm{E}}(\mathbf{r}-\mathbf{r}') - \xi\right]\phi_n^*(\mathbf{r})\phi_n(\mathbf{r}')\phi_{n'}(\mathbf{r})\phi_{n'}^*(\mathbf{r}')\,d\mathbf{r}\,d\mathbf{r}' \tag{6.25}$$

where ϕ_n is the nth single-electron LDA orbital. Since the Jastrow factor cannot alter the location of the nodes, and therefore does not affect the fixed-node DMC energy, it was included in the DMC trial wave functions to improve efficiency.

A correction of the form $(E_{\infty}^{\mathrm{LDA}} - E_N^{\mathrm{LDA}})$ has been applied to all the results to account for the independent-particle finite-size effect (see section 4.1.1).

It is evident from figures 6.1 and 6.2 that the results obtained using the MPC and Ewald interactions are in good agreement, both displaying a fairly slow convergence with respect to the system size. This is unexpected, and contrasts with the 3D case, where the MPC interaction improves the rate of convergence significantly; the

[2]This quantity is not the true exchange energy, which requires optimised Hartree-Fock orbitals; rather, it is a hybrid of the LDA and Hartree-Fock theory.

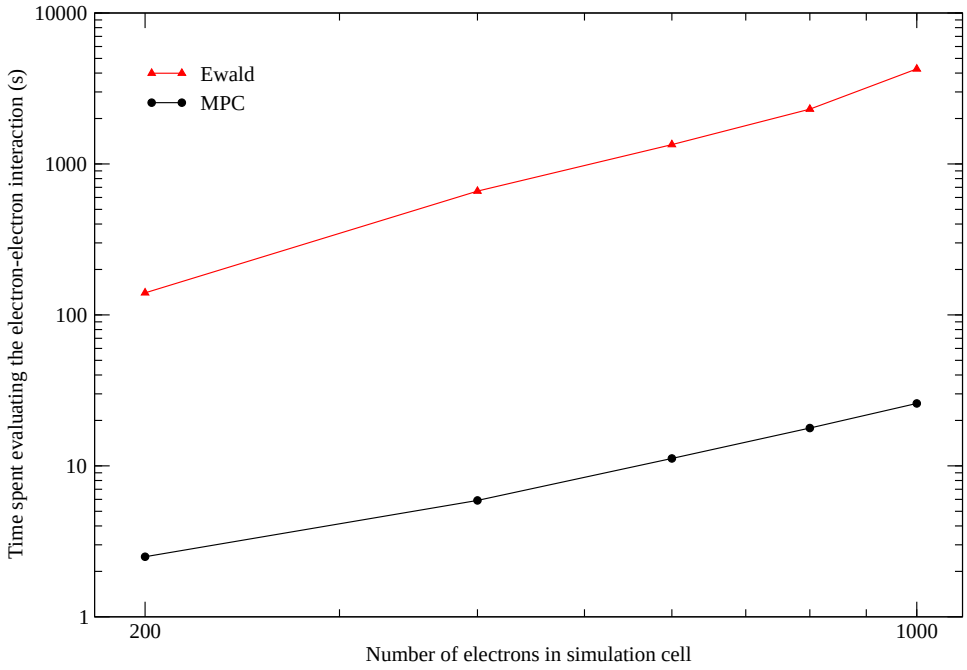


Figure 6.3: The time taken to calculate the electron-electron interaction during an 8000-move VMC simulation, comparing the quasi-2D versions of the MPC and Ewald interactions.

reasons for this will be discussed in the following sections. The slow convergence of the QMC results is reflected in the exchange energy.

An area in which the new interaction does improve on the old is computation time, which is illustrated in figure 6.3. The figure clearly shows that in these calculations, the time required to calculate the Ewald interaction is $\mathcal{O}[N^2]$. The MPC calculation should also be dominated by the $\mathcal{O}[N^2]$ term for large system sizes, but the two-body function being evaluated is much less costly; for the system sizes used here, the $\mathcal{O}[N]$ term also contributes significantly. The time required for the MPC pre-calculation is not included in the figure; for any serious calculation it is negligible in comparison with the time spent evaluating the electron-electron interaction energy during the simulation itself.

In the DMC simulations of the jellium slab, evaluating the electron-electron interaction using the quasi-2D Ewald method takes up a significant fraction of the total calculation time (often more than 50%). Figure 6.3 shows that the MPC

interaction is nearly two orders of magnitude faster than the Ewald; this is a more dramatic improvement than in the 3D case.

6.4 Analysis

The results presented in the previous section demonstrate that although the quasi-2D MPC interaction does enjoy the expected advantage in speed over the Ewald sum, it does not reduce the finite-size errors.

This may be understood by considering the expected error incurred by using the Ewald interaction, following the arguments of sections 4.1.2 and 6.1:

$$E_{\mathrm{e-e}}^{\mathrm{EW}} - E_{\mathrm{e-e}}^{\mathrm{MPC}} = \frac{1}{2} \iint_{\mathrm{cell}} n(\mathbf{r}) n_{\mathrm{XC}}(\mathbf{r}, \mathbf{r}') [v_{\mathrm{E}}(\mathbf{r} - \mathbf{r}') - \xi - f(\mathbf{r} - \mathbf{r}')] \, d\mathbf{r} \, d\mathbf{r}'. \tag{6.26}$$

The exchange-correlation hole described by n_{XC} is generally short-ranged. An approximation to equation (6.26) may therefore be obtained by using the small-r and large-L expansion

$$v_{\mathrm{E}}(\mathbf{r}) - \xi = \frac{1}{r} - \frac{C}{L^3}\left(z^2 - \frac{r_\parallel^2}{2}\right) + \mathcal{O}\left[r^4\right] + \mathcal{O}\left[e^{-L^2/\sigma^2}\right] \tag{6.27}$$

which is derived in section A.2; here C is a constant. Combining equations (6.26) and (6.27) gives

$$\begin{aligned} E_{\mathrm{e-e}}^{\mathrm{EW}} - E_{\mathrm{e-e}}^{\mathrm{MPC}} = &- \frac{C}{2L^3} \iint_{\mathrm{cell}} n(\mathbf{r}) n_{\mathrm{XC}}(\mathbf{r}, \mathbf{r}') \left[(z - z')^2 - \frac{1}{2}(\mathbf{r}_\parallel - \mathbf{r}'_\parallel)^2\right] d\mathbf{r} \, d\mathbf{r}' \\ &+ \text{higher-order terms.} \end{aligned} \tag{6.28}$$

Thus, a large cancellation of the error may be anticipated in regions where the parallel and perpendicular directions are equivalent; this is the case deep inside the slab, where the hole should be approximately spherical. Near the slab edges this is no longer true. However, the hole is expected to expand in these lower-density regions, rendering the small-r expansion inappropriate; the higher-order terms are no longer negligible. The subject of the expansion of the hole in the surface regions will be discussed in the following section.

In 3D, the small-r expansion of the Ewald interaction is spherically-symmetric [25] and there is no error cancellation; the finite-size error which results is corrected by the MPC interaction. The situation is different in quasi-2D systems; there must be an alternative mechanism giving rise to the slow convergence seen in figures 6.1 and 6.2.

6.5 The exchange-correlation hole

One possible cause of the remaining finite-size error seen with both the MPC and Ewald interactions is 'squashing' of the exchange-correlation hole by the simulation cell. It is expected that close to the slab surface, where the electron density is lower, the hole should expand in the xy-direction; if the lateral size of the hole were to approach the size of the cell, the shape of the hole would be modified, creating a finite-size error. To investigate this possibility, a simple model of the reaction to an imposed external charge will be analysed.

Let a charge $\delta\rho_{\text{ext}}$ be introduced to the slab system. This will cause some rearrangement of the slab electrons, leading to an induced charge density $\delta\rho_{\text{ind}}$. The total potential is then given by Poisson's equation:

$$\nabla^2\delta\phi_{\text{tot}} = -4\pi(\delta\rho_{\text{ext}} + \delta\rho_{\text{ind}}). \tag{6.29}$$

A second relation may be obtained by considering the electron density to be slowly-varying, as in the Thomas-Fermi theory. In this picture, when the electron energy levels are locally shifted by the small potential $\delta\phi_{\text{tot}}$, the change in the local electron density is approximately

$$\delta n_{\text{ind}} = g(\epsilon_{\text{F}})\delta\phi_{\text{tot}} \tag{6.30}$$

where $g(\epsilon_{\text{F}})$ is the density of states at the Fermi level:

$$g(\epsilon_{\text{F}}) = \left(\frac{3n}{\pi^4}\right)^{1/3}. \tag{6.31}$$

Since the induced charge density is created entirely by rearrangement of electrons, equation (6.30) becomes

$$\delta\rho_{\text{ind}} = -\left(\frac{3n}{\pi^4}\right)^{1/3}\delta\phi_{\text{tot}}. \tag{6.32}$$

Substitution in equation (6.29) leads to the following equation for $\delta\phi_{\text{tot}}$:

$$\left[\nabla^2 - 4\left(\frac{3n}{\pi}\right)^{1/3}\right]\delta\phi_{\text{tot}} = -4\pi\delta\rho_{\text{ext}}. \tag{6.33}$$

In order to solve this equation, the functions are first expanded in Fourier series in the xy-plane:

$$\delta\phi_{\text{tot}}(\mathbf{r}) = \sum_{\mathbf{k}_\parallel}\delta\tilde{\phi}_{\text{tot}}(\mathbf{k}_\parallel, z)e^{-i\mathbf{k}_\parallel\cdot\mathbf{r}_\parallel} \tag{6.34}$$

$$\delta\rho_{\text{ext}}(\mathbf{r}) = \sum_{\mathbf{k}_\parallel}\delta\tilde{\rho}_{\text{ext}}(\mathbf{k}_\parallel, z)e^{-i\mathbf{k}_\parallel\cdot\mathbf{r}_\parallel}. \tag{6.35}$$

This step has introduced in-plane periodicity into the problem, which is desirable, since the aim is to investigate the effect of cell size on the hole. The equation to be solved is now

$$\left[\frac{d^2}{dz^2} - k_\parallel^2 - 4\left(\frac{3n(z)}{\pi}\right)^{1/3}\right]\delta\tilde{\phi}_{\text{tot}}(\mathbf{k}_\parallel, z) = -4\pi\delta\tilde{\rho}_{\text{ext}}(\mathbf{k}_\parallel, z). \tag{6.36}$$

At this stage, it is useful to supply the form of the external charge. Since the idea is to investigate the hole around an electron, the appropriate form is

$$\delta\rho_{\text{ext}}(\mathbf{r}) = -\sum_{\mathbf{R}}\delta(\mathbf{r} - \mathbf{r}_0 - \mathbf{R}) + \frac{1}{L^2 s}\Theta(z)\Theta(s - z). \tag{6.37}$$

The external charge must be periodic if it is to be expanded in a Fourier series, as in equation (6.35); this is ensured by the sum over the in-plane lattice vectors $\mathbf{R}$. The positive charge (which is uniform over the slab) has been added to ensure that the cell remains charge-neutral.

The problem with this charge density is that it leads to a potential $\delta\phi_{\text{tot}}$ which is divergent at $\mathbf{r} = \mathbf{r}_0$. Instead, it is convenient to smear out the charge distribution slightly:

$$\delta\rho_{\text{ext}}(\mathbf{r}) = -\sum_{\mathbf{R}}\frac{1}{(2\pi\sigma^2)^{3/2}}e^{-(\mathbf{r}-\mathbf{r}_0-\mathbf{R})^2/2\sigma^2} + \frac{1}{L^2 s}\Theta(z)\Theta(s - z). \tag{6.38}$$

Choosing the origin so that the external charge is at $(\mathbf{r}_0)_\parallel = \mathbf{0}$ gives the Fourier components as

$$\delta\tilde{\rho}_{\text{ext}}(\mathbf{k}_\parallel, z) = -\frac{1}{L^2\sqrt{2\pi\sigma^2}} e^{-\sigma^2 k_\parallel^2/2} e^{-(z-z_0)^2/2\sigma^2} + \frac{1}{L^2 s}\Theta(z)\Theta(s-z)\delta_{\mathbf{k}_\parallel,\mathbf{0}}. \tag{6.39}$$

This reduces equation (6.36) to the form

$$\left[\hat{L}(z) - k_\parallel^2\right]\delta\tilde{\phi}_{\text{tot}}(\mathbf{k}_\parallel, z) = \frac{4\pi}{L^2\sqrt{2\pi\sigma^2}} e^{-\sigma^2 k_\parallel^2/2} e^{-(z-z_0)^2/2\sigma^2} - \frac{4\pi}{L^2 s}\Theta(z)\Theta(s-z)\delta_{\mathbf{k}_\parallel,\mathbf{0}} \tag{6.40}$$

with

$$\hat{L}(z) = \frac{d^2}{dz^2} - f(z) \tag{6.41}$$

and

$$f(z) = 4\left(\frac{3n(z)}{\pi}\right)^{1/3}. \tag{6.42}$$

This nonhomogeneous equation may be solved by computing the appropriate Green's function, which satisfies the following equation:

$$\left[\hat{L}(z) - k_\parallel^2\right] G(z, z'; k_\parallel) = \delta(z - z'). \tag{6.43}$$

In order to proceed further, an analytic expression for the electron density is required. In the case of a realistic slab, this is not available. However, it is possible to analyse a more simple case: a sharp surface, where the density profile is

$$n(z) = n_0\Theta(z). \tag{6.44}$$

The additional positive charge density which was added to maintain the neutrality of the cell disappears in this limit, which corresponds to letting $s \to \infty$. Some of the essential surface physics may be lost because of the sharp boundary: the true unbounded slab will be treated later using computational methods.

For the simplified surface, the differential equation for the Green's function is

$$\left[\frac{d^2}{dz^2} - k_\parallel^2 - 4\left(\frac{3n_0}{\pi}\right)^{1/3}\Theta(z)\right] G(z, z'; k_\parallel) = \delta(z - z'). \tag{6.45}$$

The solution must not diverge as $z \to \pm\infty$; solving the homogeneous version of the equation in the three regions separately gives

$$G(z, z'; k_\parallel) = \begin{cases} Ae^{k_\parallel z} & z < 0 \\ Be^{\kappa z} + Ce^{-\kappa z} & 0 < z < z' \\ De^{-\kappa z} & z > z' \end{cases} \tag{6.46}$$

where

$$\kappa^2 = k_\parallel^2 + 4\left(\frac{3n_0}{\pi}\right)^{1/3}. \tag{6.47}$$

It is assumed that $z' > 0$; the solution when $z' < 0$ will be given later.

The Green's function must be continuous at $z = 0$ and $z = z'$. The remaining boundary conditions are obtained by integration of equation (6.45) over an infinitesimal region about these points, giving the following constraints on the gradient:

$$\lim_{a\to 0}\left(\left[\frac{dG(z, z'; k_\parallel)}{dz}\right]_{z=a} - \left[\frac{dG(z, z'; k_\parallel)}{dz}\right]_{z=-a}\right) = 0 \tag{6.48}$$

$$\lim_{a\to 0}\left(\left[\frac{dG(z, z'; k_\parallel)}{dz}\right]_{z=z'+a} - \left[\frac{dG(z, z'; k_\parallel)}{dz}\right]_{z=z'-a}\right) = 1. \tag{6.49}$$

Applying these boundary conditions gives the four simultaneous equations required to determine the four constants:

$$A = B + C \tag{6.50}$$

$$k_\parallel A = \kappa(B - C) \tag{6.51}$$

$$Be^{\kappa z'} + Ce^{-\kappa z'} = De^{-\kappa z'} \tag{6.52}$$

$$\kappa(Be^{\kappa z'} - Ce^{-\kappa z'}) + 1 = -\kappa De^{-\kappa z'} \tag{6.53}$$

so that

$$A = -\frac{1}{(\kappa + k_\parallel)}e^{-\kappa z'} \tag{6.54}$$

$$B = -\frac{1}{2\kappa}e^{-\kappa z'} \tag{6.55}$$

$$C = -\frac{(\kappa - k_\parallel)}{2\kappa(\kappa + k_\parallel)}e^{-\kappa z'} \tag{6.56}$$

$$D = -\frac{1}{2\kappa}\left[e^{\kappa z'} + \left(\frac{\kappa - k_\parallel}{\kappa + k_\parallel}\right)e^{-\kappa z'}\right]. \tag{6.57}$$

The solution is therefore

$$G(z,z';k_\parallel) = -\frac{1}{2\kappa}\begin{cases} \dfrac{2\kappa}{\kappa+k_\parallel}e^{k_\parallel z-\kappa z'} & z<0 \\ e^{-\kappa|z-z'|}+\left(\dfrac{\kappa-k_\parallel}{\kappa+k_\parallel}\right)e^{-\kappa(z+z')} & z>0 \end{cases} \qquad (z'>0). \tag{6.58}$$

An identical method gives the Green's function when $z'<0$:

$$G(z,z';k_\parallel) = -\frac{1}{2k_\parallel}\begin{cases} e^{-k_\parallel|z-z'|}+\left(\dfrac{k_\parallel-\kappa}{k_\parallel+\kappa}\right)e^{-k_\parallel(z+z')} & z<0 \\ \dfrac{2k_\parallel}{k_\parallel+\kappa}e^{k_\parallel z'-\kappa z} & z>0 \end{cases} \qquad (z'<0). \tag{6.59}$$

The Green's function has the required symmetry property $G(z,z';k_\parallel)=G(z',z;k_\parallel)$. Once $G(z,z';k_\parallel)$ is known, the potential may be obtained by integration:

$$\delta\tilde{\phi}_{\rm tot}(\mathbf{k}_\parallel,z) = \int_{-\infty}^{\infty} G(z,z';k_\parallel)\left(\frac{4\pi}{L^2\sqrt{2\pi\sigma^2}}e^{-\sigma^2k_\parallel^2/2}e^{-(z'-z_0)^2/2\sigma^2}\right)dz'. \tag{6.60}$$

Since the aim is to calculate the induced change in the charge density (which is proportional to the original density, and hence zero outside the slab), it is only necessary to consider the region $z>0$. The integration then gives

$$\begin{aligned}
\delta\tilde{\phi}_{\rm tot}(\mathbf{k}_\parallel,z) = &-\left(\frac{4\pi}{L^2\sqrt{2\pi\sigma^2}}e^{-\sigma^2k_\parallel^2/2}\right)\Bigg[\int_{-\infty}^{0}\frac{1}{\kappa+k_\parallel}e^{k_\parallel z'-\kappa z}e^{-(z'-z_0)^2/2\sigma^2}\,dz' \\
&+\frac{1}{2\kappa}\left(\frac{\kappa-k_\parallel}{\kappa+k_\parallel}\right)\int_0^{\infty}e^{-\kappa(z+z')}e^{-(z'-z_0)^2/2\sigma^2}\,dz' \\
&+\frac{1}{2\kappa}\int_0^{z}e^{-\kappa(z-z')}e^{-(z'-z_0)^2/2\sigma^2}\,dz'+\frac{1}{2\kappa}\int_z^{\infty}e^{-\kappa(z'-z)}e^{-(z'-z_0)^2/2\sigma^2}\,dz'\Bigg] \\
= &-\frac{\pi}{L^2\kappa}\Bigg[\frac{2\kappa}{\kappa+k_\parallel}e^{-\kappa z+k_\parallel z_0}\left(1-\,\mathrm{erf}\left(\frac{z_0+k_\parallel\sigma^2}{\sqrt{2}\sigma}\right)\right) \\
&+\frac{\kappa-k_\parallel}{(\kappa+k_\parallel)}e^{(\kappa^2-k_\parallel^2)\sigma^2/2-\kappa z-\kappa z_0}\left(1+\,\mathrm{erf}\left(\frac{z_0-\kappa\sigma^2}{\sqrt{2}\sigma}\right)\right) \\
&+e^{(\kappa^2-k_\parallel^2)\sigma^2/2-\kappa z+\kappa z_0}\left(\,\mathrm{erf}\left(\frac{z-z_0-\kappa\sigma^2}{\sqrt{2}\sigma}\right)+\,\mathrm{erf}\left(\frac{z_0+\kappa\sigma^2}{\sqrt{2}\sigma}\right)\right) \\
&+e^{(\kappa^2-k_\parallel^2)\sigma^2/2+\kappa z-\kappa z_0}\left(1-\,\mathrm{erf}\left(\frac{z-z_0+\sigma^2\kappa}{\sqrt{2}\sigma}\right)\right)\Bigg].
\end{aligned} \tag{6.61}$$

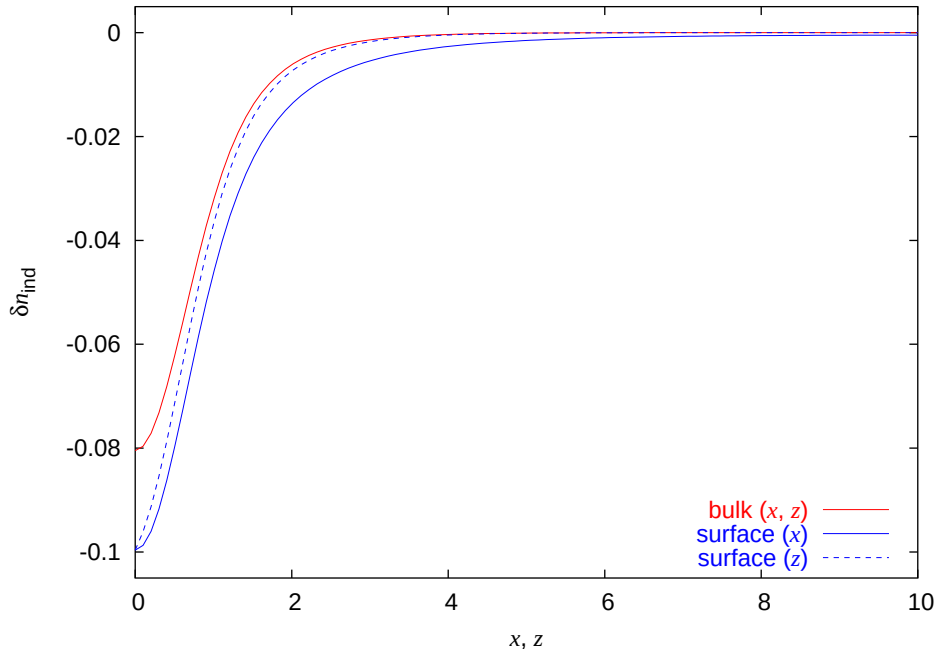


Figure 6.4: The charge disturbance created by an electron-like test charge at the surface ($z_0 = 0$) and in the bulk ($z_0 = 10$). The plots show the value of $\delta n_{\text{ind}}(\mathbf{r})$ at the points $(x, 0, z_0)$ (labelled x) and $(0, 0, z_0 + z)$ (labelled z). In the bulk, the x and z curves are identical.

The induced charge density is then readily obtained from equations (6.30), (6.31) and (6.34) and is plotted in figure 6.4. It is clear from the figure that the 'hole' spreads out in the xy-plane when the charge disturbance is near to the surface. In the bulk, the parallel and perpendicular directions are equivalent; at the surface, the hole is deeper and dies away more slowly in the x-direction. The important issue is the effect of the cell size on the shape of the hole. Figure 6.5 shows that the change is much more pronounced when the disturbance is at the surface, because of the greater extent of the hole in the xy-direction. The slow decay of the surface hole means that the result in the small system is significantly different from that in the two larger cells. The bulk hole, which decays more rapidly, does not display the same effect.

Returning to a more realistic slab, where the analytic form of $n(z)$ is not known, the procedure for calculating the induced charge density will now be outlined.

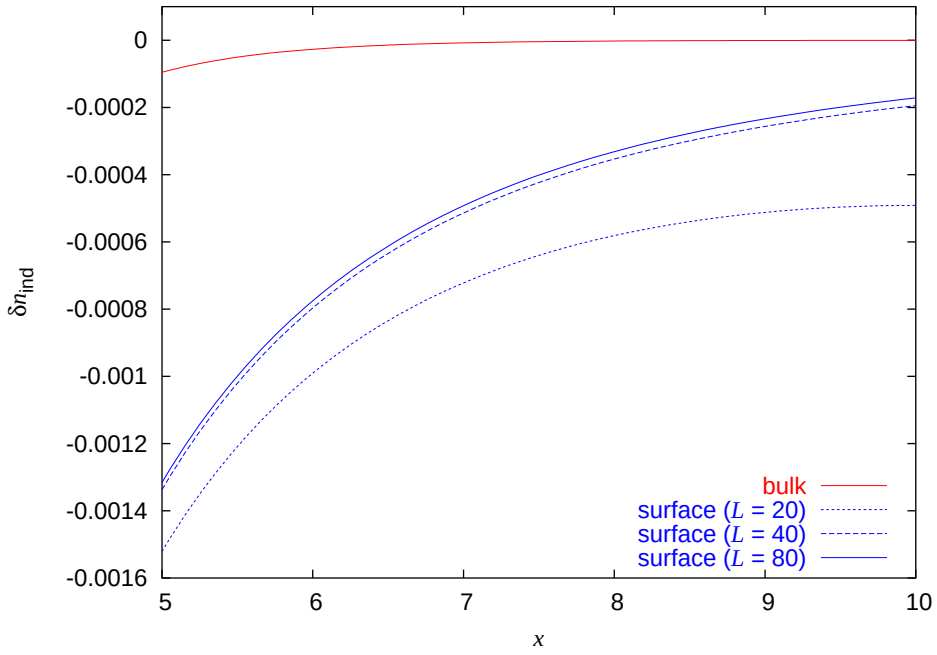


Figure 6.5: The effect of cell size on the shape of the hole. Systems of size $L = 20$, $L = 40$ and $L = 80$ are compared at a distance of 10 from the centre of the imposed charge; this corresponds to the edge of the smallest cell, where any difference should be most clearly visible. The curves for the bulk hole at each of the three sizes are indistinguishable on this scale.

One way to obtain the Green's function is through the eigenfunctions of $\hat{L}(z)$:

$$\hat{L}(z)u_n(z) = \lambda_n u_n(z). \tag{6.62}$$

Once again, a Fourier series representation is useful here. Letting

$$u_n(z) = \sum_{k_z} \tilde{u}_n(k_z) e^{-ik_z z} \tag{6.63}$$

$$f(z) = \sum_{k_z} \tilde{f}(k_z) e^{-ik_z z} \tag{6.64}$$

turns the eigenvalue equation into

$$-k_z^2 \tilde{u}_n(k_z) - \sum_{k_z'} \tilde{f}(k_z')\tilde{u}_n(k_z - k_z') = \lambda_n \tilde{u}_n(k_z). \tag{6.65}$$

Finding the $\tilde{u}_n$ now amounts to finding the eigenvectors of the following matrix:

$$\begin{pmatrix} -\tilde{f}(0) & -\tilde{f}\left(-\frac{2\pi}{w}\right) & -\tilde{f}\left(-\frac{4\pi}{w}\right) & \cdots & -\tilde{f}\left(\frac{2\pi}{w}\right) \\ -\tilde{f}\left(\frac{2\pi}{w}\right) & -\left(\frac{2\pi}{w}\right)^2 - \tilde{f}(0) & -\tilde{f}\left(-\frac{2\pi}{w}\right) & \cdots & -\tilde{f}\left(\frac{4\pi}{w}\right) \\ -\tilde{f}\left(\frac{4\pi}{w}\right) & -\tilde{f}\left(\frac{2\pi}{w}\right) & -\left(\frac{4\pi}{w}\right)^2 - \tilde{f}(0) & \cdots & -\tilde{f}\left(\frac{6\pi}{w}\right) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -\tilde{f}\left(-\frac{2\pi}{w}\right) & -\tilde{f}\left(-\frac{4\pi}{w}\right) & -\tilde{f}\left(-\frac{6\pi}{w}\right) & \cdots & -\left(-\frac{2\pi}{w}\right)^2 - \tilde{f}(0) \end{pmatrix}. \tag{6.66}$$

Here w is the chosen cell size in the z-direction, so that $k_z = 2m\pi/w$. In order that the matrix be finite, m must be restricted. Using the standard discrete Fourier ordering (as in the matrix itself) gives $m = 0, 1, 2, \ldots, N/2, -N/2+1, \ldots, -1$; the matrix is then of size[3] $N \times N$.

Obtaining the eigenvalues and eigenvectors of this finite matrix is a standard computational operation. The Green's function can now be constructed:

$$\begin{aligned} G(z, z'; k_\parallel) &= \sum_n \frac{u_n(z) u_n^*(z')}{\lambda_n - k_\parallel^2} \\ &= \sum_n \frac{1}{\lambda_n - k_\parallel^2} \sum_{k_z} \sum_{k_z'} e^{-i(k_z z - k_z' z')} \tilde{u}_n(k_z) \tilde{u}_n^*(k_z'). \end{aligned} \tag{6.67}$$

Once the Green's function is known, the response to the imposed charge distribution can be calculated, as in equation (6.60):

$$\begin{aligned} \delta\tilde{\phi}_{\text{tot}}(\mathbf{k}_\parallel, z) &= -4\pi \int_0^w G(z, z'; k_\parallel) \delta\tilde{\rho}_{\text{ext}}(\mathbf{k}_\parallel, z')\, dz' \\ &= -4\pi \sum_n \frac{1}{\lambda_n - k_\parallel^2} \sum_{k_z} \sum_{k_z'} e^{-ik_z z} \tilde{u}_n(k_z) \tilde{u}_n^*(k_z') \int_0^w e^{ik_z' z'} \delta\tilde{\rho}_{\text{ext}}(\mathbf{k}_\parallel, z')\, dz'. \end{aligned} \tag{6.68}$$

The charge distribution $\delta\rho_{\text{ext}}$ is not quite the same as the one used previously, since it must now be periodic in all three dimensions. The definition given in equation (6.38) may still be used, with the understanding that $\mathbf{R}$ now represents a vector of

[3] N is assumed to be even.

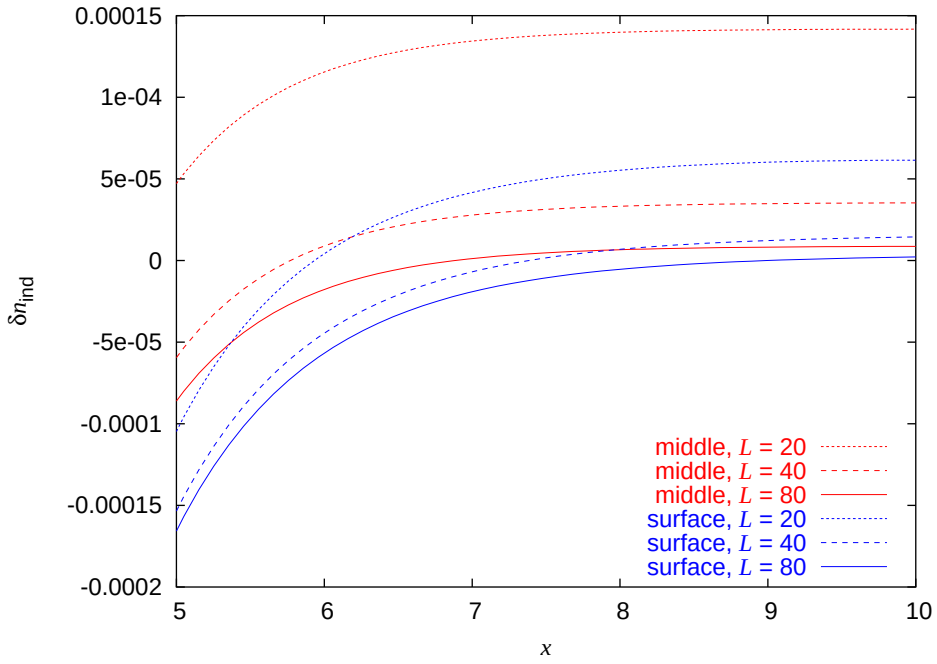


Figure 6.6: The Thomas-Fermi hole for different cell sizes in a more realistic model of the slab. The induced electron density δn_{ind} is plotted at the points $(x, 0, z_0)$, where z_0 corresponds to the slab centre (labelled 'middle') or the conventional slab edge (labelled 'surface').

the 3D lattice and that the Θ-functions now also become periodic. The transform of $\delta\tilde{\rho}_{\text{ext}}$ in the z-direction appearing in equation (6.68) is then

$$\int_0^w e^{ik'_z z'}\,\delta\tilde{\rho}_{\text{ext}}(\mathbf{k}_\parallel, z')\,dz' = -\frac{1}{L^2}e^{-\sigma^2(k_\parallel^2+k_z'^2)/2+ik'_z z_0} + \frac{1}{L^2}e^{ik'_z s/2}\left(\frac{\sin(k'_z s/2)}{k'_z s/2}\right)\delta_{\mathbf{k}_\parallel,\mathbf{0}}. \tag{6.69}$$

Combining this with equations (6.30), (6.31), (6.34) and (6.68) gives the induced charge density:

$$\delta n_{\text{ind}}(\mathbf{r}) = \frac{4}{L^2}\left(\frac{3n(z)}{\pi}\right)^{1/3}\sum_{\mathbf{k}} e^{-i\mathbf{k}\cdot\mathbf{r}}\sum_n \frac{\tilde{u}_n(k_z)}{\lambda_n - k_\parallel^2}\sum_{k'_z}\tilde{u}_n^*(k'_z)\left[e^{ik'_z z_0-\sigma^2(k_\parallel^2+k_z'^2)/2}\right.$$
$$\left. - e^{ik'_z s/2}\left(\frac{\sin(k'_z s/2)}{k'_z s/2}\right)\delta_{\mathbf{k}_\parallel,\mathbf{0}}\right]. \tag{6.70}$$

This is plotted in figure 6.6. On comparing figures 6.5 and 6.6, one major difference stands out immediately: far away from the imposed charge, the induced electron

density tends to zero in the idealised surface system, but to some non-zero positive value in the more realistic slab model.

This is because the net change in the electron density is forced to be zero in the slab system (which is finite); since the electron density is reduced in the vicinity of the charge disturbance, it must be increased elsewhere. This constraint does not apply to the idealised surface system (which is infinite).

To see this, consider integrating the original differential equation (6.29) over the cell:

$$\int_{\text{cell}} \nabla^2 \delta\phi_{\text{tot}}(\mathbf{r})\, d^3r - 4\pi \int_{\text{cell}} \delta n_{\text{ind}}(\mathbf{r})\, d^3r = -4\pi \int_{\text{cell}} \delta\rho_{\text{ext}}(\mathbf{r})\, d^3r. \tag{6.71}$$

The external charge density was chosen to be neutral overall, giving

$$\int_{\text{cell}} \delta\rho_{\text{ext}}(\mathbf{r})\, d^3r = 0. \tag{6.72}$$

When the potential $\delta\phi_{\text{tot}}$ is forced to be periodic, the first integral in equation (6.71) also gives zero. This is evident on substitution of the Fourier series representation:

$$\begin{aligned} \int_{\text{cell}} \nabla^2 \delta\phi_{\text{tot}}(\mathbf{r})\, d^3r &= \int_{\text{cell}} \nabla^2 \sum_{\mathbf{k}} \delta\tilde{\phi}_{\text{tot}}(\mathbf{k}) e^{-i\mathbf{k}\cdot\mathbf{r}}\, d^3r \\ &= -\sum_{\mathbf{k}} k^2 \delta\tilde{\phi}_{\text{tot}}(\mathbf{k}) \int_{\text{cell}} e^{-i\mathbf{k}\cdot\mathbf{r}}\, d^3r \\ &= 0. \end{aligned} \tag{6.73}$$

Thus, for a finite system with periodic boundary conditions, the net change to the electron density is zero:

$$\int_{\text{cell}} \delta n_{\text{ind}}(\mathbf{r})\, d^3r = 0. \tag{6.74}$$

This effect becomes smaller as the system size increases, as can be seen in figure 6.6; in an infinite system like the idealised surface, it disappears.

Since the aim is to investigate the exchange-correlation hole, it is helpful to establish a link between this entity and δn_{ind}. For this, a new notation is required. The electron density in a system of N electrons at the point $\mathbf{r}$ is labelled $n(\mathbf{r}; N)$. When one electron is fixed at the position $\mathbf{r}'$, the density at $\mathbf{r}$ is $n(\mathbf{r}|\mathbf{r}'; N)$; this is

similar to a conditional probability. The exchange-correlation hole for an electron at $\mathbf{r}'$ is then defined by the formula

$$n(\mathbf{r}|\mathbf{r}';N) = n(\mathbf{r};N) + n_{\mathrm{XC}}(\mathbf{r},\mathbf{r}';N). \tag{6.75}$$

In the calculations above, the external charge density was chosen to model an electron; after fixing this electron at $\mathbf{r}'$, the change to the density of the remaining free electrons was estimated:

$$\delta n_{\mathrm{ind}}(\mathbf{r}|\mathbf{r}';N) \approx n(\mathbf{r}|\mathbf{r}';N+1) - n(\mathbf{r};N). \tag{6.76}$$

Note the change from N to $N+1$ in the conditional density, which arises because an electron has been added to the system. Combining these equations then gives

$$\begin{aligned} n_{\mathrm{XC}}(\mathbf{r},\mathbf{r}';N+1) &= n(\mathbf{r}|\mathbf{r}';N+1) - n(\mathbf{r};N+1) \\ &\approx \delta n_{\mathrm{ind}}(\mathbf{r}|\mathbf{r}';N) + n(\mathbf{r};N) - n(\mathbf{r};N+1). \end{aligned} \tag{6.77}$$

For large N, adding an extra electron should not change the density dramatically. The following approximation may therefore be used:

$$n(\mathbf{r};N+1) \approx \left(\frac{N+1}{N}\right) n(\mathbf{r};N). \tag{6.78}$$

The exchange-correlation hole is then

$$n_{\mathrm{XC}}(\mathbf{r},\mathbf{r}';N+1) \approx \delta n_{\mathrm{ind}}(\mathbf{r}|\mathbf{r}';N) - \frac{1}{N} n(\mathbf{r};N). \tag{6.79}$$

This function is plotted in figure 6.7, which shows the characteristic spreading-out of the hole near the surface. As in the idealised system, the surface hole is curtailed by the cell boundary, although the effect is weaker. This supports the idea that the remaining finite-size error is caused by the exchange-correlation hole being squashed into the simulation cell. This error would act to reduce the energy of the system, and therefore has the correct sign.

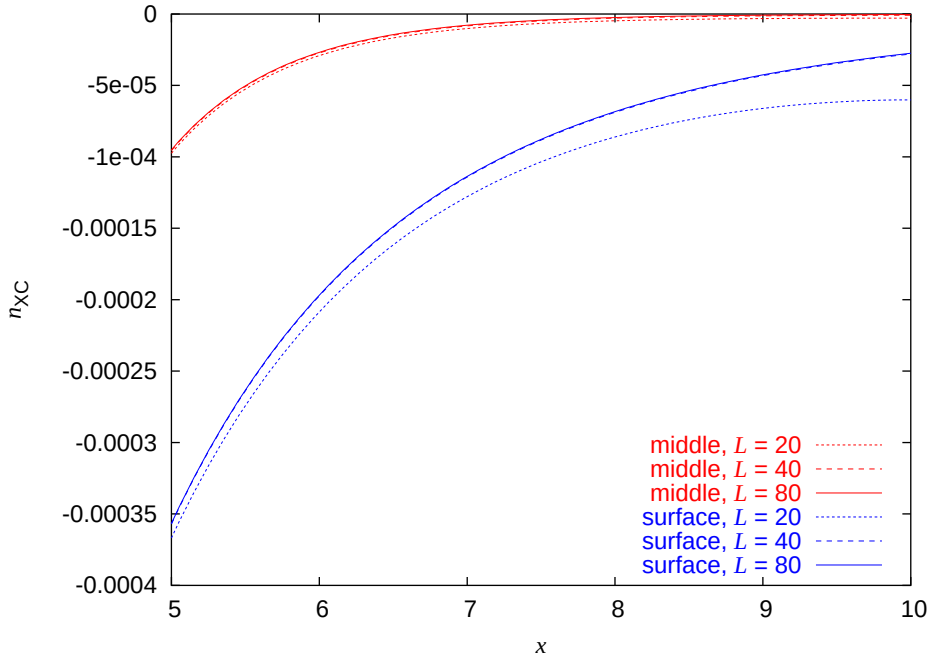


Figure 6.7: The exchange-correlation hole for a realistic slab.

Chapter 7

The electronic ground-state wave function from classical plasmon normal modes

A prerequisite for a successful quantum Monte Carlo simulation is a trial wave function of good quality. Conventionally, the many-electron trial wave function takes the form

$$\Psi_T(\mathbf{X}) = e^{J(\mathbf{X})} D^{\uparrow}(\mathbf{R}^{\uparrow}) D^{\downarrow}(\mathbf{R}^{\downarrow}) \tag{7.1}$$

where $D^{\uparrow}$ and $D^{\downarrow}$ are up- and down-spin determinants made up of density-functional or Hartree-Fock single-electron orbitals. The quality of these orbitals sets a hard limit on the accuracy of the QMC simulation in variational or fixed-node diffusion Monte Carlo. The Jastrow factor, J, cannot alter the position of the nodes, and therefore cannot improve the fixed-node DMC energy; however, a good Jastrow factor significantly improves both the ground-state energy in VMC and the efficiency of fixed-node DMC simulations. In certain circumstances, a poor-quality Jastrow factor can make DMC simulations impossible, because the consequent population fluctuations become unmanageable. This is the case for the jellium slab system, which also suffers from the problem that optimising the Jastrow factor is very difficult; in systems such as this, it is very important to obtain as much information as

possible about the correct form of the wave function by analytical means. One way to obtain a better Jastrow factor is through a consideration of plasmons.

A plasmon is a collective excitation of electrons. In a system with a homogeneous electron density, plasmons are well-defined for wavelengths above some critical value; for wavelengths below this value, the plasmons are able to decay to form electron-hole pairs.

It is not surprising that the correlation of pairs of electrons can be related to the collective electron motion described by plasmons; in the past, several authors have studied this relationship for homogeneous [9] and inhomogeneous systems [28].

The aim of this chapter is to clarify this connection between long-wavelength plasmons and the electronic ground-state wave function. A prescription for obtaining a Jastrow factor for general inhomogeneous systems will be presented which is consistent with the previous work, though based on a more 'physical' approach; in addition, it will be shown how a knowledge of the classical plasmon normal modes allows an alternative method for writing down J.

The electronic Schrödinger equation (equation (2.4)) contains only the electrostatic interaction energy: any interaction arising from currents (or mediated by a magnetic field) is neglected. This has important consequences for the plasmon Hamiltonian, and has a strong bearing on the analysis in the following sections.

7.1 Derivation of the classical plasmon Hamiltonian and Lagrangian

In general, a plasmon is associated with a change in the electron density. The equilibrium electron density will be denoted by $\bar{n}(\mathbf{r})$, while the time-dependent perturbation to the charge density associated with the plasmon will be written as $\rho(\mathbf{r}, t)$; this is assumed to be small, along with all other time-dependent quantities. The total electron density is therefore

$$n(\mathbf{r}, t) = \bar{n}(\mathbf{r}) - \frac{\rho(\mathbf{r}, t)}{e} \tag{7.2}$$

In what follows, terms of order ρ^2 will be neglected: the system will be linearised.

Classically, the (linearised) kinetic energy of the electron gas is

$$T = \int_V \frac{1}{2} m_e v^2(\mathbf{r}, t) \bar{n}(\mathbf{r}) \, d^3r \tag{7.3}$$

where $\mathbf{v}(\mathbf{r}, t)$ is the electron velocity field. The current density is given by

$$\mathbf{J}(\mathbf{r}, t) = -\bar{n}(\mathbf{r}) e \mathbf{v}(\mathbf{r}, t). \tag{7.4}$$

The kinetic energy may be written in terms of the current density as

$$T = \frac{1}{2\epsilon_0} \int_V \frac{J^2(\mathbf{r}, t)}{\omega_p^2(\mathbf{r})} \, d^3r \tag{7.5}$$

where

$$\omega_p(\mathbf{r}) = \sqrt{\frac{\bar{n}(\mathbf{r}) e^2}{m_e \epsilon_0}} \tag{7.6}$$

is the plasma frequency.

The potential energy is the electrostatic self-interaction energy of the *plasmon* charge density:

$$V = \frac{1}{2} \int_V d^3r \int_V d^3r' \, \frac{\rho(\mathbf{r}, t) \rho(\mathbf{r}', t)}{4\pi\epsilon_0 |\mathbf{r} - \mathbf{r}'|}. \tag{7.7}$$

Working in the Coulomb gauge, the charge density is related to the electrostatic potential $\phi(\mathbf{r}, t)$ in the usual way:

$$-\nabla^2 \phi(\mathbf{r}, t) = \frac{\rho(\mathbf{r}, t)}{\epsilon_0}, \tag{7.8}$$

or equivalently

$$\phi(\mathbf{r}, t) = \int_V \frac{\rho(\mathbf{r}', t)}{4\pi\epsilon_0 |\mathbf{r} - \mathbf{r}'|} \, d^3r'. \tag{7.9}$$

This allows the potential energy to be rewritten in the alternative form:

$$V = \frac{1}{2} \int_V \epsilon_0 \left[\nabla \phi(\mathbf{r}, t)\right]^2 d^3r. \tag{7.10}$$

The Hamiltonian is the sum of the kinetic and potential energies:

$$H = \frac{1}{2} \int_V \left(\frac{J^2(\mathbf{r}, t)}{\epsilon_0 \omega_p^2(\mathbf{r})} + \epsilon_0 \left[\nabla \phi(\mathbf{r}, t)\right]^2 \right) d^3r. \tag{7.11}$$

It is possible to relate the current density to the scalar potential by considering a microscopic model of the system. Each electron is subject to the electrostatic force, so that

$$m_e \dot{\mathbf{v}}(\mathbf{r},t) = e\nabla\phi(\mathbf{r},t). \tag{7.12}$$

Combining this with equation (7.4) then leads to the required relation:

$$\dot{\mathbf{J}}(\mathbf{r},t) = -\epsilon_0\omega_p^2(\mathbf{r})\nabla\phi(\mathbf{r},t) \tag{7.13}$$

The Hamiltonian contains $\mathbf{J}$, not $\dot{\mathbf{J}}$; it is therefore convenient to introduce a second scalar function $f(\mathbf{r},t)$, where

$$\dot{f}(\mathbf{r},t) = \phi(\mathbf{r},t). \tag{7.14}$$

From equation (7.13), it follows that

$$\mathbf{J}(\mathbf{r},t) = -\epsilon_0\omega_p^2(\mathbf{r})\nabla f(\mathbf{r},t), \tag{7.15}$$

subject to reasonable restrictions on $\mathbf{J}$.

In terms of f, the Hamiltonian becomes

$$H = \frac{\epsilon_0}{2}\int_V \left(\omega_p^2(\mathbf{r})\big[\nabla f(\mathbf{r},t)\big]^2 + \big[\nabla\dot{f}(\mathbf{r},t)\big]^2\right) d^3r, \tag{7.16}$$

while the Lagrangian is

$$L = \frac{\epsilon_0}{2}\int_V \left(-\omega_p^2(\mathbf{r})\big[\nabla f(\mathbf{r},t)\big]^2 + \big[\nabla\dot{f}(\mathbf{r},t)\big]^2\right) d^3r. \tag{7.17}$$

Note that the signs in the Lagrangian are the opposite of what would naively be expected: this is because the kinetic energy term contains f, while the potential energy term contains $\dot{f}$.

To check these results, the equations of motion may be determined. The field variable conjugate to f is

$$\frac{\delta L[f,\dot{f}]}{\delta\dot{f}(\mathbf{r},t)} = -\epsilon_0\nabla^2\dot{f}(\mathbf{r},t) = \rho(\mathbf{r},t). \tag{7.18}$$

In terms of ρ and f, the Hamiltonian is

$$H = \frac{1}{2}\int_V \left(\epsilon_0\omega_p^2(\mathbf{r})\big[\nabla f(\mathbf{r},t)\big]^2 + \int_V \frac{\rho(\mathbf{r},t)\rho(\mathbf{r}',t)}{4\pi\epsilon_0|\mathbf{r}-\mathbf{r}'|}\, d^3r'\right)\, d^3r. \tag{7.19}$$

The first equation of motion is obtained by setting

$$\begin{aligned}\dot{f}(\mathbf{r},t) &= \frac{\delta H[f,\rho]}{\delta\rho(\mathbf{r},t)} \\ &= \int_V \frac{\rho(\mathbf{r}',t)}{4\pi\epsilon_0|\mathbf{r}-\mathbf{r}'|}\, d^3r'.\end{aligned}$$

The right-hand side is equal to $\phi(\mathbf{r},t)$ (see equation (7.9)); comparison with equation (7.14) shows that this relation is correct. The second equation of motion is given by

$$\begin{aligned}-\dot{\rho}(\mathbf{r},t) &= \frac{\delta H[f,\rho]}{\delta f(\mathbf{r},t)} \\ &= -\epsilon_0\nabla\cdot\left(\omega_p^2(\mathbf{r})\nabla f(\mathbf{r},t)\right).\end{aligned} \tag{7.20}$$

Equation (7.15) then allows f to be replaced, giving

$$-\dot{\rho}(\mathbf{r},t) = \nabla\cdot\mathbf{J}, \tag{7.21}$$

which is the equation of continuity. The Lagrangian and Hamiltonian defined in equations (7.17) and (7.16) therefore lead to the correct equations of motion.

Introducing the Fourier transforms

$$\begin{aligned}f_{\mathbf{k}} &= \frac{1}{\sqrt{V}}\int_V f(\mathbf{r})e^{-i\mathbf{k}\cdot\mathbf{r}}\, d^3r \\ \rho_{\mathbf{k}} &= \frac{1}{\sqrt{V}}\int_V \rho(\mathbf{r})e^{i\mathbf{k}\cdot\mathbf{r}}\, d^3r\end{aligned} \tag{7.22}$$

and inverse transforms

$$\begin{aligned}f(\mathbf{r}) &= \frac{1}{\sqrt{V}}\sum_{\mathbf{k}} f_{\mathbf{k}}e^{i\mathbf{k}\cdot\mathbf{r}} \\ \rho(\mathbf{r}) &= \frac{1}{\sqrt{V}}\sum_{\mathbf{k}} \rho_{\mathbf{k}}e^{-i\mathbf{k}\cdot\mathbf{r}},\end{aligned} \tag{7.23}$$

the Hamiltonian may be written as

$$\begin{aligned}H &= \frac{1}{2}\sum_{\mathbf{k},\mathbf{k}'}\left(\epsilon_0 f_{\mathbf{k}}f_{-\mathbf{k}'}\mathbf{k}\cdot\mathbf{k}'\frac{1}{V}\int_V \omega_p^2(\mathbf{r})e^{i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{r}}\, d^3r\right. \\ &\qquad \left.+\rho_{-\mathbf{k}}\rho_{\mathbf{k}'}\frac{1}{V}\int_V\int_V \frac{e^{i(\mathbf{k}\cdot\mathbf{r}-\mathbf{k}'\cdot\mathbf{r}')}}{4\pi\epsilon_0|\mathbf{r}-\mathbf{r}'|}\, d^3r\, d^3r'\right) \\ &= \frac{1}{2}\sum_{\mathbf{k},\mathbf{k}'}\left(\epsilon_0 f_{\mathbf{k}}kM_{\mathbf{k}\mathbf{k}'}k'f_{-\mathbf{k}'} + \rho_{-\mathbf{k}}\frac{\delta_{\mathbf{k}\mathbf{k}'}}{\epsilon_0 k^2}\rho_{\mathbf{k}'}\right)\end{aligned} \tag{7.24}$$

where

$$M_{\mathbf{k}\mathbf{k}'} = \frac{\hat{\mathbf{k}} \cdot \hat{\mathbf{k}}'}{V} \int_V \omega_p^2(\mathbf{r}) e^{i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{r})} \, d^3r. \tag{7.25}$$

The reason for choosing opposite signs in the transforms of f and ρ is to ensure that $f_{\mathbf{k}}$ and $\rho_{\mathbf{k}}$ are conjugate variables. This conjugate variable requirement implies that

$$\begin{aligned} \rho_{\mathbf{k}} &= \frac{\partial L(\{f_{\mathbf{k}}\}, \{\dot{f}_{\mathbf{k}}\})}{\partial \dot{f}_{\mathbf{k}}} \\ &= \int_V \frac{\delta L[f, \dot{f}]}{\delta \dot{f}(\mathbf{r}, t)} \frac{\partial \dot{f}(\mathbf{r}, t)}{\partial \dot{f}_{\mathbf{k}}} \, d^3r \\ &= \frac{1}{\sqrt{V}} \int_V \rho(\mathbf{r}, t) e^{i\mathbf{k}\cdot\mathbf{r}} \, d^3r. \end{aligned} \tag{7.26}$$

The $\mathbf{k} = 0$ terms are excluded; this will be the case for all sums from now on. The justification for this is that the average charge density fluctuation is constrained to be zero, and the average potential (which corresponds to $\dot{f}_{\mathbf{k}=\mathbf{0}}$) may be chosen to be zero.

7.2 Quantisation and diagonalisation of the Hamiltonian

In equation (7.24), the Hamiltonian is expressed in terms of pairs of discrete conjugate variables. Quantisation proceeds by letting $f_{\mathbf{k}}$ and $\rho_{\mathbf{k}}$ become operators, subject to the commutation relation

$$\left[\rho_{\mathbf{k}}, f_{\mathbf{k}'}\right] = -i\hbar\delta_{\mathbf{k}\mathbf{k}'}. \tag{7.27}$$

These are not Hermitian operators; however, they are the Fourier components of operators corresponding to real, observable fields, and must obey the symmetry rules

$$\rho_{\mathbf{k}}^\dagger = \rho_{-\mathbf{k}}, \; f_{\mathbf{k}}^\dagger = f_{-\mathbf{k}}. \tag{7.28}$$

The quantum-mechanical Hamiltonian operator is therefore

$$H = \frac{1}{2} \sum_{\mathbf{k},\mathbf{k}'} \left(\rho_{\mathbf{k}}^\dagger \frac{\delta_{\mathbf{k}\mathbf{k}'}}{\epsilon_0 k^2} \rho_{\mathbf{k}'} + \epsilon_0 f_{\mathbf{k}} k M_{\mathbf{k}\mathbf{k}'} k' f_{\mathbf{k}'}^\dagger \right). \tag{7.29}$$

which describes a set of coupled harmonic oscillators; the coupling between the oscillations of different **k**-vector is governed by the Hermitian matrix M.

To uncouple the oscillators, M must be diagonalised:

$$M_{\mathbf{k}\mathbf{k}'} = \sum_i U_{\mathbf{k}i} \lambda_i U^*_{\mathbf{k}'i}. \tag{7.30}$$

Here, U is the unitary matrix of eigenvectors and λ represents the eigenvalues of M. From equation (7.25), it is clear that

$$M_{\mathbf{k}\mathbf{k}'} = M^*_{\mathbf{k}'\mathbf{k}} = M^*_{(-\mathbf{k})(-\mathbf{k}')} = M_{(-\mathbf{k}')(-\mathbf{k})}. \tag{7.31}$$

These symmetry relations have implications for U which will prove useful. The eigenvalue equation is

$$\sum_{\mathbf{k}'} M_{\mathbf{k}\mathbf{k}'} U_{\mathbf{k}'i} = \lambda_i U_{\mathbf{k}i}. \tag{7.32}$$

Replacing $\mathbf{k}$ with $-\mathbf{k}$, $\mathbf{k}'$ with $-\mathbf{k}'$ and taking the complex conjugate gives

$$\sum_{\mathbf{k}'} M^*_{(-\mathbf{k})(-\mathbf{k}')} U^*_{(-\mathbf{k}')i} = \lambda^*_i U^*_{(-\mathbf{k})i}. \tag{7.33}$$

Equation (7.31) then implies that

$$\sum_{\mathbf{k}'} M_{\mathbf{k}\mathbf{k}'} U^*_{(-\mathbf{k}')i} = \lambda_i U^*_{(-\mathbf{k})i}. \tag{7.34}$$

Comparison with equation (7.32) shows that $U^*_{(-\mathbf{k})i}$ is also an eigenvector of M, with the same eigenvalue λ_i. If λ_i is non-degenerate, the two eigenvectors are the same, and

$$U^*_{(-\mathbf{k})i} = U_{\mathbf{k}i} \tag{7.35}$$

to within a phase factor, which is chosen to be zero. If λ_i is degenerate, it is possible to choose linear combinations of the degenerate eigenvectors to ensure that this condition is met.

Using equation (7.25), and the fact that

$$\sum_i U_{\mathbf{k}i} U^*_{\mathbf{k}'i} = \sum_i U_{\mathbf{k}i} U^\dagger_{i\mathbf{k}'} = \delta_{\mathbf{k}\mathbf{k}'}, \tag{7.36}$$

the Hamiltonian may be recast as

$$H = \frac{1}{2}\sum_i \left(\frac{1}{\epsilon_0}\sum_{\mathbf{k}} \frac{U_{\mathbf{k}i}\rho^\dagger_{\mathbf{k}}}{k} \sum_{\mathbf{k}'} \frac{U^*_{\mathbf{k}'i}\rho_{\mathbf{k}'}}{k'} \right. \tag{7.37}$$

$$\left. + \epsilon_0\lambda_i \sum_{\mathbf{k}} kU_{\mathbf{k}i} f_{\mathbf{k}} \sum_{\mathbf{k}'} k' U^*_{\mathbf{k}'i} f^\dagger_{\mathbf{k}'} \right)$$

$$= \frac{1}{2}\sum_i \left(\frac{1}{\epsilon_0} p_i^\dagger p_i + \epsilon_0\lambda_i q_i q_i^\dagger \right). \tag{7.38}$$

The new operators correspond to the normal coordinates:

$$p_i^\dagger = \sum_{\mathbf{k}} \frac{U_{\mathbf{k}i}\rho^\dagger_{\mathbf{k}}}{k} \qquad \rho^\dagger_{\mathbf{k}} = \sum_i U^*_{\mathbf{k}i} k p_i^\dagger \tag{7.39a}$$

$$q_i^\dagger = \sum_{\mathbf{k}} U^*_{\mathbf{k}i} k f^\dagger_{\mathbf{k}} \qquad f^\dagger_{\mathbf{k}} = \sum_i \frac{U_{\mathbf{k}i} q_i^\dagger}{k}. \tag{7.39b}$$

One can show that $p_i^\dagger$ is an Hermitian operator; beginning with the complex conjugate of equation (7.39a), and using equations (7.28) and (7.35), it follows that

$$\begin{aligned} p_i &= \sum_{\mathbf{k}} \frac{U^*_{\mathbf{k}i}\rho_{\mathbf{k}}}{k} \\ &= \sum_{\mathbf{k}} \frac{U_{(-\mathbf{k})i}\rho_{\mathbf{k}}}{k} \\ &= \sum_{\mathbf{k}'} \frac{U_{\mathbf{k}'i}\rho^\dagger_{\mathbf{k}'}}{k'} \\ &= p_i^\dagger. \end{aligned} \tag{7.40}$$

A similar proof demonstrates that q_i is also Hermitian. The commutation relation for these operators is

$$\left[p_i, q_j\right] = -i\hbar\delta_{ij}. \tag{7.41}$$

The Hamiltonian is now

$$H = \frac{1}{2}\sum_i \left(\frac{1}{\epsilon_0} p_i^2 + \lambda_i\epsilon_0 q_i^2 \right). \tag{7.42}$$

In terms of the coordinates $\{p_i\}$ and $\{q_i\}$, the diagonalised Hamiltonian is completely separable. The eigenfunctions are products of the form

$$\Psi(\{p_i\}) = \prod_i \psi_i(p_i) \tag{7.43}$$

where

$$\frac{1}{2}\left(\frac{1}{\epsilon_0}p_i^2 + \lambda_i\epsilon_0 q_i^2\right)\psi_i = \mathcal{E}_i\psi_i. \tag{7.44}$$

Diagonalisation is now trivial. The usual operators

$$a_i = \sqrt{\frac{1}{2\epsilon_0\hbar\omega_i}}p_i - i\sqrt{\frac{\epsilon_0\omega_i}{2\hbar}}q_i \tag{7.45a}$$

$$a_i^\dagger = \sqrt{\frac{1}{2\epsilon_0\hbar\omega_i}}p_i + i\sqrt{\frac{\epsilon_0\omega_i}{2\hbar}}q_i \tag{7.45b}$$

are introduced, which obey the commutation relation

$$\left[a_i, a_j^\dagger\right] = \delta_{ij}. \tag{7.46}$$

Defining $\omega_i = \sqrt{\lambda_i}$, the Hamiltonian becomes

$$\begin{aligned} H &= \sum_i \hbar\omega_i\left(a_i^\dagger a_i + \frac{i}{2\hbar}\left[p_i, q_i\right]\right) \\ &= \sum_i \hbar\omega_{\mathbf{k}}\left(a_i^\dagger a_i + \frac{1}{2}\right) \\ &= \sum_i \hbar\omega_i a_i^\dagger a_i + \text{zero-point energy}. \end{aligned} \tag{7.47}$$

The zero-point energy, though infinite, is constant; it therefore does not affect the form of the wave function, and will be omitted from now on. It appears as the consequence of having an infinite number of oscillators, each of which contributes its own zero-point energy. This may be clearly seen in equation (7.47).

7.3 The ground-state wave function

The equation which determines the form of ψ_i for the ground-state wave function is

$$a_i\psi_i = 0. \tag{7.48}$$

To express a_i in terms of q_i alone, the representation of p_i in the q_i basis is required. The commutation relation for these operators, equation (7.41), implies that

$$q_i(p_i) = i\hbar\frac{\partial}{\partial p_i}. \tag{7.49}$$

Substituting equations (7.49) and (7.45a) into equation (7.48) and solving the resulting differential equation gives

$$\psi_i(p_i) = \exp\left(-\frac{p_i^2}{2\epsilon_0\hbar\omega_i}\right). \tag{7.50}$$

The full ground-state solution is therefore

$$\Psi(\{p_i\}) = \exp\left(-\frac{1}{2\epsilon_0\hbar}\sum_i \frac{1}{\omega_i}p_i^2\right). \tag{7.51}$$

Alternatively, equation (7.39) may be used to obtain the solution in terms of the Fourier components of the plasmon charge density:

$$\Psi(\{\rho_{\mathbf{k}}\}) = \exp\left(-\frac{1}{2\epsilon_0\hbar}\sum_{\mathbf{k},\mathbf{k}'} \rho_{\mathbf{k}}^* \frac{\left(M^{-1/2}\right)_{\mathbf{k}\mathbf{k}'}}{kk'}\rho_{\mathbf{k}'}\right). \tag{7.52}$$

Equation (7.2) gives the relationship between the plasmon charge density and the electron density in real space. In Fourier space, this becomes

$$\rho_{\mathbf{k}} = -e(n_{\mathbf{k}} - \bar{n}_{\mathbf{k}}) \tag{7.53}$$

where $n_{\mathbf{k}}$ and $\bar{n}_{\mathbf{k}}$ are defined analogously[1] to $\rho_{\mathbf{k}}$. Substituting for $\rho_{\mathbf{k}}$ allows the ground-state wave function to be written in terms of the electron density:

$$\Psi(\{n_{\mathbf{k}}\}) = \exp\left(-\frac{e^2}{2\epsilon_0\hbar}\sum_{\mathbf{k},\mathbf{k}'}\left(n_{\mathbf{k}}^* - \bar{n}_{\mathbf{k}}^*\right)\frac{\left(M^{-1/2}\right)_{\mathbf{k}\mathbf{k}'}}{kk'}\left(n_{\mathbf{k}'} - \bar{n}_{\mathbf{k}}\right)\right). \tag{7.54}$$

The electron density operator is

$$n(\mathbf{r}) = \sum_i \delta(\mathbf{r} - \mathbf{r_i}), \tag{7.55}$$

or, in $\mathbf{k}$-space,

$$n_{\mathbf{k}} = \frac{1}{\sqrt{V}}\sum_i e^{i\mathbf{k}\cdot\mathbf{r_i}}. \tag{7.56}$$

The final step is to write the ground-state wave function in terms of the electron coordinates:

$$\Psi(\{\mathbf{r}_i\}) = \exp\left(-\frac{1}{2}\sum_{i,j} u(\mathbf{r}_i, \mathbf{r}_j) + \sum_i \chi(\mathbf{r}_i)\right) \tag{7.57}$$

[1] See equation (7.22).

where

$$\chi(\mathbf{r}_i) = \frac{e^2}{\epsilon_0 \hbar \sqrt{V}} \sum_{\mathbf{k},\mathbf{k}'} e^{-i\mathbf{k}\cdot\mathbf{r}_i} \frac{\left(M^{-1/2}\right)_{\mathbf{k}\mathbf{k}'}}{kk'} \bar{n}_{\mathbf{k}'} \tag{7.58}$$

and

$$u(\mathbf{r}_i, \mathbf{r}_j) = \frac{e^2}{\epsilon_0 V \hbar} \sum_{\mathbf{k},\mathbf{k}'} e^{-i\mathbf{k}\cdot\mathbf{r}_i} \frac{\left(M^{-1/2}\right)_{\mathbf{k}\mathbf{k}'}}{kk'} e^{i\mathbf{k}'\cdot\mathbf{r}_j}. \tag{7.59}$$

In this notation, the double sum over i and j is unrestricted and includes the case $i = j$; this means that a part of the one-body term is incorporated in the sum over u.

It is useful to obtain the ground-state wave function in terms of f, rather than ρ. Instead of solving equation (7.48) for p_i, it may be solved for q_i, giving

$$\psi_i(q_i) = \exp\left(-\frac{\epsilon_0 \omega_i q_i^2}{2\hbar}\right) \tag{7.60}$$

The full ground-state wave function is then

$$\Psi(\{f_{\mathbf{k}}\}) = \exp\left(-\frac{\epsilon_0}{2\hbar} \sum_{\mathbf{k},\mathbf{k}'} f_{\mathbf{k}} k \left(M^{1/2}\right)_{\mathbf{k}\mathbf{k}'} k' f^*_{\mathbf{k}'}\right). \tag{7.61}$$

7.4 Normal modes

An alternative approach to the analysis of the previous section is to diagonalise the Hamiltonian before quantising; this is perhaps neater, and emphasises the rôle played by the normal modes of the classical system, which will be useful later. To diagonalise the classical Hamiltonian, it is necessary to establish the equation of motion satisfied by the field f; this is obtained by combining equations (7.18) and (7.20):

$$-\nabla^2 \ddot{f} = \nabla \cdot \left(\omega_p^2 \nabla f\right). \tag{7.62}$$

The defining characteristic of a normal mode is harmonic time dependence; the normal modes for f therefore satisfy the equation

$$\omega_i^2 \nabla^2 f_i = \nabla \cdot \left(\omega_p^2 \nabla f_i\right). \tag{7.63}$$

Multiplying by f_j and integrating gives

$$\omega_i^2 \int_V f_j \nabla^2 f_i \; d^3r = \int_V f_j \nabla \cdot \left(\omega_p^2 \nabla f_i\right) \; d^3r, \tag{7.64}$$

or, on rearranging,

$$\omega_i^2 \int_V \nabla f_j \cdot \nabla f_i \, d^3r = \int_V \omega_p^2 \nabla f_j \cdot \nabla f_i \, d^3r. \tag{7.65}$$

Swapping the indices gives the corresponding expression

$$\omega_j^2 \int_V \nabla f_i \cdot \nabla f_j \, d^3r = \int_V \omega_p^2 \nabla f_i \cdot \nabla f_j \, d^3r. \tag{7.66}$$

Taking the difference of the two previous equations shows that

$$\left(\omega_i^2 - \omega_j^2\right) \int_V \nabla f_i \cdot \nabla f_j \, d^3r = 0. \tag{7.67}$$

In the non-degenerate case, this implies that ∇f_i and ∇f_j are orthogonal, in the sense that

$$\int_V \nabla f_i \cdot \nabla f_j \, d^3r = 0 \quad (\omega_i \neq \omega_j). \tag{7.68}$$

When the modes are degenerate, it is always possible to construct combinations of the (linearly independent) functions which are orthogonal. If, additionally, the modes are taken to be normalised, the general result

$$\int_V \nabla f_i \cdot \nabla f_j \, d^3r = \delta_{ij} \tag{7.69}$$

is obtained. A secondary consequence is that

$$\int_V \omega_p^2 \nabla f_i \cdot \nabla f_j \, d^3r = \omega_i^2 \delta_{ij}. \tag{7.70}$$

Any solution of the original problem (equation (7.62)) may therefore be expanded in terms of the normal modes as follows:

$$\nabla f = \sum_i \alpha_i \nabla f_i. \tag{7.71}$$

The normal mode amplitudes are determined by the inverse relation

$$\alpha_i = \int_V \nabla f_i \cdot \nabla f \, d^3r. \tag{7.72}$$

This converts the classical Hamiltonian of equation (7.24) into the form

$$H = \frac{\epsilon_0}{2} \sum_i \left(\dot{\alpha}_i^2 + \omega_i^2 \alpha_i^2\right). \tag{7.73}$$

The Lagrangian becomes

$$L = \frac{\epsilon_0}{2} \sum_i \left(\dot{\alpha}_i^2 - \omega_i^2 \alpha_i^2 \right), \tag{7.74}$$

leading to the set of conjugate variables

$$\beta_i = \frac{\partial L}{\partial \dot{\alpha}_i} = \epsilon_0 \dot{\alpha}_i. \tag{7.75}$$

In terms of the new variables, the Hamiltonian is

$$H = \frac{1}{2} \sum_i \left(\frac{1}{\epsilon_0} \beta_i^2 + \epsilon_0 \omega_i^2 \alpha_i^2 \right). \tag{7.76}$$

As before, the Hamiltonian may be quantised; this time, the commutation relation is between α_i and β_j:

$$\left[\beta_j, \alpha_i \right] = -i\hbar \delta_{ij}. \tag{7.77}$$

If the normal modes are chosen to be real, the operators $\{\alpha_i, \beta_i\}$ are Hermitian. The Hamiltonian operator is therefore

$$\begin{aligned} \hat{H} &= \sum_i \hbar\omega_i \left[\left(\sqrt{\frac{1}{2\epsilon_0 \hbar \omega_i}} \beta_i + i \sqrt{\frac{\epsilon_0 \omega_i}{2\hbar}} \alpha_i \right) \left(\sqrt{\frac{1}{2\epsilon_0 \hbar \omega_i}} \beta_i - i \sqrt{\frac{\epsilon_0 \omega_i}{2\hbar}} \alpha_i \right) - \frac{i}{2\hbar} \left[\alpha_i, \beta_i \right] \right] \\ &= \sum_i \hbar\omega_i \left(a_i^\dagger a_i + \frac{1}{2} \right) \\ &= \sum_i \hbar\omega_i a_i^\dagger a_i + \text{zero-point energy} \end{aligned} \tag{7.78}$$

where the operators $a_i^\dagger$ and a_i are defined in equation (7.45); they obey the commutation relation (7.46).

Comparison with the results of section 7.3 gives the ground-state wave function as

$$\Psi(\{\beta_i\}) = \exp\left(-\frac{1}{2\epsilon_0 \hbar} \sum_i \frac{1}{\omega_i} \beta_i^2 \right), \tag{7.79}$$

or, using equations (7.75) and (7.72),

$$\Psi[\dot{f}] = \exp\left(-\frac{\epsilon_0}{2\hbar}\sum_i \frac{1}{\omega_i}\left(\int_V \nabla f_i \cdot \nabla \dot{f}\; d^3r\right)^2\right) \tag{7.80a}$$

$$\Psi[\rho] = \exp\left(-\frac{1}{2\epsilon_0\hbar}\sum_i \frac{1}{\omega_i}\left(\int_V f_i\rho\; d^3r\right)^2\right) \tag{7.80b}$$

$$\Psi(\{\mathbf{r}_i\}) = \exp\left(-\frac{e^2}{2\epsilon_0\hbar}\sum_i \frac{1}{\omega_i}\left(-\sum_m f_i(\mathbf{r}_m) + \int_V f_i\bar{n}\; d^3r\right)^2\right). \tag{7.80c}$$

Several previous results have been used in equation (7.80), including equations (7.2), (7.18) and (7.55).

The wave function has now been reduced to the standard form given in equation (7.57), with

$$\chi(\mathbf{r}) = \frac{e^2}{\epsilon_0\hbar}\sum_i \frac{1}{\omega_i} f_i(\mathbf{r}) \int_V f_i(\mathbf{r}')\bar{n}(\mathbf{r}')\; d^3r' \tag{7.81}$$

and

$$u(\mathbf{r},\mathbf{r}') = \frac{e^2}{\epsilon_0\hbar}\sum_i \frac{1}{\omega_i} f_i(\mathbf{r}) f_i(\mathbf{r}'). \tag{7.82}$$

Note that ϕ and f satisfy the same equation of motion. The boundary conditions used in determining the normal modes for these two quantities are also identical; in fact, the normal modes differ only by a numerical factor, and not in functional form. The numerical factor may be eliminated by suitable normalisation, with the result that f_i can be replaced by ϕ_i in equations (7.81) and (7.82).

Chapter 8

Applying the plasmon normal mode theory to slab systems

In this chapter, the application of the techniques developed in chapter 7 to jellium slab systems will be discussed. The first task is to establish the form of the plasmon normal modes; this will be dealt with in section 8.1. In section 8.2, the equations of chapter 7 will be applied to these modes, giving the plasmon contribution to the ground-state wave functions. Because studying plasmons only gives information about long-wavelength correlations, it is necessary to combine the plasmonic form with knowledge of the correlation between electrons at short distances in order to generate good trial wave functions; this is discussed in section 8.3, along with the results of using this method.

8.1 Classical plasmons in slabs

The system to be considered is illustrated in figure 8.1. It consists of a metal slab of width s, characterised by a simple non-absorptive frequency-dependent conductivity, surrounded by vacuum.

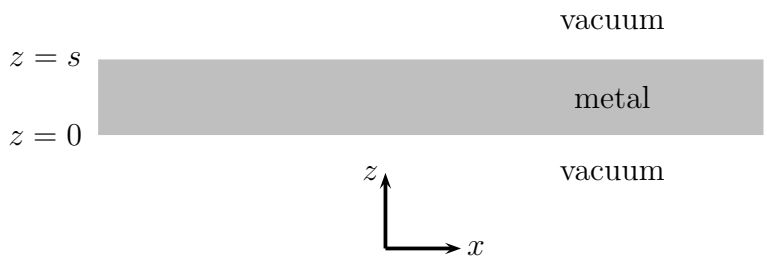


Figure 8.1: The metal slab.

The appropriate versions of Maxwell's equations are

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0} \qquad \nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \tag{8.1}$$

$$\nabla \cdot \mathbf{B} = 0 \qquad \nabla \times \mathbf{B} = \mu_0 \mathbf{J} + \frac{1}{c^2}\frac{\partial \mathbf{E}}{\partial t}. \tag{8.2}$$

8.1.1 The metal conductivity

In the jellium model of a bulk metal, the (smoothed-out) charge density is zero in the steady state: the electron and background charges cancel each other out completely. In equation (8.1), ρ refers to the net charge.

Any departure of ρ from zero will be caused by motion of electrons, because the background charge is not free to move; the same is true for any current density. The current density at the point $\mathbf{r}$ is therefore given by

$$\mathbf{J}(\mathbf{r}) = -n(\mathbf{r}, t)\, e\, \mathbf{v}(\mathbf{r}, t) \tag{8.3}$$

where $n(\mathbf{r}, t)$ is the electron density and $\mathbf{v}(\mathbf{r}, t)$ is the velocity of the electrons at $\mathbf{r}$. For small amplitude oscillations, it is reasonable to make the linearising approximation[1] $n(\mathbf{r}, t) = n_0(\mathbf{r})$, where n_0 is the steady-state electron density.

[1]The formal approach is to treat any time-dependent quantity as a small perturbation. In fact, $n(\mathbf{r}, t) = n_0(\mathbf{r}) - \frac{\rho(\mathbf{r},t)}{e}$. In the expression for the current density, the second-order term $\rho(\mathbf{r}, t)\mathbf{v}(\mathbf{r}, t)$ is then discarded.

If there are no collisions then the classical equation of motion for each electron is

$$m_e \dot{\mathbf{v}} = -e\mathbf{E} \tag{8.4}$$

which, under the assumption of harmonic time dependence,[2] may be manipulated to obtain the conductivity:

$$i\omega m_e \mathbf{v} = -e\mathbf{E} \tag{8.5}$$

$$-n_0 e\mathbf{v} = \frac{n_0 e^2}{i\omega m_e}\mathbf{E} \tag{8.6}$$

$$\mathbf{J} = \sigma\mathbf{E} \tag{8.7}$$

where

$$\sigma = \frac{ne^2}{im_e\omega} = \frac{\epsilon_0\omega_p^2}{i\omega}. \tag{8.8}$$

The quantity

$$\omega_p = \sqrt{\frac{ne^2}{m_e\epsilon_0}} \tag{8.9}$$

is the plasma frequency.

8.1.2 Solving Maxwell's equations

Applying the results of the preceding section to equations (8.1) and (8.2) gives

$$\nabla\cdot\mathbf{E} = \frac{\rho}{\epsilon_0} \qquad \nabla\times\mathbf{E} = -i\omega\mathbf{B} \tag{8.10}$$

$$\nabla\cdot\mathbf{B} = 0 \qquad \nabla\times\mathbf{B} = \frac{i}{\omega c^2}(\omega^2 - \omega_p^2)\mathbf{E}. \tag{8.11}$$

Conveniently, these apply both to the metal and the vacuum, with the understanding that $\omega_p = 0$ in the vacuum.

Combining the 'curl' equations gives

$$\nabla\times\nabla\times\mathbf{E} = \frac{1}{c^2}(\omega^2 - \omega_p^2)\mathbf{E}. \tag{8.12}$$

[2] In the search for normal modes, all time-varying quantities are taken to have the form $e^{i\omega t}$.

To look for solutions propagating in a direction parallel to the slab, we set

$$\mathbf{E}(\mathbf{r},t) = \mathbf{E}(z)e^{i(\omega t - kx)}, \tag{8.13}$$

which, on substitution into equation (8.12) leads to the set of equations

$$-ik\frac{dE_z}{dz} - \frac{d^2E_x}{dz^2} = \frac{\omega^2-\omega_p^2}{c^2}E_x \tag{8.14a}$$

$$k^2E_y - \frac{d^2E_y}{dz^2} = \frac{\omega^2-\omega_p^2}{c^2}E_y \tag{8.14b}$$

$$-ik\frac{dE_x}{dz} + k^2E_z = \frac{\omega^2-\omega_p^2}{c^2}E_z \tag{8.14c}$$

The y equation is decoupled from the others, and has solutions

$$E_y = E_0e^{\pm K_y z} \tag{8.15}$$

where

$$k^2 - K_y^2 = \frac{\omega^2-\omega_p^2}{c^2}. \tag{8.16}$$

If $k = 0$, the solution is symmetrical in x and y; equation (8.14a) corresponds to equation (8.14b), with E_x replacing E_y. Equation (8.14c) shows that there is no field in the z-direction in this case, unless $\omega = \omega_p$.

When $k \neq 0$ and $\omega \neq \omega_p$, equation (8.14c) may be rearranged to give

$$E_z = \frac{-ikc^2}{\omega^2-\omega_p^2-k^2c^2}\frac{dE_x}{dz}, \tag{8.17}$$

which, on substitution into equation (8.14a), leads to

$$\frac{d^2E_x}{dz^2} = \left(k^2 - \frac{\omega^2}{c^2} + \frac{\omega_p^2}{c^2}\right)E_x. \tag{8.18}$$

The fields may then be calculated:

$$E_x = E_0e^{\pm Kz} \tag{8.19}$$

$$E_z = \pm E_0\frac{ik}{K}e^{\pm Kz} \tag{8.20}$$

where

$$K = \sqrt{k^2 - \frac{\omega^2}{c^2} + \frac{\omega_p^2}{c^2}} \quad \text{(with } \omega_p = 0 \text{ in vacuum)}. \tag{8.21}$$

These solutions are transverse, because $\nabla \cdot \mathbf{E} = 0$ (except at the boundaries).

Equations (8.14a), (8.14b) and (8.14c) make it clear that $\omega = \omega_p$ constitutes a special case. The oscillations for which this condition is satisfied are the bulk plasmons; note that there are no bulk plasmon solutions in the vacuum.[3]

Equation (8.11) shows that there is no **B**-field at this frequency; then, from equation (8.10), $\nabla \times \mathbf{E} = \mathbf{0}$, which implies that the **E**-field is longitudinal. The only restriction on the functional form of the field is that

$$ik\frac{dE_x}{dz} + k^2 E_z = 0. \tag{8.22}$$

The remaining special case is when $\omega^2 = \omega_p^2 + k^2c^2$ (or in vacuum, $\omega = kc$). Then

$$E_x = E_0 \tag{8.23a}$$

$$E_z = ikzE_0 + E_1 \tag{8.23b}$$

where E_0 and E_1 are constants.

8.1.3 Boundary conditions

At any plane interface, Maxwell's equations may be used to demonstrate that the component of the **E**-field parallel to the interface is continuous, while there is a discontinuity in the component perpendicular to the interface equal to the surface charge density.

The same free-electron model which was used to estimate the conductivity is useful here. For the interface illustrated in figure 8.2, the surface charge density at the interface is $-n_0eq$, where q represents the displacement (in the direction normal to the interface) of each electron at the boundary from its equilibrium position:

$$E_z^{(v)} - E_z^{(m)} = -\frac{n_0 eq}{\epsilon_0}. \tag{8.24}$$

[3]The equivalent of a bulk plasmon in vacuum is the zero-frequency solution: the electric field generated by a charge density which is constant in time.

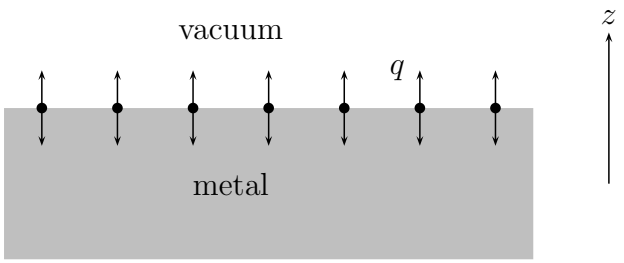


Figure 8.2: The surface charge density.

To relate this to the previous calculation, note that $v_z = \dot{q}$. It then follows that

$$E_z^{(v)} - E_z^{(m)} = -\frac{n_0 e v_z}{i\epsilon_0 \omega} \tag{8.25}$$

$$= \frac{\sigma}{i\epsilon_0 \omega} E_z^{(m)} \tag{8.26}$$

$$= -\frac{\omega_p^2}{\omega^2} E_z^{(m)}. \tag{8.27}$$

To summarise, the boundary conditions at an interface between metal and vacuum are:

$$E_x^{(m)} = E_x^{(v)} \tag{8.28a}$$

$$E_y^{(m)} = E_y^{(v)} \tag{8.28b}$$

$$\left(1 - \frac{\omega_p^2}{\omega^2}\right) E_z^{(m)} = E_z^{(v)}. \tag{8.28c}$$

Note that from now on, ω_p is being used exclusively to refer to the plasma frequency of the metal.

In the limit $z \to \pm\infty$, all field components are required to tend to zero. This excludes both the non-physical (constant or increasing) and the radiative (oscillatory) solutions.

8.1.4 Surface plasmons

In order to maintain the relationship specified by equation (8.28) at the interfaces, k and ω must take the same values in the metal and the vacuum. One consequence

of the boundary conditions at infinity is that K must be real if any field is to exist in the vacuum; equation (8.21) then implies that

$$k > \omega/c. \tag{8.29}$$

This condition also ensures that K is real in the metal.[4]

The surface plasmon solutions ($\omega \neq \omega_p$) for E_x and E_z are therefore

$$E_x = \begin{cases} E_0 e^{-K_v z} & \text{when } z > s \\ E_1 e^{-K_m z} + E_2 e^{K_m z} & \text{when } 0 < z < s \\ E_3 e^{K_v z} & \text{otherwise} \end{cases} \tag{8.30}$$

$$E_z = \begin{cases} -\frac{ik}{K_v} E_0 e^{-K_v z} & \text{when } z > s \\ \frac{ik}{K_m}\left(-E_1 e^{-K_m z} + E_2 e^{K_m z}\right) & \text{when } 0 < z < s \\ \frac{ik}{K_v} E_3 e^{K_v z} & \text{otherwise} \end{cases} \tag{8.31}$$

with

$$K_v = \sqrt{k^2 - \frac{\omega^2}{c^2}} \tag{8.32}$$

$$K_m = \sqrt{k^2 - \frac{\omega^2}{c^2} + \frac{\omega_p^2}{c^2}}. \tag{8.33}$$

The y-component of the field, as was mentioned previously, is not coupled to the others. This component is transverse everywhere, because $\nabla \cdot \mathbf{E} = 0$, even at the interfaces. Thus, E_y is not related to any change in the charge density, and will not be dealt with further.

The boundary conditions at $z = 0$ and $z = s$ impose the relations:

$$E_0 e^{-K_v s} = E_1 e^{-K_m s} + E_2 e^{K_m s} \tag{8.34a}$$

$$E_3 = E_1 + E_2 \tag{8.34b}$$

$$-\frac{ik}{K_v} E_0 e^{-K_v s} = \left(1 - \frac{\omega_p^2}{\omega^2}\right) \frac{ik}{K_m}\left(-E_1 e^{-K_m s} + E_2 e^{K_m s}\right) \tag{8.34c}$$

$$\frac{ik}{K_v} E_3 = \left(1 - \frac{\omega_p^2}{\omega^2}\right) \frac{ik}{K_m}\left(-E_1 + E_2\right). \tag{8.34d}$$

[4]This is not required, merely a consequence of equation (8.29).

Eliminating E_0 and E_3 gives

$$E_1 e^{-K_m s} + E_2 e^{K_m s} = \frac{1}{R}\left(E_1 e^{-K_m s} - E_2 e^{K_m s}\right) \tag{8.35a}$$

$$-E_1 - E_2 = \frac{1}{R}\left(E_1 - E_2\right). \tag{8.35b}$$

where

$$R = \frac{K_m \omega^2}{K_v\left(\omega^2 - \omega_p^2\right)}. \tag{8.36}$$

A final rearrangement gives

$$E_1(1-R)e^{-K_m s} = E_2(1+R)e^{K_m s} \tag{8.37a}$$

$$E_1(1+R) = E_2(1-R) \tag{8.37b}$$

so that the dispersion relation is

$$\frac{(1-R)^2}{(1+R)^2} = e^{2K_m s}. \tag{8.38}$$

This may be solved numerically: results are shown in figure 8.3. To obtain an approximate analytical solution, equation (8.38) is first expanded:

$$k = \frac{\omega}{c}\sqrt{\frac{\left[\tanh\left(K_m s/2\right)\right]^{\pm 2} + \omega^2/\left(\omega_p^2 - \omega^2\right)}{\left[\tanh\left(K_m s/2\right)\right]^{\pm 2} - \omega^4/\left(\omega_p^2 - \omega^2\right)^2}} \tag{8.39}$$

Note that the right-hand side contains K_m, which depends on k. However, equation (8.33) shows that for small k (and therefore small ω) $K_m \approx k_p$, where $k_p = \omega_p/c$. More precisely, let

$$\Delta = \left(\frac{k}{k_p}\right)^2 - \left(\frac{\omega}{\omega_p}\right)^2. \tag{8.40}$$

Because $k > \omega/c$, Δ is always positive and is of order $(k/k_p)^2$ at most; when $k \ll k_p$, $\Delta \ll 1$ and an expansion in terms of Δ is reasonable. Substituting for K_m gives

$$\tanh\left(\frac{K_m s}{2}\right) = \tanh\left(\frac{k_p s}{2}\right) + \Delta\left(\frac{k_p s}{4}\right)\left[\operatorname{sech}\left(\frac{k_p s}{2}\right)\right]^2 + \mathcal{O}[\Delta^2]. \tag{8.41}$$

Inserting this expression into equation (8.39) and expanding then leads to the following approximate form of the dispersion relation:

$$\frac{k}{k_p} = \frac{\omega}{\omega_p}\sqrt{1 + \left[\tanh\left(k_p s/2\right)\right]^{\mp 2}\left(\frac{\omega^2}{\omega_p^2} + 2\frac{\omega^4}{\omega_p^4}\right)} + \mathcal{O}\left[\left(\frac{\omega}{\omega_p}\right)^7\right] + \mathcal{O}\left[\left(\frac{\omega}{\omega_p}\right)^3 \Delta^2\right]. \tag{8.42}$$

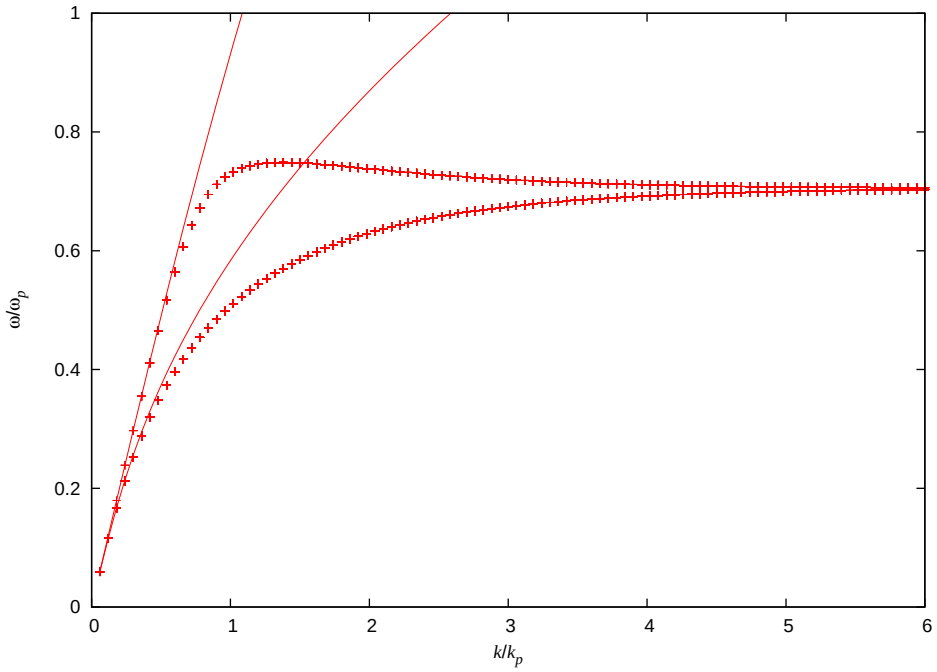


Figure 8.3: The dispersion relation for a slab with $s = 200$ and density parameter $r_s = 2.0$ (in Hartree atomic units). This density corresponds to a plasma frequency $\omega_p = 0.6124$. The crosses represent the results of applying a numerical root-finding algorithm to equation (8.39); the smooth lines correspond to the small-k approximation given in equation (8.42). In the limit of large k, ω tends to $\omega_p/\sqrt{2}$.

This demonstrates that Δ is in fact of order $(\omega/\omega_p)^4$. Rearranging equation (8.42) to give an expression for ω leads to

$$\frac{\omega}{\omega_p} \approx \sqrt{\frac{1}{2}\left[\tanh\left(\frac{k_p s}{2}\right)\right]^{\pm 2}\left(\sqrt{1+4\left[\tanh\left(\frac{k_p s}{2}\right)\right]^{\mp 2}\left(\frac{k}{k_p}\right)^2}-1\right)}. \qquad (8.43)$$

As will be demonstrated below, the two branches of the dispersion curve represent symmetric and antisymmetric charge oscillations; as the slab is made wider, $\tanh(k_p s/2) \to 1$ and the two modes become degenerate.

The surface plasmons are transverse: the only places at which the charge density changes (i.e., $\nabla \cdot \mathbf{E} \neq 0$) are the boundaries between vacuum and metal. The surface charge density is related via the boundary condition (8.28c) to the change in E_z at

the interface:

$$\rho_s = \epsilon_0 \begin{cases} \frac{ik}{K_m}\left(-E_1 + E_2\right) - \frac{ik}{K_v}E_3 & \text{at } z = 0 \\ -\frac{ik}{K_v}E_0 e^{-K_v s} - \frac{ik}{K_m}\left(-E_1 e^{-K_m s} + E_2 e^{K_m s}\right) & \text{at } z = s. \end{cases} \tag{8.44}$$

The dependence on x and t is exactly as for the **E**-field. Using equations (8.34), (8.36) and (8.38), the various field amplitudes may be related to each other and thus eliminated (except for one):

$$E_0 = \pm e^{K_v s} E_3 \tag{8.45a}$$

$$E_1 = \frac{1 - R}{2} E_3 \tag{8.45b}$$

$$E_2 = \frac{1 + R}{2} E_3. \tag{8.45c}$$

The surface charge density is then

$$\rho_s = \pm \frac{\epsilon_0 ik}{K_v} \frac{\omega_p^2}{\omega^2 - \omega_p^2} E_3. \tag{8.46}$$

The positive sign always applies at $z = 0$. For symmetric oscillation, it is also taken at $z = s$: the surface charge density variations on the two interfaces are then in phase. The negative sign applies at $z = s$ in the case of antisymmetric oscillation, so that the two surface charge densities are in antiphase. The lower branch of the dispersion graph corresponds to symmetric oscillation.

The electric field created by the surface charge is given by the solution of equation (8.1), with a charge density of

$$\rho(x, z, t) = \rho_0(\omega)\big(\delta(z) \pm \delta(z - s)\big) e^{i(\omega t - kx)}. \tag{8.47}$$

This longitudinal (electrostatic) field may be expressed as the gradient of a scalar potential, ϕ. Assuming a solution of the form

$$\phi(x, z, t) = \phi_z(z) e^{i(\omega t - kx)} \tag{8.48}$$

means that the scalar potential must satisfy the equation

$$\left(-k^2 + \frac{d^2}{dz^2}\right) \phi_z(z) = -\frac{\rho_0}{\epsilon_0} \left(\delta(z) \pm \delta(z - s)\right). \tag{8.49}$$

The left-hand side is the modified Helmholtz equation in one dimension, which has the Green's function

$$G(z, z') = \frac{1}{2k} e^{-k|z-z'|}. \tag{8.50}$$

Adapting this to equation (8.49) gives the full solution

$$\phi_z(z) = -\frac{\rho_0}{2k\epsilon_0} \left(e^{-k|z|} \pm e^{-k|z-s|} \right). \tag{8.51}$$

The electric field is then easy to calculate:

$$\mathbf{E} = -\nabla\phi \tag{8.52}$$

$$= -\frac{\rho_0}{2\epsilon_0} e^{i(\omega t - kx)} \left[f_x(z)\hat{\mathbf{x}} + f_z(z)\hat{\mathbf{z}} \right] \tag{8.53}$$

where

$$f_x(z) = i \left(e^{-k|z|} \pm e^{-k|z-s|} \right) \tag{8.54}$$

$$f_z(z) = \mathrm{sgn}(z)\, e^{-k|z|} \pm \mathrm{sgn}(z-s)\, e^{-k|z-s|} \tag{8.55}$$

The components of the field are plotted in figure 8.4.

For all surface plasmons, $\omega < \omega_p$.

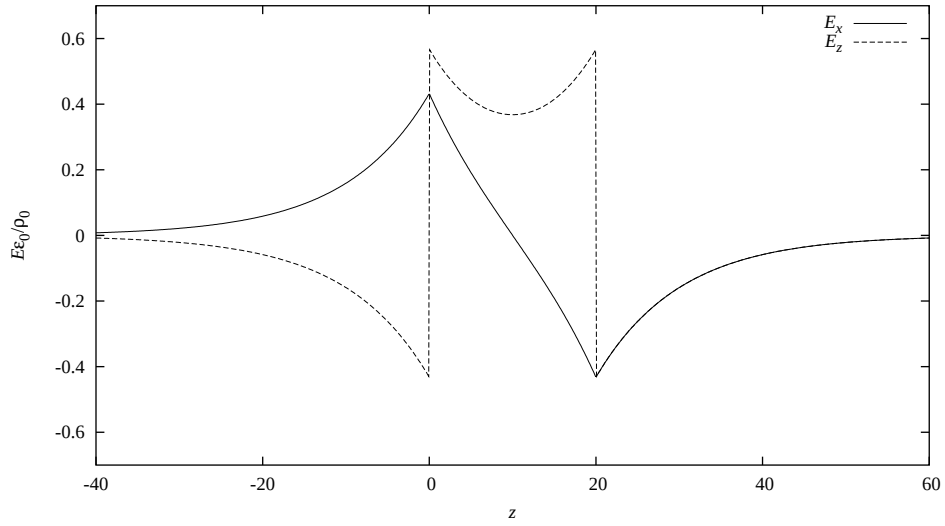
8.1.5 Bulk plasmons

Solutions also exist for the bulk plasmons. When $\omega = \omega_p$, the boundary condition on E_z (equation (8.28c)) implies that at the interface $E_z^{(v)} = 0$. The general form of the field in the vacuum is

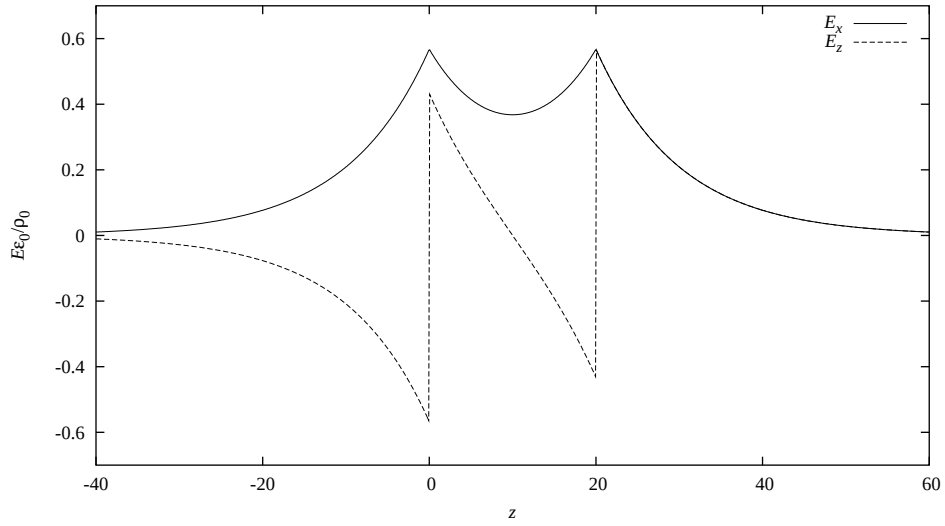
$$E_z^{(v)} = \begin{cases} E_0 e^{-K_v z} & \text{when } z > s \\ E_1 e^{K_v z} & \text{when } z < 0 \end{cases} \tag{8.56}$$

$$E_x^{(v)} = \begin{cases} \frac{ik}{K_v} E_0 e^{-K_v z} & \text{when } z > s \\ -\frac{ik}{K_v} E_1 e^{K_v z} & \text{when } z < 0 \end{cases} \tag{8.57}$$

with K_v given by equation (8.32). If $E_z^{(v)} = 0$ at $z = s$ and $z = 0$ then both E_0 and E_1 are also zero, and consequently there is no field in the vacuum. Consideration of equation (8.23) shows that this is also true in the special case $k = k_p = \omega_p/c$.



(a) Antisymmetric



(b) Symmetric

Figure 8.4: The x- and z-components of the longitudinal electric field associated with the surface plasmon oscillations.

The bulk plasmons (defined by $\omega = \omega_p$) must therefore satisfy the condition

$$E_x^{(m)}(z=0) = E_x^{(m)}(z=s) = 0, \tag{8.58}$$

along with equation (8.22). Suitable normal modes in the slab are:

$$E_x(z) = E_0 \sin k_z z \tag{8.59a}$$

$$E_z(z) = \frac{ik_z}{k} E_0 \cos k_z z \tag{8.59b}$$

where

$$k_z = \frac{n\pi}{s}, \quad n \in \mathbb{Z}^+. \tag{8.60}$$

There is a special case to consider: when $k = 0$, equation (8.14a) gives

$$\frac{d^2 E_x}{dz^2} = 0 \tag{8.61}$$

in the slab, which, when combined with condition (8.58), means that $E_x = 0$ everywhere. The z-component of the field is a function of z only, and is not restricted at the interfaces.

A more general solution, in which the in-plane propagation is no longer restricted to the x-direction, is obtained by a rotation about the z-axis:

$$\mathbf{E}_{\mathbf{k}_\parallel, k_z}(\mathbf{r}, t) = E_0 \left(\hat{\mathbf{k}}_\parallel \sin k_z z + \hat{\mathbf{z}} \frac{ik_z}{k_\parallel} \cos k_z z \right) e^{i(\omega_p t - \mathbf{k}_\parallel \cdot \mathbf{r}_\parallel)}. \tag{8.62}$$

The vector $\mathbf{k}_\parallel$ has taken the place of k; $\hat{\mathbf{k}}_\parallel$ and $\hat{\mathbf{z}}$ are unit vectors. It will be convenient to write $\mathbf{k}$ to represent the vector $(\mathbf{k}_\parallel, k_z)$.

In order to normalise the plasmon modes, it is required that

$$\int_V |\mathbf{E}_\mathbf{k}(\mathbf{r}, t)|^2 \, d^3 r = 1. \tag{8.63}$$

The volume over which the plasmons are normalised is $V = L^2 s$. This gives

$$E_0^2 = 2k_\parallel^2 / k^2 V. \tag{8.64}$$

In the new notation, k^2 means $k_\parallel^2 + k_z^2$. The normalised field is then

$$\mathbf{E}_\mathbf{k}(\mathbf{r}, t) = \frac{\sqrt{2}}{k\sqrt{V}} \left(\mathbf{k}_\parallel \sin k_z z + i \mathbf{k}_z \cos k_z z \right) e^{i(\omega_p t - \mathbf{k}_\parallel \cdot \mathbf{r}_\parallel)}. \tag{8.65}$$

This solution includes the case $\mathbf{k}_\parallel = \mathbf{0}$ discussed previously. Note that k_z is restricted to positive values.

Any electric field within the slab may be expressed as a sum of the normal modes described in equation (8.65). However, any physical field is real; it is useful to convert the set of complex modes to an equivalent set of real modes by forming appropriate linear combinations.[5]

The symmetry relation

$$\mathbf{E}_{\mathbf{k}_\parallel \mathbf{k}_z} = -\mathbf{E}^*_{(-\mathbf{k}_\parallel)\mathbf{k}_z} \tag{8.66}$$

which follows from equation (8.65) shows that real modes may be obtained by taking the combinations

$$\begin{aligned}
\mathbf{E}_{1\mathbf{k}}(\mathbf{r}) &= \frac{i}{\sqrt{2}}\left(\mathbf{E}_{\mathbf{k}_\parallel \mathbf{k}_z}(\mathbf{r}) + \mathbf{E}_{(-\mathbf{k}_\parallel)\mathbf{k}_z}(\mathbf{r})\right) \\
&= \frac{2}{k\sqrt{V}}\left(\mathbf{k}_\parallel \sin k_z z \sin \mathbf{k}_\parallel \cdot \mathbf{r}_\parallel - \mathbf{k}_z \cos k_z z \cos \mathbf{k}_\parallel \cdot \mathbf{r}_\parallel\right) \qquad (8.67a) \\
\mathbf{E}_{2\mathbf{k}}(\mathbf{r}) &= \frac{1}{\sqrt{2}}\left(\mathbf{E}_{\mathbf{k}_\parallel \mathbf{k}_z}(\mathbf{r}) - \mathbf{E}_{(-\mathbf{k}_\parallel)\mathbf{k}_z}(\mathbf{r})\right) \\
&= \frac{2}{k\sqrt{V}}\left(\mathbf{k}_\parallel \sin k_z z \cos \mathbf{k}_\parallel \cdot \mathbf{r}_\parallel + \mathbf{k}_z \cos k_z z \sin \mathbf{k}_\parallel \cdot \mathbf{r}_\parallel\right). \qquad (8.67b)
\end{aligned}$$

There are now two modes for each $\mathbf{k}$-vector, so the number of labels required is reduced by half. The restriction applies to the xy-plane: if $(\mathbf{k}_\parallel + \mathbf{k}_z)$ is a valid label for a mode, then $(-\mathbf{k}_\parallel + \mathbf{k}_z)$ is not. With this restriction, the modes described in equation (8.67) form a complete orthonormal set, in the sense that

$$\int_V \mathbf{E}_{j\mathbf{k}}(\mathbf{r}) \cdot \mathbf{E}_{j'\mathbf{k}'}(\mathbf{r})\, d^3r = \delta_{jj'}\delta_{\mathbf{k}\mathbf{k}'}. \tag{8.68}$$

The special case $\mathbf{k}_\parallel = \mathbf{0}$ is included in equation (8.67), with the understanding that there is no type 2 mode.

The potential associated with these normal modes may be obtained by integra-

[5]When quantising a Hamiltonian expressed in terms of the normal mode coordinates, it is convenient for them to represent real quantities, and thus to be associated with Hermitian operators.

tion:

$$\begin{aligned}\phi_{1\mathbf{k}}(\mathbf{r}) &= \int_{\mathbf{r}'=\mathbf{0}}^{\mathbf{r}'=\mathbf{r}} \frac{2}{k\sqrt{V}} \left(\mathbf{k}_\parallel \sin k_z z \sin \mathbf{k}_\parallel \cdot \mathbf{r}_\parallel - \mathbf{k}_z \cos k_z z \cos \mathbf{k}_\parallel \cdot \mathbf{r}_\parallel\right) \cdot d\mathbf{r}' \\ &= -\frac{2}{k\sqrt{V}} \sin k_z z \cos \mathbf{k}_\parallel \cdot \mathbf{r}_\parallel \qquad (8.69) \\ \phi_{2\mathbf{k}}(\mathbf{r}) &= \int_{\mathbf{r}'=\mathbf{0}}^{\mathbf{r}'=\mathbf{r}} \frac{2}{k\sqrt{V}} \left(\mathbf{k}_\parallel \sin k_z z \cos \mathbf{k}_\parallel \cdot \mathbf{r}_\parallel + \mathbf{k}_z \cos k_z z \sin \mathbf{k}_\parallel \cdot \mathbf{r}_\parallel\right) \cdot d\mathbf{r}' \\ &= \frac{2}{k\sqrt{V}} \sin k_z z \sin \mathbf{k}_\parallel \cdot \mathbf{r}_\parallel. \qquad (8.70)\end{aligned}$$

8.1.6 Surface plasmons in electrostatic theory

The bulk plasmons described in section 8.1.5 do not have any associated magnetic field, and the electric field is purely longitudinal. They are based solely on the Coulomb interaction. In contrast, the surface plasmons of section 8.1.4 cannot be described using the electrostatic theory alone. The normal-mode theory of chapter 7 uses only the Coulomb interaction, because this is all that is contained in the standard Schrödinger equation; the surface plasmons corresponding to the electrostatic theory must therefore be obtained.

In this theory, equation (8.4) is replaced by

$$m_e \dot{\mathbf{v}} = e\nabla\phi. \qquad (8.71)$$

The electrons now interact only via the Coulomb force. The current-field relation becomes

$$\dot{\mathbf{J}} = -\epsilon_0 \omega_p^2 \nabla\phi. \qquad (8.72)$$

Combining this expression with the equation of continuity,

$$\nabla \cdot \mathbf{J} = -\frac{\partial \rho}{\partial t}, \qquad (8.73)$$

gives the electrostatic equation of motion for the scalar potential:

$$\nabla^2 \ddot{\phi} = -\nabla \cdot \left(\omega_p^2 \nabla\phi\right). \qquad (8.74)$$

This reproduces equation (7.62). In the slab and the vacuum, ω_p^2 is constant; the equation of motion in each of these regions becomes

$$\nabla^2 \ddot{\phi} = -\omega_p^2 \nabla^2 \phi. \tag{8.75}$$

Searching for a travelling-wave solution of the form

$$\phi(\mathbf{r}, t) = \phi_z(z) e^{i(\omega t - kx)} \tag{8.76}$$

leads to the following equation for ϕ_z:

$$\omega^2 \left(k^2 \phi_z \; - \frac{d^2 \phi_z}{dz^2} \right) = \omega_p^2 \left(k^2 \phi_z - \frac{d^2 \phi_z}{dz^2} \right). \tag{8.77}$$

As noted above, the bulk plasmons (with $\omega = \omega_p$) are solutions of this equation. When $\omega \neq \omega_p$, the solutions must satisfy Laplace's equation:

$$\frac{d^2 \phi_z}{dz^2} - k^2 \phi_z = 0. \tag{8.78}$$

In the case of the slab, the solutions take the simple form

$$\phi_z = \begin{cases} Ae^{kz} & \text{when } z < 0 \\ Be^{kz} + Ce^{-kz} & \text{when } 0 < z < s \\ De^{-kz} & \text{otherwise.} \end{cases} \tag{8.79}$$

This ensures that $\phi \to 0$ as $z \to \pm\infty$, as long as the trivial solution with $k = 0$ is excluded. The other boundary conditions (as derived in section 8.1.3) are:

$$\frac{\partial \phi^{(m)}}{\partial x} = \frac{\partial \phi^{(v)}}{\partial x} \tag{8.80}$$

$$\left(1 - \frac{\omega_p^2}{\omega^2} \right) \frac{\partial \phi^{(m)}}{\partial z} = \frac{\partial \phi^{(v)}}{\partial z}. \tag{8.81}$$

Applying these conditions to equation (8.79) leads to the dispersion relation

$$\frac{\omega}{\omega_p} = \sqrt{\frac{1 \pm e^{-ks}}{2}} \tag{8.82}$$

which is plotted in figure 8.5. The relationship between the amplitudes A, B, C,

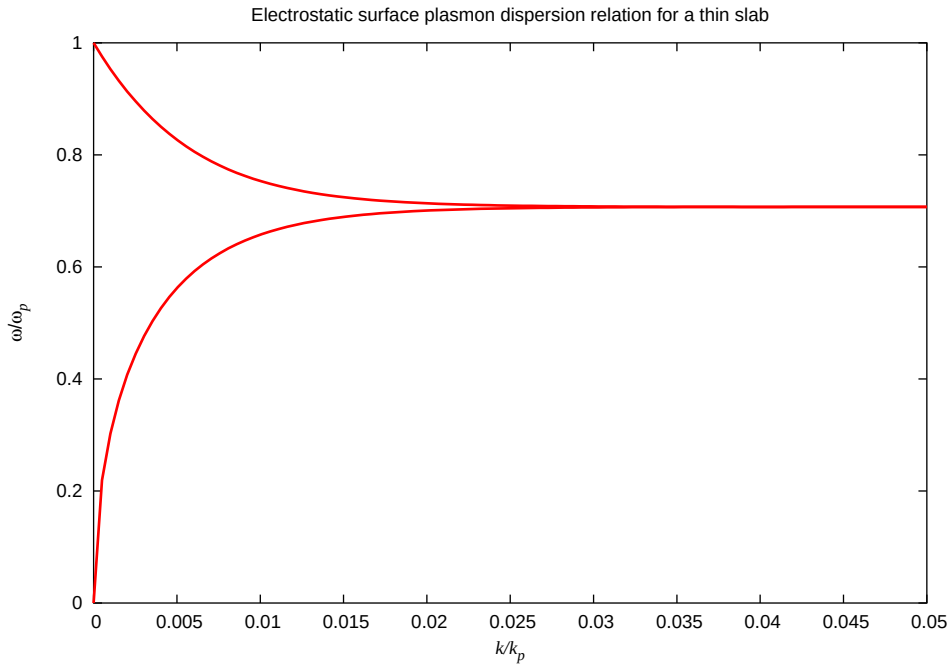


Figure 8.5: The dispersion relation for surface plasmons in the electrostatic theory. The slab width is 20 au. For metallic densities, the Fermi wave vector is of order 1, whereas $k_p \sim 0.01$. The plasmon frequency very rapidly reaches the large-k limit of $\omega_p/\sqrt{2}$; this is the result obtained for a semi-infinite system, and indicates that the coupling between surface plasmon modes on the two sides of the slab is weaker than that obtained when using the full dynamical theory.

and D is also determined, giving

$$\phi_k(\mathbf{r}) = A_k e^{-ikx} \begin{cases} e^{kz} & \text{when } z < 0 \\ \frac{e^{kz} \mp e^{-k(z-s)}}{1 \mp e^{ks}} & \text{when } 0 < z < s \\ \mp e^{-k(z-s)} & \text{otherwise.} \end{cases} \tag{8.83}$$

As with the bulk plasmons earlier, this result can be generalised to allow propagation in any direction parallel to the slab by a rotation of the axes. The form of

the scalar potential is then

$$\phi^{\pm}_{1\mathbf{k}_\parallel}(\mathbf{r}) = \left(\frac{1 \mp e^{-k_\parallel s}}{2L^2 k_\parallel}\right)^{1/2} \cos \mathbf{k}_\parallel \cdot \mathbf{r}_\parallel \begin{cases} e^{k_\parallel z} & \text{when } z < 0 \\ \frac{e^{k_\parallel z} \mp e^{-k_\parallel (z-s)}}{1 \mp e^{k_\parallel s}} & \text{when } 0 < z < s \\ \mp e^{-k_\parallel (z-s)} & \text{otherwise} \end{cases} \tag{8.84a}$$

$$\phi^{\pm}_{2\mathbf{k}_\parallel}(\mathbf{r}) = \left(\frac{1 \mp e^{-k_\parallel s}}{2L^2 k_\parallel}\right)^{1/2} \sin \mathbf{k}_\parallel \cdot \mathbf{r}_\parallel \begin{cases} e^{k_\parallel z} & \text{when } z < 0 \\ \frac{e^{k_\parallel z} \mp e^{-k_\parallel (z-s)}}{1 \mp e^{k_\parallel s}} & \text{when } 0 < z < s \\ \mp e^{-k_\parallel (z-s)} & \text{otherwise.} \end{cases} \tag{8.84b}$$

The complex solutions have been combined into pairs of real modes; in addition, these have been normalised so that

$$\int_V \left[\nabla \phi_{\mathbf{k}_\parallel}(\mathbf{r})\right]^2 d^3r = 1. \tag{8.85}$$

8.2 The plasmon wave function for slab systems

Equations (7.81) and (7.82) give a prescription for calculating a Jastrow factor based on plasmon normal modes. The normal modes appropriate to jellium slab systems have been obtained in the preceding sections; here, the prescription of chapter 7 will be applied to these modes.

The bulk plasmon modes given in equations (8.69) and (8.70) are degenerate in frequency: all oscillate at $\omega = \omega_p$. The restrictions on $\mathbf{k}$ for these modes are as follows:

- each $\mathbf{k}_\parallel$ is a reciprocal lattice vector;
- if $\mathbf{k}_\parallel$ is included, $-\mathbf{k}_\parallel$ is not;
- $k_z = n\pi/s$, where $n = 1, 2, 3, \ldots$ and s is the slab width;
- the magnitude is limited: $|\mathbf{k}| < k_c$;
- there is no type 2 mode when $\mathbf{k}_\parallel = \mathbf{0}$.

The cut-off wave vector k_c is related to the electron density [73]:

$$k_c \sim \frac{1}{\sqrt{r_s}}. \tag{8.86}$$

In a homogeneous system, plasmons with wave vector greater than k_c are no longer well-defined; at $k \sim k_c$, the (almost flat) plasmon dispersion curve merges with the particle-hole continuum [17]. In general, it is not obvious how to apply the cut-off in an inhomogeneous system. However, in the slab system under consideration here, the electron density is approximately constant within the slab and zero outside; it is therefore reasonable to apply the same cut-off as would be used in a homogeneous system of equivalent density.

Substituting these functions in place of the $\{f_i\}$ in equation (7.81) gives

$$\begin{aligned}\chi_{\text{bulk}}(\mathbf{r}) &= \frac{e^2}{\hbar\omega_p\epsilon_0}\sum_{\mathbf{k}}\Big(\phi_{1\mathbf{k}}(\mathbf{r})\int_V \phi_{1\mathbf{k}}(\mathbf{r}')\bar{n}(z')\,d^3r' \\ &\quad + \phi_{2\mathbf{k}}(\mathbf{r})\int_V \phi_{2\mathbf{k}}(\mathbf{r}')\bar{n}(z')\,d^3r'\Big) \\ &= \frac{e^2}{\hbar\omega_p\epsilon_0}\sum_{k_z}\frac{4}{k_z^2 s}\sin k_z z\int_0^s \bar{n}(z')\sin k_z z'\,dz'\;\Theta(z)\Theta(s-z)\end{aligned} \tag{8.87}$$

while substitution in equation (7.82) gives

$$\begin{aligned}u_{\text{bulk}}(\mathbf{r},\mathbf{r}') &= \frac{e^2}{\hbar\omega_p\epsilon_0}\sum_{\mathbf{k}}\Big(\phi_{1\mathbf{k}}(\mathbf{r})\phi_{1\mathbf{k}}(\mathbf{r}') + \phi_{2\mathbf{k}}(\mathbf{r})\phi_{2\mathbf{k}}(\mathbf{r}')\Big) \\ &= \frac{e^2}{\hbar\omega_p\epsilon_0 V}\sum_{\mathbf{k}}\frac{4}{k^2}\cos\mathbf{k}_\parallel\cdot(\mathbf{r}_\parallel-\mathbf{r}'_\parallel)\sin k_z z\sin k_z z' \\ &\quad\times\Theta(z)\Theta(s-z)\Theta(z')\Theta(s-z').\end{aligned} \tag{8.88}$$

The sum is understood to include the term with $\mathbf{k}_\parallel = \mathbf{0}$. The Heaviside functions appear because the bulk plasmons have zero amplitude outside the slab. Note that only the modes with $k_\parallel = 0$ contribute to the χ-function; all others integrate to zero. Taking the density to be constant (to be consistent with the derivation of the plasmon modes), $\bar{n}(z) = n_0$, giving

$$\chi_{\text{bulk}}(\mathbf{r}) = \frac{e^2}{\hbar\omega_p\epsilon_0}\sum_{k_z}\frac{4n_0}{k_z^3 s}\sin k_z z(1-\cos k_z s)\Theta(z)\Theta(s-z). \tag{8.89}$$

Unlike the bulk plasmons, the electrostatic surface plasmon modes (equations (8.84a) and (8.84b)) are not degenerate in frequency; the dispersion relation is given in equation (8.82), and plotted in figure 8.5. The same restrictions on $\mathbf{k}_\parallel$ apply as in the bulk case,[6] with the additional constraint that $\mathbf{k}_\parallel \neq \mathbf{0}$. The surface plasmon contribution to the two-body function is therefore

$$\begin{aligned} u_{\text{surf}}(\mathbf{r},\mathbf{r}') &= \frac{e^2}{\hbar\omega_p\epsilon_0}\sum_{\mathbf{k}_\parallel}\Bigg(\sqrt{\frac{2}{1+e^{-k_\parallel s}}}\Big[\phi^+_{1\mathbf{k}_\parallel}(\mathbf{r})\phi^+_{1\mathbf{k}_\parallel}(\mathbf{r}')+\phi^+_{2\mathbf{k}_\parallel}(\mathbf{r})\phi^+_{2\mathbf{k}_\parallel}(\mathbf{r}')\Big] \\ &\quad+\sqrt{\frac{2}{1-e^{-k_\parallel s}}}\Big[\phi^-_{1\mathbf{k}_\parallel}(\mathbf{r})\phi^-_{1\mathbf{k}_\parallel}(\mathbf{r}')+\phi^-_{2\mathbf{k}_\parallel}(\mathbf{r})\phi^-_{2\mathbf{k}_\parallel}(\mathbf{r}')\Big]\Bigg) \\ &= \frac{e^2}{\hbar\omega_p\epsilon_0 V}\sum_{\mathbf{k}_\parallel}\cos\mathbf{k}_\parallel\cdot\left(\mathbf{r}_\parallel-\mathbf{r}'_\parallel\right)F_{k_\parallel}(z,z') \end{aligned} \tag{8.90}$$

where $F_{k_\parallel}$ is defined in table 8.1. The constants used in this definition are

$$\begin{aligned} A_{k_\parallel} &= \frac{s\sqrt{2}}{2k_\parallel\sqrt{1+e^{-k_\parallel s}}} \\ B_{k_\parallel} &= \frac{s\sqrt{2}}{2k_\parallel\sqrt{1-e^{-k_\parallel s}}} \\ C_{k_\parallel} &= 1-e^{k_\parallel s} \\ D_{k_\parallel} &= 1+e^{k_\parallel s}. \end{aligned} \tag{8.91}$$

There is no surface plasmon contribution to the one-body term; the integral

$$\int_V \phi^\pm_{j\mathbf{k}_\parallel}(\mathbf{r}')\bar{n}(z')\,d^3r' \tag{8.92}$$

is zero for all surface plasmon modes $\phi^\pm_{j\mathbf{k}_\parallel}$.

[6]The surface plasmon dispersion curve tends to $\omega_p/\sqrt{2}$ when $k \gg 1/s$; it should therefore meet the electron-hole continuum earlier than the bulk plasmon dispersion curve (which is flat, lying at $\omega = \omega_p$). It could be argued that a smaller cut-off should be used for the surface plasmon modes, but because the cut-off is only expected to be correct to within an order of magnitude, this seems unnecessary.

z	z'	$F_{k_\parallel}(z,z')$
<0	<0	$e^{k_\parallel(z+z'-s)}\left(-A_{k_\parallel}C_{k_\parallel}+B_{k_\parallel}D_{k_\parallel}\right)$
<0	$0\to s$	$e^{k_\parallel(z+z'-s)}\left(-A_{k_\parallel}+B_{k_\parallel}\right)+e^{k_\parallel(z-z')}\left(A_{k_\parallel}+B_{k_\parallel}\right)$
<0	$>s$	$e^{k_\parallel(z-z')}\left(A_{k_\parallel}C_{k_\parallel}+B_{k_\parallel}D_{k_\parallel}\right)$
$0\to s$	$0\to s$	$2\cosh\left[k_\parallel(z+z'-s)\right]\left(-A_{k_\parallel}/C_{k_\parallel}+B_{k_\parallel}/D_{k_\parallel}\right)$ $+2\cosh\left[k_\parallel(z-z')\right]\left(A_{k_\parallel}/C_{k_\parallel}+B_{k_\parallel}/D_{k_\parallel}\right)$
$0\to s$	$>s$	$e^{k_\parallel(s-z-z')}\left[-A_{k_\parallel}+B_{k_\parallel}\right]+e^{k_\parallel(z-z')}\left[A_{k_\parallel}+B_{k_\parallel}\right]$
$>s$	$>s$	$e^{k_\parallel(s-z-z')}\left(-A_{k_\parallel}C_{k_\parallel}+B_{k_\parallel}D_{k_\parallel}\right)$

Table 8.1: The function $F_{k_\parallel}(z,z')$. This contains all the z- and z'-dependence of the part of $u_{\rm pl}$ arising from the surface plasmon contribution. The constants $A_{k_\parallel}$, $B_{k_\parallel}$, $C_{k_\parallel}$ and $D_{k_\parallel}$ are defined in equation (8.91).

Combining the bulk and surface terms gives the full plasmon two-body function:

$$\begin{aligned} u_{\rm pl}(\mathbf{r},\mathbf{r}') &= u_{\rm bulk}(\mathbf{r},\mathbf{r}')+u_{\rm surf}(\mathbf{r},\mathbf{r}') \\ &= \frac{e^2}{\hbar\omega_p\epsilon_0 V}\sum_{\mathbf{k}_\parallel}\cos\mathbf{k}_\parallel\cdot(\mathbf{r}_\parallel-\mathbf{r}'_\parallel)\Bigg(F_{k_\parallel}(z,z')+ \\ &\quad+\sum_{k_z}\frac{4}{k^2}\sin k_z z\sin k_z z'\Theta(z)\Theta(s-z)\Theta(z')\Theta(s-z')\Bigg). \end{aligned} \tag{8.93}$$

8.2.1 Approximate analytic solution for an infinite slab

It is possible to obtain an approximation to $u_{\rm pl}$ by taking the slab width s and the cell size L to infinity, in which case the sums become integrals:

$$\sum_{k_z>0} \rightarrow \frac{s}{\pi}\int_0^\infty dk_z \tag{8.94}$$

$$\sum_{\mathbf{k}_\parallel} \rightarrow \frac{L^2}{(2\pi)^2}\int_{-\infty}^\infty dk_x \int_0^\infty dk_y \tag{8.95}$$

The system is no longer a slab but a single surface (at $z = 0$) of infinite extent in the x- and y-directions.

Note that a further approximation has also been introduced here: the **k**-space cut-off described in the previous section has been neglected in order to enable the integrals to be solved analytically.

In this limit, the bulk term becomes

$$\begin{aligned} u^\infty_{\rm bulk}(\mathbf{r},\mathbf{r}') &= \frac{e^2}{\hbar\omega_p\epsilon_0 V}\int_{\mathbf{k}} \frac{4}{k^2}\cos\mathbf{k}_\parallel\cdot(\mathbf{r}_\parallel-\mathbf{r}'_\parallel)\sin k_z z \sin k_z z' \,\frac{V d^3k}{\pi(2\pi)^2}\,\Theta(z)\Theta(z') \\ &= \frac{e^2}{2\pi^3\epsilon_0\hbar\omega_p}\int_0^\infty dk_z\,\Big(\cos k_z(z-z') - \cos k_z(z+z')\Big) \\ &\quad\times\int_0^\infty dk_\parallel\,\frac{k_\parallel}{k_\parallel^2+k_z^2}\int_0^\pi d\theta\,\cos\big(k_\parallel\Delta r_\parallel\cos(\theta-\phi)\big)\,\Theta(z)\Theta(z') \end{aligned} \tag{8.96}$$

where $\Delta\mathbf{r}_\parallel = \mathbf{r}_\parallel - \mathbf{r}'_\parallel$ and $\phi = \tan^{-1}(\Delta y/\Delta x)$. Performing the integration gives

$$\begin{aligned} u^\infty_{\rm bulk}(\mathbf{r},\mathbf{r}') &= \frac{e^2}{2\pi^3\epsilon_0\hbar\omega_p}\int_0^\infty dk_z\,\Big(\cos k_z(z-z') - \cos k_z(z+z')\Big)\Theta(z)\Theta(z') \\ &\quad\times\int_0^\infty dk_\parallel\,\frac{k_\parallel}{k_\parallel^2+k_z^2}\pi J_0\left(k_\parallel\Delta r_\parallel\right)\Theta(z)\Theta(z') \\ &= \frac{e^2}{2\pi^2\epsilon_0\hbar\omega_p}\int_0^\infty dk_z\,\Big(\cos k_z(z-z') - \cos k_z(z+z')\Big)K_0\left(k_z\Delta r_\parallel\right) \\ &\quad\times\Theta(z)\Theta(z') \\ &= \frac{e^2}{2\pi^2\epsilon_0\hbar\omega_p}\left(\frac{\pi}{2\sqrt{(z-z')^2+(\Delta r_\parallel)^2}} - \frac{\pi}{2\sqrt{(z+z')^2+(\Delta r_\parallel)^2}}\right)\Theta(z)\Theta(z') \\ &= \frac{e^2}{4\pi\epsilon_0\hbar\omega_p|\mathbf{r}-\mathbf{r}'|}\left(1-\sqrt{\frac{1}{1+\frac{4zz'}{|\mathbf{r}-\mathbf{r}'|^2}}}\right)\Theta(z)\Theta(z'). \end{aligned} \tag{8.97}$$

Several standard integrals involving Bessel functions have been used.

In the limit $s \to \infty$, the surface term simplifies to

$$u_{\text{surf}}^{s\to\infty}(\mathbf{r},\mathbf{r}') = \frac{\sqrt{2}e^2}{\epsilon_0\hbar\omega_p L^2}\sum_{\mathbf{k}_\parallel}\frac{1}{k_\parallel}\cos\mathbf{k}_\parallel\cdot\Delta\mathbf{r}_\parallel\; e^{-k_\parallel(|z|+|z'|)}. \tag{8.98}$$

When the limit $L \to \infty$ is also taken, the prescription of equation (8.95) for converting the summation to an integral gives

$$\begin{aligned} u_{\text{surf}}^{\infty}(\mathbf{r},\mathbf{r}') &= \frac{\sqrt{2}e^2}{4\pi^2\epsilon_0\hbar\omega_p}\int_0^\infty k\;dk\int_0^\pi d\theta\;\frac{e^{-k_\parallel(|z|+|z'|)}}{k}\cos\mathbf{k}_\parallel\cdot\Delta\mathbf{r}_\parallel \\ &= \frac{\sqrt{2}e^2}{4\epsilon_0\pi^2\hbar\omega_p}\int_0^\infty dk\; e^{-k_\parallel(|z|+|z'|)}\int_0^\pi d\theta\;\cos\left(k_\parallel\Delta r_\parallel\cos(\theta-\phi)\right) \\ &= \frac{\sqrt{2}e^2}{4\pi^2\epsilon_0\hbar\omega_p}\int_0^\infty dk\; e^{-k_\parallel(|z|+|z'|)}\pi J_0\left(k_\parallel\Delta r_\parallel\right) \\ &= \frac{\sqrt{2}e^2}{4\pi\epsilon_0\hbar\omega_p\sqrt{(|z|+|z'|)^2+(\Delta r_\parallel)^2}}. \end{aligned} \tag{8.99}$$

The full plasmon two-body function in this limit is therefore

$$\begin{aligned} u_{\text{pl}}^{\infty}(\mathbf{r},\mathbf{r}') = \frac{e^2}{4\pi\epsilon_0\hbar\omega_p}\Bigg[&\Theta(z)\Theta(z')\left(\frac{1}{|\mathbf{r}-\mathbf{r}'|}-\frac{1}{\sqrt{(z+z')^2+(\Delta r_\parallel)^2}}\right) \\ &+\frac{\sqrt{2}}{\sqrt{(|z|+|z'|)^2+(\Delta r_\parallel)^2}}\Bigg]. \end{aligned} \tag{8.100}$$

This function is plotted in figures 8.6, 8.7, 8.8 and 8.9. The contribution from the bulk plasmons is only relevant when both electrons are inside the metal; when the electrons are deep inside, u_{pl}^{∞} tends to the expected homogeneous electron gas form. The correlation is boosted for electrons closer to the boundary.

The singularities present in equation (8.100) are a feature of the approximations which have been made; they do not appear in the original expressions relevant to a finite slab and cell. In any case, the plasmon theory is not expected to predict electron-electron correlation accurately at short range; the cusp conditions described in appendix B and discussed in the next section give more information about this regime.

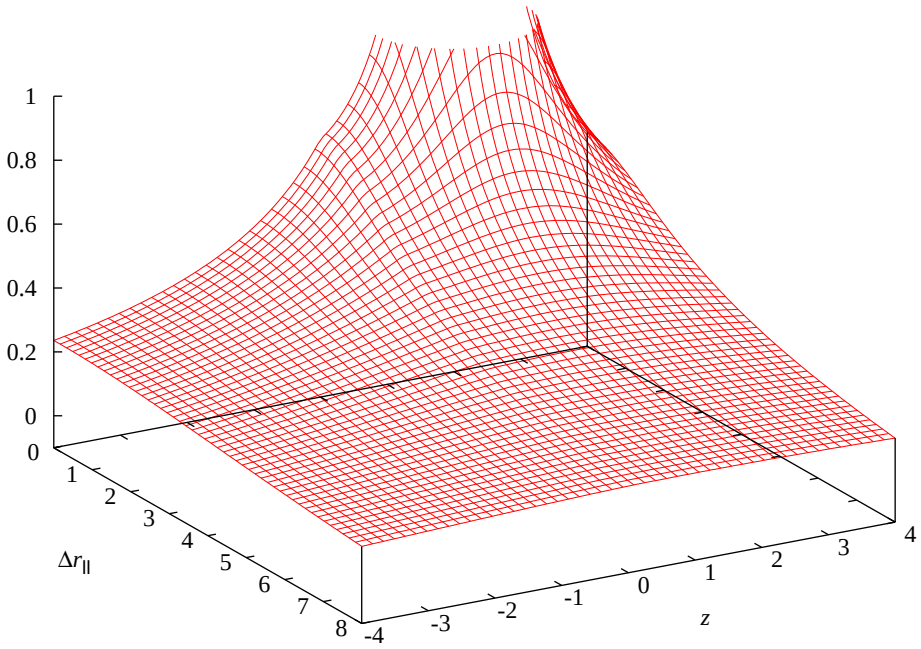


Figure 8.6: The plasmon contribution to the two-body function in the Jastrow factor. In this and subsequent graphs, one electron is fixed (here at $z' = 2.0$, inside the slab); the plot shows the dependency of u_{pl}^{∞} on the parallel separation $\Delta r_{\|}$ and the z-coordinate of the other electron.

8.3 Creating a realistic Jastrow factor

Equation (8.89) and (8.93) are the final results of applying the plasmon theory to the jellium slab system, and they are obtained from a consideration of the plasmon degrees of freedom only.

The full electronic Hamiltonian (equation (2.9)) can be separated into long-range plasmonic and short-range terms:

$$\hat{H} = \hat{H}_{\mathrm{pl}} + \hat{H}_{\mathrm{sr}}. \tag{8.101}$$

The plasmon Hamiltonian $\hat{H}_{\mathrm{pl}}$ is the one described in chapter 7 and appearing in different forms in equations (7.11), (7.16), (7.19) and (7.47). The long-range part of

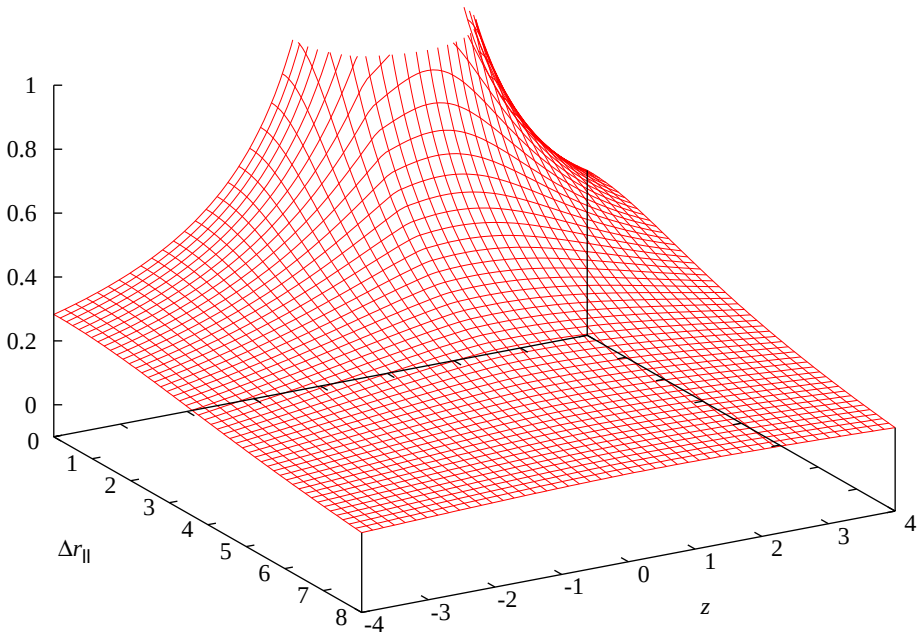


Figure 8.7: The fixed electron is at $z' = 1.0$: inside the slab but close to the surface.

the Coulomb interaction is subsumed into $\hat{H}_{\rm pl}$; the remaining *screened* interaction stays in $\hat{H}_{\rm sr}$. The Schrödinger equation has been effectively separated; the ground-state eigenfunction is then a product of the form

$$\Psi_{\rm pr} = \Psi_{\rm pl}\Psi_{\rm sr}. \tag{8.102}$$

The basis for the separation is the physical observation that plasmons exist.

The short-range wave function does not contain information about the long-range correlations between electrons; however, it should contain the short-range correlations. The conventional trial wave functions used in QMC have

$$\Psi_{\rm sr} = D^{\uparrow}D^{\downarrow}, \tag{8.103}$$

where the determinants are made up of orbitals obtained from a mean-field calculation. This wave function does not include the necessary information about short-

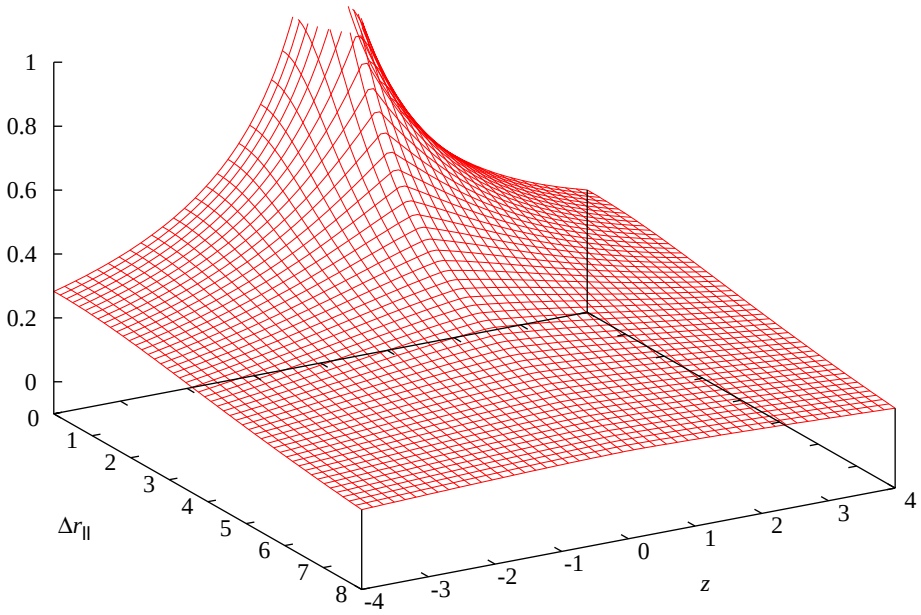


Figure 8.8: The fixed electron is now just outside the slab at $z' = -1.0$; the bulk plasmons no longer contribute to u_{pl}^{∞}.

range correlation between electrons; this must therefore be introduced artificially.

The behaviour of the wave function when two electrons are in close proximity is analysed in appendix B; the divergence in the potential energy in this limit must be compensated by a cancelling divergence in the kinetic energy, which can only arise as a consequence of a cusp (gradient discontinuity) in the wave function. The *cusp conditions*, originally due to Kato [37], specify only the gradient with respect to the separation of the two electrons in the limit of this separation tending to zero; they do not determine the value of the wave function.

In appendix B, it is shown that a wave function with the desired short-range behaviour is

$$\Psi = \exp\left[-\frac{1}{2}\sum_{i\neq j} u_{\mathrm{cusp}}(\mathbf{r}_i, \mathbf{r}_j)\right]\Psi_{\mathrm{smooth}}, \tag{8.104}$$

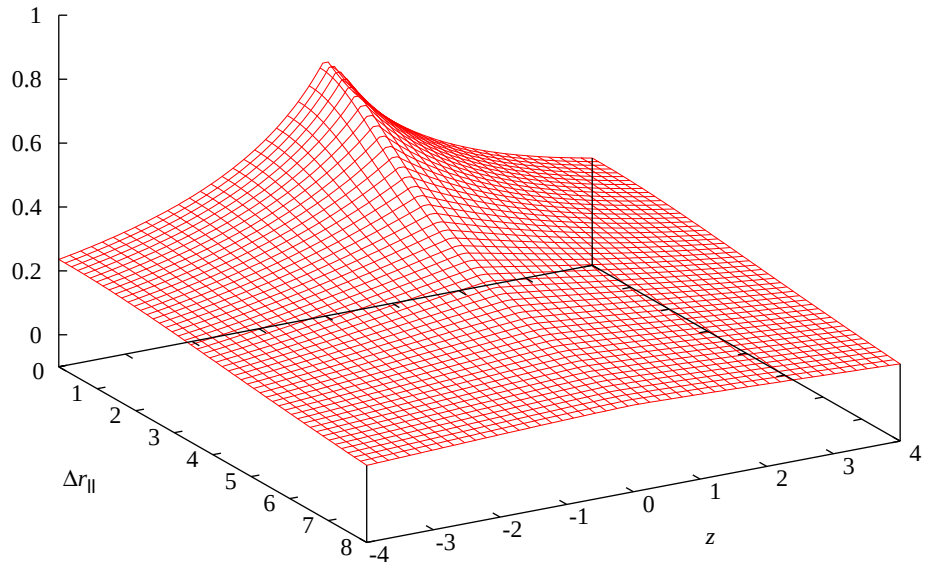


Figure 8.9: The fixed electron is now further outside the slab ($z' = -2.0$), and the strength of the correlation is reduced.

where Ψ_{smooth} is the unmodified cuspless wave function and

$$u_{\text{cusp}}(\mathbf{x}_i, \mathbf{x}_j) = \left(\frac{m_e e^2}{4\pi\epsilon_0\hbar^2}\right)\beta_{\sigma_i\sigma_j} e^{-\alpha_{\sigma_i\sigma_j} r_{ij}}. \tag{8.105}$$

The analysis depends on the original wave function Ψ_{smooth} being smooth (hence the subscript): more precisely, it must be possible to expand this function in a Taylor series about the point $\mathbf{r}_i = \mathbf{r}_j$, irrespective of the positions of the other electrons. This is true for the two-determinant wave function Ψ_{sr}, and almost true for the product wave function Ψ_{pr}. The problem is that the plasmon contribution Ψ_{pl} has cusps at the slab boundaries.

In fact, the cusps in the plasmon wave function create a more serious problem in QMC. These extended gradient discontinuities lead to singularities in the second derivative of Ψ_{pl} which should make a finite contribution to the energy expectation

value. However, because these points (which constitute a region of zero volume) are never sampled in a QMC simulation, their contribution is missed, and any expectation values involving the Laplacian will not be accurate. This is not the case for the integrable point singularities in the Laplacian of the electron-electron cusps, which do not generate errors in QMC.

Therefore, in order to achieve proper sampling in QMC, and also to implement the electron-electron cusp conditions correctly, it is necessary to smooth out the plasmon wave function: both u_{pl} and χ_{bulk} must be modified to have continuous first and second derivatives.

8.3.1 Removing undesirable cusps

The cusps in u_{bulk}, u_{surf} and χ_{bulk} are a consequence of the piecewise way in which these functions are constructed. A simple and appealing method of removing the cusps is to blur the boundaries between the regions of definition.

For the purpose of illustration, consider an arbitrary function $f(x)$ with the following form near the point $x = x_0$:

$$f(x) = f_1(x)\Theta(x - x_0) + f_2(x)\Theta(x_0 - x). \tag{8.106}$$

A smooth approximation to f is

$$f^{\mathrm{s}}(x) = f_1(x)T(x - x_0) + f_2(x)T(x_0 - x) \tag{8.107}$$

where the smoothing function T has the following properties:

- the value and first and second derivatives are continuous;
- $\lim_{x\to\infty} T(x) = 1$;
- $\lim_{x\to-\infty} T(x) = 0$;
- the transition between these limiting values takes place over some quantifiable distance Δx.

In addition, it is desirable that $T(0) = \frac{1}{2}$; then if $f(x)$ is continuous, the original value of f on the boundary is preserved. A function with these characteristics is

$$T(x) = \frac{1 - \tanh k_c x}{2}. \tag{8.108}$$

This function has a transition region of size $\Delta x \sim k_c^{-1}$, which is the shortest length-scale available for plasmons with a cut-off of k_c in reciprocal space.

Replacing Θ with T in equations (8.88) and (8.89) gives the corrected bulk plasmon formulae:

$$\chi^{\mathrm{s}}_{\mathrm{bulk}}(\mathbf{r}) = \frac{e^2}{\hbar\omega_p\epsilon_0}\sum_{k_z}\frac{4n_0}{k_z^3 s}\sin k_z z(1-\cos k_z s)T(z)T(s-z) \tag{8.109}$$

$$\begin{aligned} u^{\mathrm{s}}_{\mathrm{bulk}}(\mathbf{r},\mathbf{r}') = {} & \frac{e^2}{\hbar\omega_p\epsilon_0 V}\sum_{\mathbf{k}}\frac{4}{k^2}\cos\mathbf{k}_\parallel\cdot(\mathbf{r}_\parallel-\mathbf{r}'_\parallel)\sin k_z z\sin k_z z' \\ & \times T(z)T(z')T(s-z)T(s-z'). \end{aligned} \tag{8.110}$$

The cusps of the surface plasmon two-body term are contained in the function $F_{k_\parallel}(z,z')$ (given in table 8.1). The smooth version of $F_{k_\parallel}$ is

$$\begin{aligned} F^{\mathrm{s}}_{k_\parallel}(z,z') = {} & F_{k_\parallel}(z<0,z'<0)T(z)T(z') \\ & + F_{k_\parallel}(z<0,0<z'<s)T(z)T(-z')T(z'-s) \\ & + F_{k_\parallel}(z<0,z'>s)T(z)T(s-z') \\ & + \cdots . \end{aligned} \tag{8.111}$$

Replacing $F_{k_\parallel}$ with $F^{\mathrm{s}}_{k_\parallel}$ in equation (8.90) renders u_{surf} cusp-free.

Figures 8.10 and 8.11 illustrate the effect of removing the cusps from χ_{bulk} and u_{pl}; some detail is lost when the electrons are close to the slab edges.

8.3.2 Applying the electron-electron cusps

Having removed the undesirable cusps in the plasmon wave function, the next step is to insert the desirable ones! This proceeds as indicated in equation (8.104):

$$\Psi = \exp\left[-\frac{1}{2}\sum_{i\neq j}u_{\mathrm{cusp}}(\mathbf{x}_i,\mathbf{x}_j) - \frac{1}{2}\sum_{i,j}u^{\mathrm{s}}_{\mathrm{pl}}(\mathbf{r}_i,\mathbf{r}_j) + \sum_i \chi^{\mathrm{s}}_{\mathrm{bulk}}(\mathbf{r}_i)\right]D^{\uparrow}(\mathbf{R}^{\uparrow})D^{\downarrow}(\mathbf{R}^{\downarrow}). \tag{8.112}$$

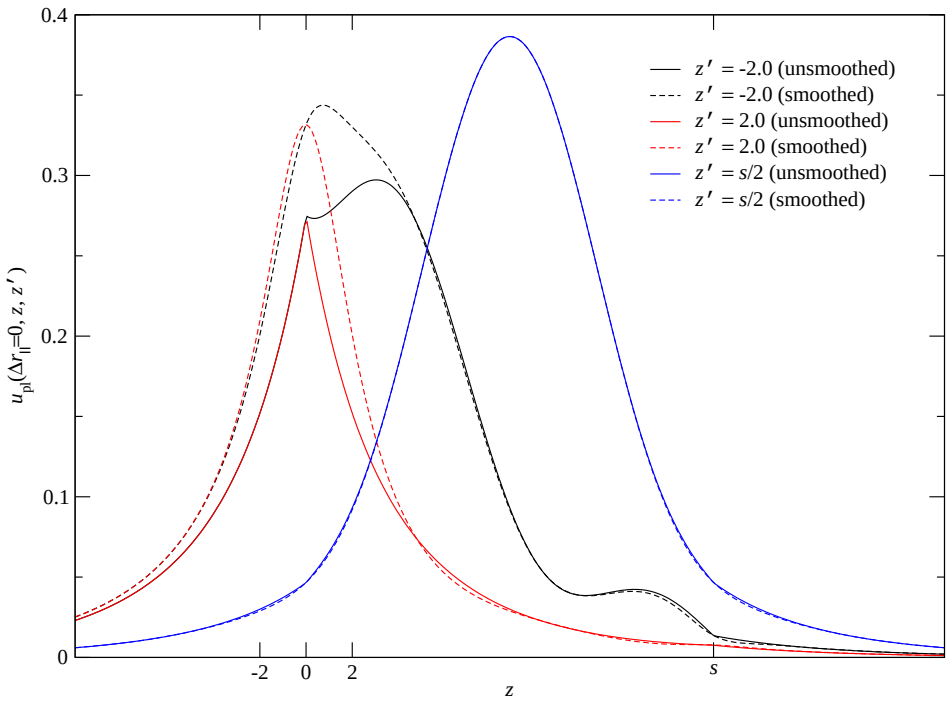


Figure 8.10: Removing the cusps from u_{pl} at the slab boundaries. In each plot, one electron is fixed while the other is scanned along a line in the z-direction; the x- and y-coordinates are chosen to be the same, so that $\Delta r_{\|} = 0$. The smoothed and unsmoothed curves are almost indistinguishable when the electrons are far from the slab edges; the most pronounced difference appears when one electron is just inside the slab edge (the black curves). This and the following plots were calculated for a cell containing 600 electrons, with $s = 17.64248$ and $r_s = 2.07$ (in Hartree atomic units).

The free parameter $\alpha_{\sigma_i\sigma_j}$ in the electron-electron cusp function u_{cusp} determines the distance at which the short-range cusp-dominated behaviour is replaced by the long-range plasmonic form. Following the discussion in the previous section, this distance should be $\sim k_c^{-1}$, which means that $\alpha_{\sigma_i\sigma_j} = k_c$. The remaining parameter, $\beta_{\sigma_i\sigma_j}$ is determined by equation (B.21).

There is one further subtlety related to the use of periodic boundary conditions. The plasmon two-body function, by construction, is periodic in the xy-plane; however, the cusp function is not. It is therefore important to ensure that u_{cusp} is effectively zero before the electron-electron separation reaches a maximum;[7] not

[7]The electron-electron distance has a maximum because all electron-electron interactions (in-

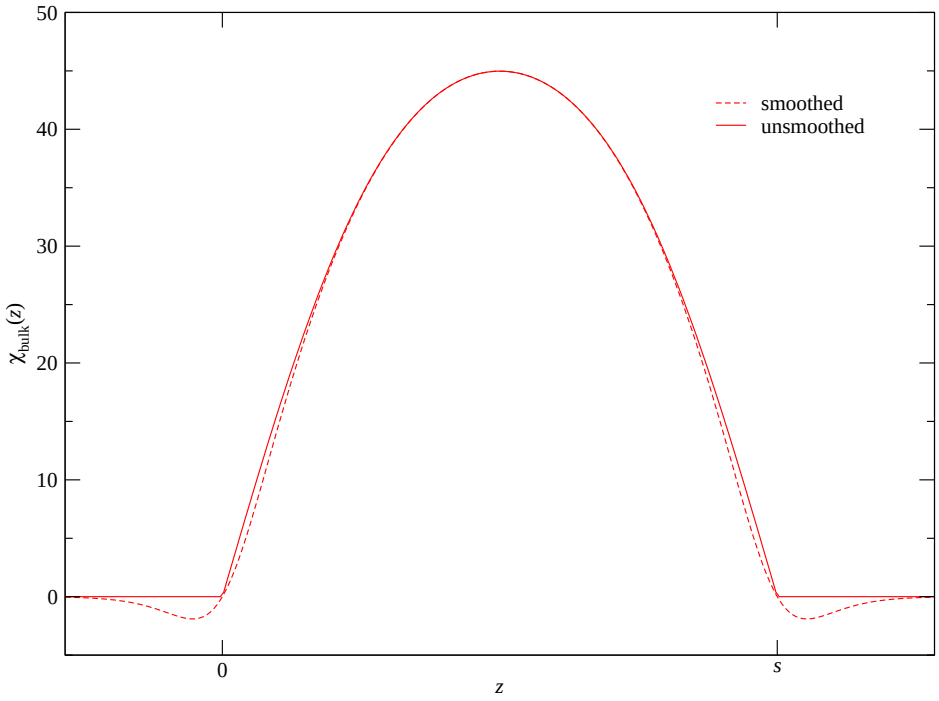


Figure 8.11: Removing the cusps from χ_{bulk} at the slab boundaries; $\chi^{\text{s}}_{\text{bulk}}$ is compared with χ_{bulk}.

taking this precaution would risk the introduction of yet more unwanted cusps in the wave function. A factor of $e^{-r_{ij}^2/L_c^2}$ is included for this purpose. The cut-off distance L_c should be larger than k_c^{-1}, to avoid interfering with the short-range behaviour, but significantly less than the Wigner-Seitz radius of the cell. The final form of the cusp function is

$$u_{\text{cusp}}(\mathbf{x}_i, \mathbf{x}_j) = \left(\frac{m_e e^2}{4\pi\epsilon_0\hbar^2}\right)\frac{1}{2k_c}\left(\frac{1}{1+\delta_{\sigma_i\sigma_j}}\right)e^{-k_c r_{ij} - r_{ij}^2/L_c^2}. \tag{8.113}$$

The full two-body function is plotted in figure 8.12, along with the usual homogeneous form (equation (3.79)) for comparison.

The purpose of the two-body terms is to incorporate correlations into the wave function: principally to keep electrons apart. However, a secondary effect (in non-homogeneous systems) is to alter the electron density. In the case of the slab, using only a u term forces electrons away from the centre, towards the slab edges;

cluding correlations) are treated in the minimum-image scheme described in chapter 6.

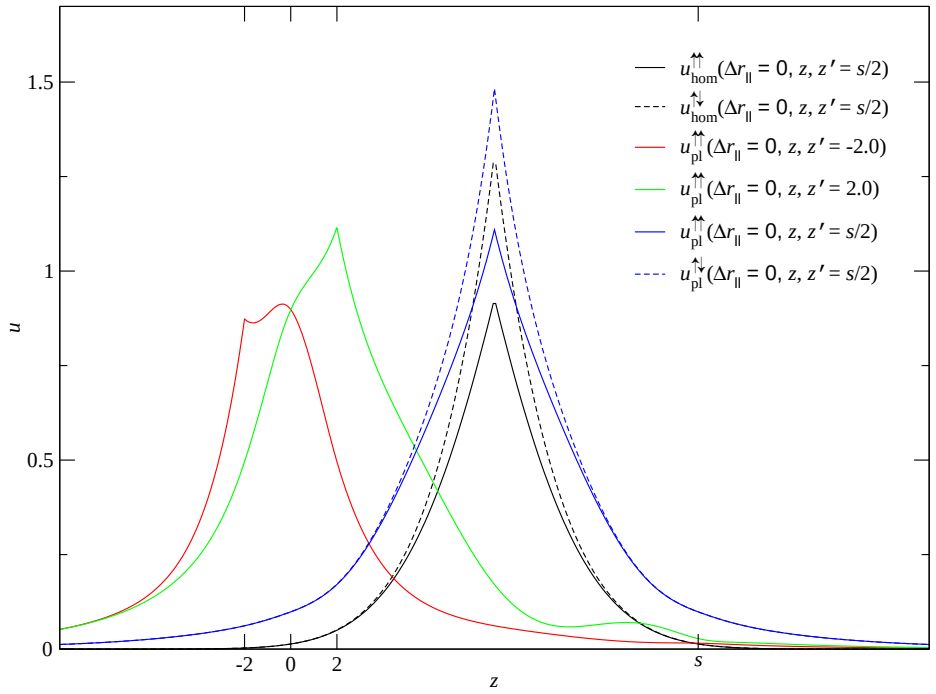


Figure 8.12: The smoothed, cusp-corrected plasmon two-body function. The homogeneous curve is plotted for comparison.

the density spreads out more. This effect is undesirable, because the initial two-determinant wave function usually gives an accurate estimate of the density. The χ function serves to correct this problem, restoring the original density.

The plasmon χ function should be expected to correct the density changes created by $u_{\rm pl}$. It should not be expected to account for the additional effect of using $u_{\rm cusp}$; thus, it is necessary to make a final modification to the wave function.

Returning to equation (8.87), it is evident that the plasmon one-body function has the form

$$\chi_{\rm bulk}(\mathbf{r}) = \int_V u_{\rm bulk}(\mathbf{r}, \mathbf{r}')\bar{n}(z')\, d^3r'. \tag{8.114}$$

To within a factor of $(2N-1)/2N$, this is in agreement with the findings of Malatesta et al. [56], which are based on a plausibility argument. The difference arises because in that work, the terms with $i = j$ in the double sums used to define u are neglected; in the explicit plasmon formulation used here and in the work of Gaudoin et al. [28], it is more natural to include these terms.

An improved one-body function may be obtained by replacing u_{bulk} in equation (8.114) with the full two-body term, giving

$$\chi(\mathbf{r}) = \chi_{\text{bulk}}(\mathbf{r}) + \chi_{\text{cusp}}(\mathbf{r}) \tag{8.115}$$

where

$$\begin{aligned}\chi_{\text{cusp}}(\mathbf{r}) &= \frac{1}{2}\int_V \Big[u_{\text{cusp}}(\mathbf{r}\uparrow, \mathbf{r}'\uparrow) + u_{\text{cusp}}(\mathbf{r}\uparrow, \mathbf{r}'\downarrow)\Big]\bar{n}(z')\; d^3r'. \\ &= \left(\frac{m_e e^2}{4\pi\epsilon_0\hbar^2}\right)\frac{3}{8k_c}\int_V \bar{n}(z') e^{-k_c|\mathbf{r}-\mathbf{r}'|-(\mathbf{r}-\mathbf{r}')^2/L_c^2}\; d^3r'.\end{aligned} \tag{8.116}$$

This step is not rigorously justified; the relationship (8.114) has only been shown to apply to the plasmon part of the Jastrow factor. However, this relationship has proven to be accurate in the past. The surface plasmon term u_{surf} is not included; the in-plane integral is zero[8] because there is no surface plasmon with $k_\parallel = 0$.

The integral in equation (8.116) is over the cell. However, the factor of $e^{-(\mathbf{r}-\mathbf{r}')^2/L_c^2}$ in u_{cusp} is designed to ensure that u_{cusp} becomes zero well before $|\mathbf{r}-\mathbf{r}'|$ approaches the size of the cell. Conveniently, this means that the integral may equally well be evaluated over the entire xy-plane. Switching to cylindrical polar coordinates gives

$$\begin{aligned}\int_V \bar{n}(z') e^{-k_c|\mathbf{r}-\mathbf{r}'|-(\mathbf{r}-\mathbf{r}')^2/L_c^2}\; d^3r' = 2\pi\int_{z'=-\infty}^{\infty}\int_{\rho'=0}^{\infty}\bar{n}(z')\exp\Bigg(-k_c\sqrt{\rho'^2+(z-z')^2} \\ - \frac{\rho'^2+(z-z')^2}{L_c^2}\Bigg)\rho'\; d\rho'\; dz'.\end{aligned} \tag{8.117}$$

There is no dependence on the in-plane components of $\mathbf{r}$.

The integral is solved in appendix C. Without specifying the form of $\bar{n}$, the result is

$$\begin{aligned}\chi(\mathbf{r}) = \chi_{\text{bulk}}(\mathbf{r}) + \left(\frac{m_e e^2}{4\pi\epsilon_0\hbar^2}\right)\frac{3\pi L_c^2}{8k_c}e^{k_c^2L_c^2/4}\int_{z=-\infty}^{\infty}\bar{n}(z')\Bigg\{\exp\left(-\left[\frac{|z-z'|}{L_c}+\frac{k_cL_c}{2}\right]^2\right) \\ - \frac{k_cL_c\sqrt{\pi}}{2}\operatorname{erfc}\left(\frac{|z-z'|}{L_c}+\frac{k_cL_c}{2}\right)\Bigg\}\; dz'.\end{aligned} \tag{8.118}$$

[8] See equation (8.92).

Making the approximation (once again) that $\bar{n}$ is constant within the slab then gives

$$\begin{aligned}\chi(\mathbf{r}) = \chi_{\text{bulk}}(\mathbf{r}) + \left(\frac{m_e e^2}{4\pi\epsilon_0\hbar^2}\right)\frac{3}{16}n_0 L_c^4 \pi\sqrt{\pi}e^{k_c^2 L_c^2/4} \\ \times\Bigg\{-\left(\frac{z}{L_c}+\frac{k_cL_c}{2}+\frac{1}{k_cL_c}\right)\text{erfc}\left(\frac{z}{L_c}+\frac{k_cL_c}{2}\right) \\ -\left(\frac{s-z}{L_c}+\frac{k_cL_c}{2}+\frac{1}{k_cL_c}\right)\text{erfc}\left(\frac{s-z}{L_c}+\frac{k_cL_c}{2}\right) \\ +\left(\frac{z-a(z)}{L_c}+\frac{k_cL_c}{2}+\frac{1}{k_cL_c}\right)\text{erfc}\left(\frac{z-a(z)}{L_c}+\frac{k_cL_c}{2}\right) \\ +\left(\frac{a(z)-z}{L_c}+\frac{k_cL_c}{2}+\frac{1}{k_cL_c}\right)\text{erfc}\left(\frac{a(z)-z}{L_c}+\frac{k_cL_c}{2}\right) \\ +\frac{1}{\sqrt{\pi}}e^{-(z/L_c+k_cL_c/2)^2}+\frac{1}{\sqrt{\pi}}e^{-[(s-z)/L_c+k_cL_c/2]^2} \\ -\frac{1}{\sqrt{\pi}}e^{-([z-a(z)]/L_c+k_cL_c/2)^2}-\frac{1}{\sqrt{\pi}}e^{-([a(z)-z]/L_c+k_cL_c/2)^2}\Bigg\}\end{aligned} \tag{8.119}$$

where

$$a(z) = \begin{cases} 0 & \text{when } z < 0 \\ z & \text{when } 0 < z < s \\ s & \text{otherwise.}\end{cases} \tag{8.120}$$

This version of the modification to χ is plotted in figure 8.13; the relative size of the correction is small ($\sim 1\%$ in the centre of the slab). Fortunately, the correction is cuspless. The modification obtained by using the true density and integrating equation (8.118) is very similar to that shown in figure 8.13 for the unbounded[9] slab; the difference is slightly more noticeable for the bounded slab.

The sum containing u_{pl} in equation (8.112) is unrestricted: it includes terms with $i = j$. It is perhaps clearer to organise one- and two-body functions separately, so that the full plasmon wave function (with appropriate cusps, and without

[9]See chapter 5 for a description of these two versions of the jellium slab.

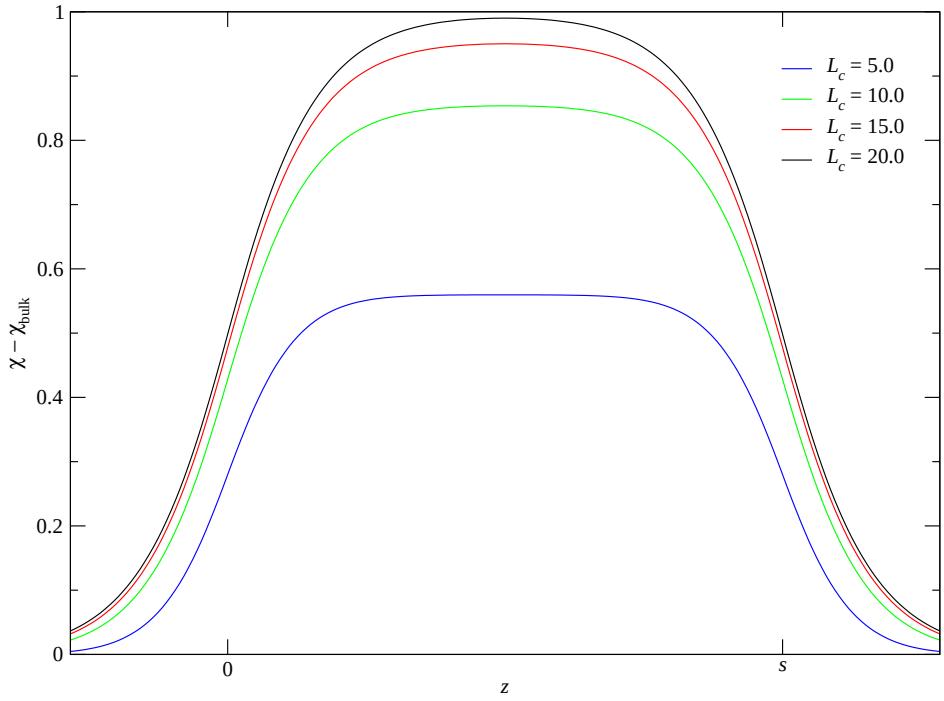


Figure 8.13: The one-body term χ_{cusp}, for various values of the cut-off distance L_c, calculated using equation (8.119); this equation was derived using the approximation of constant electron density within the slab. This term is designed to combat the density-changing effects of u_{cusp}. As L_c becomes large, the curves tend to a limit, because the decay of u_{cusp} is then dominated by k_c rather than L_c.

inappropriate ones) is

$$\Psi = \exp\Bigg[-\frac{1}{2}\sum_{i\neq j}\Big(u_{\text{cusp}}(\mathbf{x}_i,\mathbf{x}_j) + u^{\text{s}}_{\text{pl}}(\mathbf{r}_i,\mathbf{r}_j)\Big) + \sum_i\Big(\chi^{\text{s}}_{\text{bulk}}(\mathbf{r}_i) + \chi_{\text{cusp}}(\mathbf{r}_i) - \frac{1}{2}u^{\text{s}}_{\text{pl}}(\mathbf{r}_i,\mathbf{r}_i)\Big)\Bigg] D^{\uparrow}(\mathbf{R}^{\uparrow})D^{\downarrow}(\mathbf{R}^{\downarrow}). \tag{8.121}$$

8.4 Results

In order to test the plasmon-derived Jastrow factor, variational Monte Carlo simulations were carried out for the two versions of the jellium slab system described in chapter 5. For both systems, a cell containing 600 electrons was used, with the

slab width $s = 17.64248$ and density parameter $r_s = 2.07$.[10] At this electron density (which corresponds to aluminium), the plasma frequency is $\omega_p = 0.581574$ and the plasmon reciprocal-space cut-off is $k_c = 0.695048$. Using 600 electrons means that the length of side of the cell is $L = 35.5464$; in this relatively large cell, the long-range plasmon correlations should be important, whereas in a small cell with $L \sim k_c^{-1}$, the short-range behaviour would be expected to dominate. In addition, it means that $L_c = 5.5$, so that $L_c \gg k_c^{-1}$ and the short-range function u_{cusp} is allowed to decay naturally. As a further check, simulations were also carried out for an even larger system containing 1600 electrons.

The trial wave functions used in the test were of the standard two-determinant Slater-Jastrow form described in equation (3.77), with determinantal orbitals obtained from LDA calculations. In addition to the VMC energies, it is instructive to compare the electron density profiles generated by the different Jastrow factors:

$$n(z) = \int |\Psi(\mathbf{r}_1, \ldots, \mathbf{r}_N)|^2 \sum_{i=1}^{N} \delta(z - z_i)\; d^3r_1 \;\cdots\; d^3r_N. \tag{8.122}$$

During the VMC simulation, the z-positions of the electrons are sampled periodically. These coordinates are taken from the distribution with probability density function $n(z)/N$. To see this, consider $P(a < z_1 < b)$ (the probability that the first electron lies in a given z-range):

$$\begin{aligned} P(a < z_1 < b) &= \int_{z_1=a}^{b} \left(\int |\Psi(\mathbf{r}_1, \ldots, \mathbf{r}_N)|^2 \; dx_1 \; dy_1 \; d^3r_2 \;\cdots\; d^3r_N \right) dz_1 \\ &= \int_{z=a}^{b} \frac{n(z)}{N} \; dz. \end{aligned} \tag{8.123}$$

The problem of reconstructing the probability density function $n(z)/N$ from the set of sampled points $\{z_i\}$ is addressed in appendix D.

8.4.1 Unbounded slab

Figure 8.14 shows the density profile of the original two-determinant wave function. Because the orbitals used in the determinants were taken from an LDA calculation,

[10] For the rest of this chapter, all measurements will once again be quoted in atomic units.

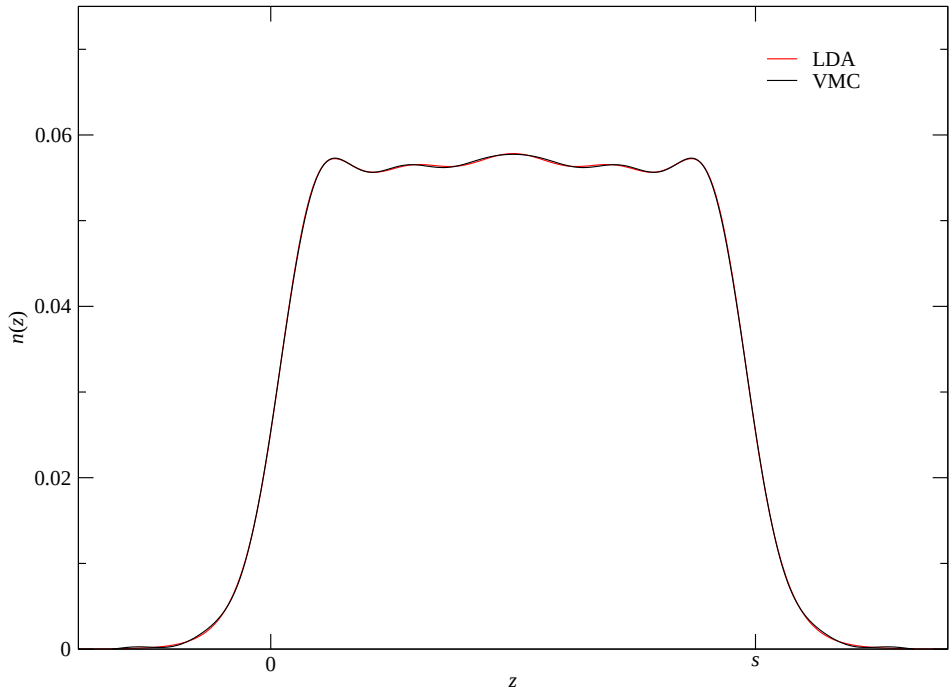


Figure 8.14: The electron density profile, from LDA and VMC. The wave function used in VMC was a product of determinants made up of LDA orbitals, with no Jastrow factor; the densities are the same, to within the VMC noise.

the profile should match the LDA density, which is also shown. Assuming that the LDA density is reasonably accurate, this provides a reference for the results to follow.

First, the effect of introducing only the homogeneous short-ranged two-body term u_{cusp} is investigated. Fahy and coworkers [21] were the first to point out in the context of quantum Monte Carlo simulations that adding a homogeneous[11] two-body term causes the electron density to become more uniform; this is seen in figure 8.15. However, the effect is small.

The change in electron density brought about by using the plasmon two-body term is much more dramatic, and is illustrated in figure 8.16. The long-range correlations cause the electron density to be pushed almost entirely into bands outside

[11]The term 'homogeneous' here refers to the fact that u only depends on the relative spin and separation of the electrons, and not on their individual positions.

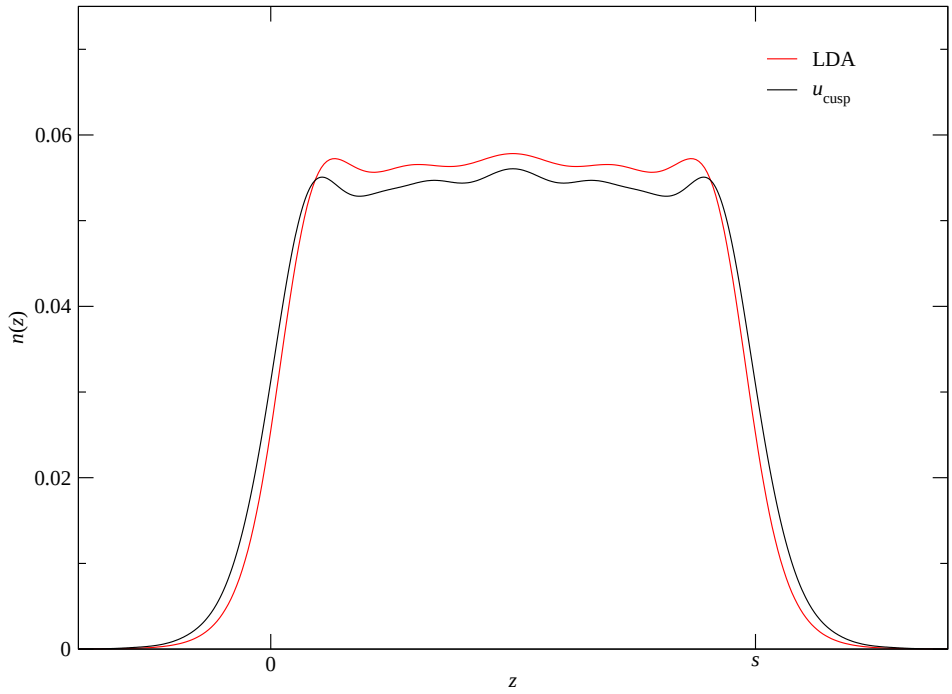


Figure 8.15: The effect on the electron density of including a homogeneous two-body term of the form given in equation (8.113) in the wave function. The density becomes slightly more homogeneous, with more electrons being pushed into the vacuum regions outside the slab.

the original slab. In these regions, u_{pl} is much weaker than in the centre.

The function of the one-body term is to restore the correct electron density. Remarkably, given the dramatic separation observed in figure 8.16, this is achieved by χ_{pl}; figure 8.17 shows the density profile for the full wave function given in equation (8.121).

An important test for the plasmonic wave function is to compare it with a wave function containing only the short-range electron-electron cusp terms, u_{cusp} and χ_{cusp}. The density is plotted in figure 8.18, from which it is evident that the original LDA density is almost exactly recovered. This is an indication that the expression (8.119) for χ_{cusp}, which was not rigorously derived but constructed based on a plausibility argument, works very well for this system.

The results for a larger system of 1600 electrons are almost identical to those for the 600-electron system; table 8.2 shows the effect of the different Jastrow factors

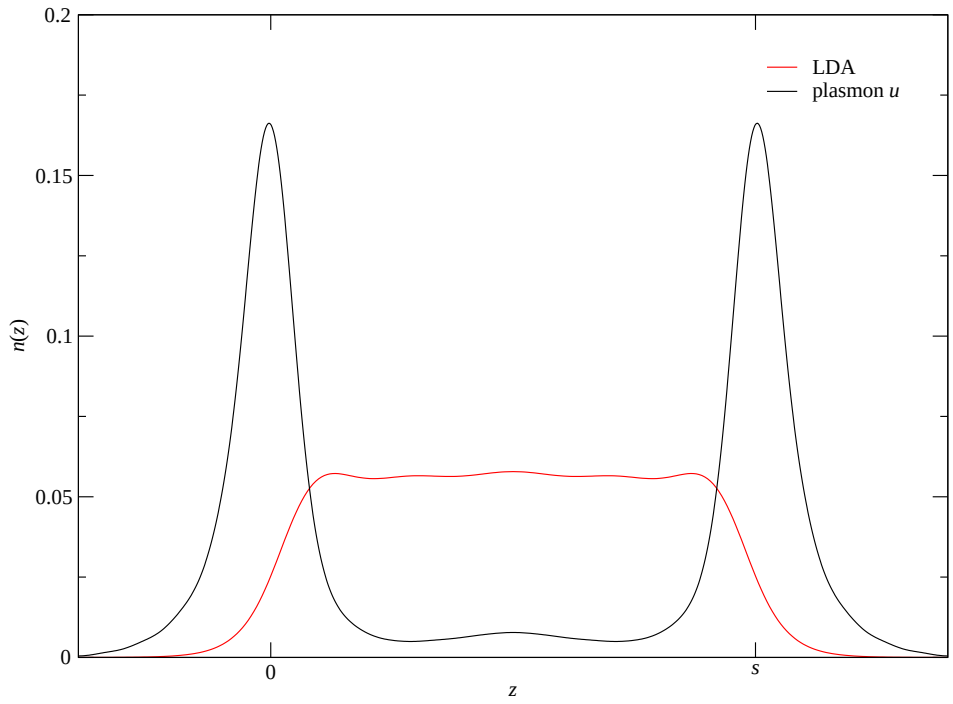


Figure 8.16: The electron density profile when the plasmon two-body term ($u_{\mathrm{cusp}} + u_{\mathrm{pl}}^{\mathrm{s}}$) is included in the VMC wave function. The profile is completely different to the original curve, with almost all the electrons now at the slab edges. The change is much more pronounced than when the homogeneous (and short-ranged) u_{cusp} alone was used.

on the VMC energy for the two sizes. The lowest energy is obtained when only the short-ranged part of the Jastrow factor is used. This is because, in this case, the electron density profile is very close to the (presumably optimal) original LDA profile. The short range of the two-body term means that it disrupts the density less, and there is consequently less work for the one-body function to do.

In contrast, the plasmon two-body function is very long-ranged, and has an enormous impact on the electron density (as seen in figure 8.16); the plasmon one-body function is correspondingly large. Although this function comes close to restoring the original density, it is not perfect: this may be due to the way in which the functions have been smoothed, or the fact that the Jastrow factor was derived for an idealised slab of constant density.

These results suggest that having the correct electron density is more important

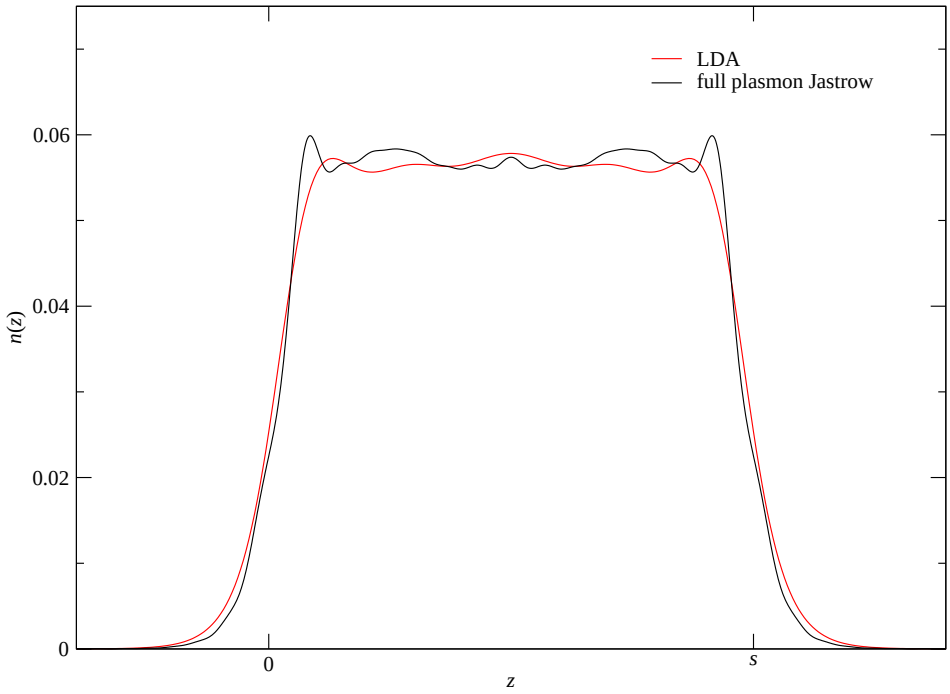


Figure 8.17: The density profile when the full plasmon Jastrow factor is included in the VMC wave function. The result is very close to the original density, although not perfect.

Number of electrons in cell	Jastrow factor used	Energy per electron (mHa)
600	none	32.5 ± 0.2
	u_{cusp}	12.5 ± 0.1
	$u_{\mathrm{pl}}^{\mathrm{s}} + u_{\mathrm{cusp}}$	3361 ± 3
	short-ranged Jastrow	6.14 ± 0.07
	full plasmon Jastrow	7.95 ± 0.06
1600	none	32.9 ± 0.2
	short-ranged Jastrow	4.8 ± 0.2
	full plasmon Jastrow	8.4 ± 0.4

Table 8.2: The energy per electron, calculated in VMC, for the unbounded jellium slab. Results for two different cell sizes and various forms of Jastrow factor are shown. The 'full plasmon Jastrow' refers to a wave function of the type described in equation (8.121); the 'short-ranged Jastrow' uses only u_{cusp} and χ_{cusp}.

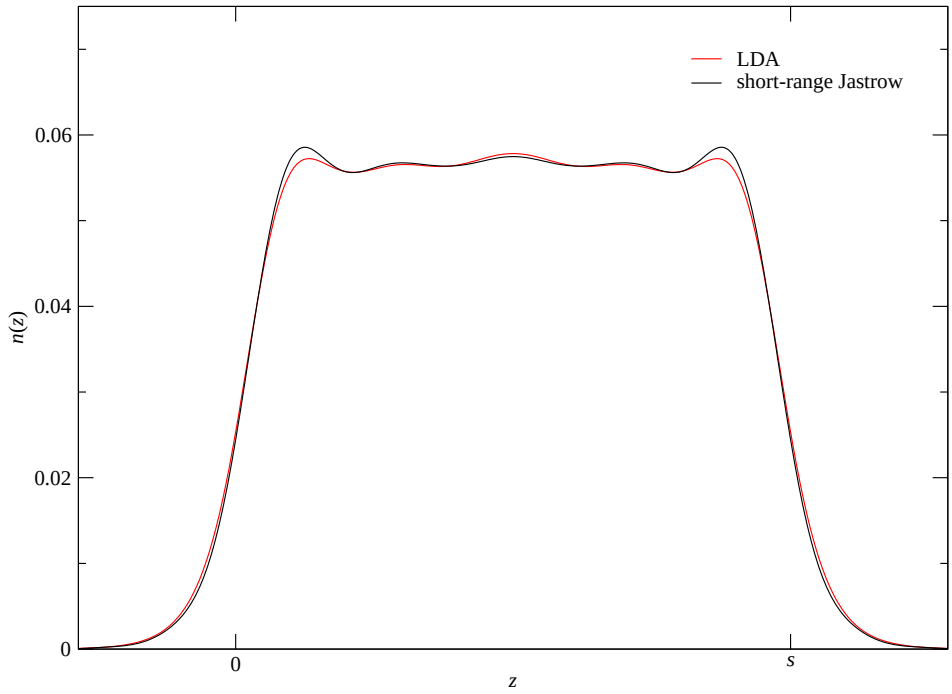


Figure 8.18: The electron density profile for a Jastrow factor consisting of only the short-range two-body function u_{cusp} and the corresponding one-body term χ_{cusp}.

than having the precise form of u. Up to this point, no optimisation has been carried out: all parameters have been calculated in advance, based only on theoretical considerations. To improve the plasmon Jastrow factor, a reasonable approach is to take the wave function of equation (8.121) as a starting point, and then to add a small one-body term with adjustable parameters, with the aim of restoring the initial density.

8.4.2 Bounded slab

Although the plasmon Jastrow factor for the unbounded slab performed well, and should provide a good starting point for optimisation, the unoptimised form did not improve on the simple short-ranged Jastrow factor, which came closer to maintaining the LDA density and hence generated a lower energy.

The bounded jellium slab is in some ways a more accurate representation of

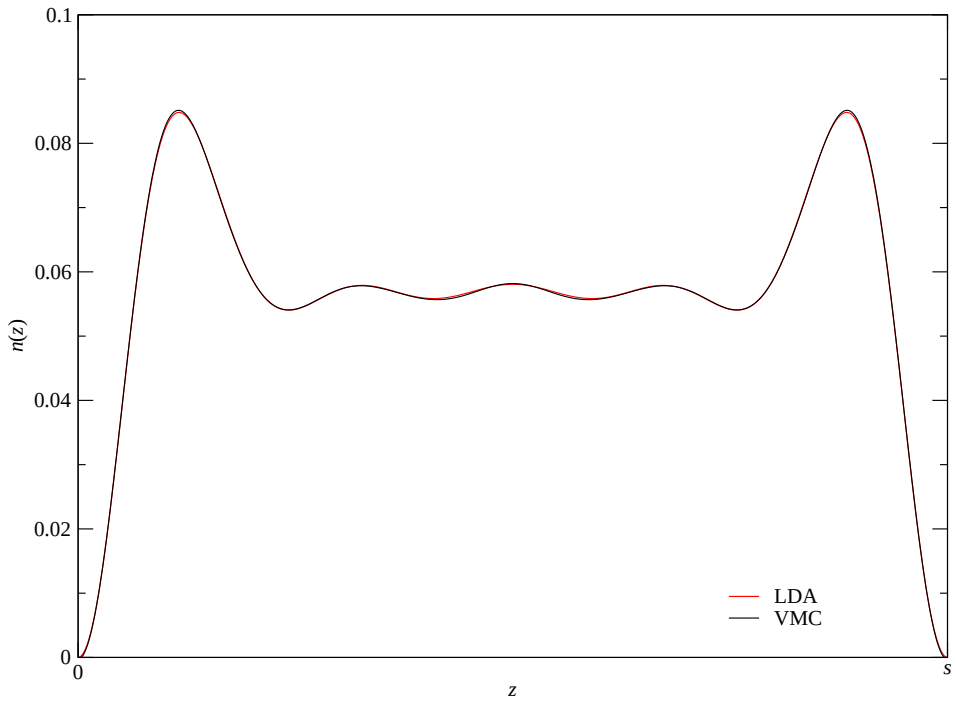


Figure 8.19: Electron density profile for the bounded slab, in LDA and VMC (with no Jastrow factor).

the system for which the plasmon normal modes were derived; electrons are truly confined to the slab, as in the original model. In the bounded jellium slab, it is not necessary to smooth out the cusps at the slab edges, because the determinantal part of the wave function already goes to zero here. The smoothing is presumably one of the areas which contribute to the small but significant errors in the Jastrow factor for the unbounded slab; the fact that it is not required for the bounded slab suggests that the results for the plasmon Jastrow factor should be better here.

Figures 8.19, 8.20, 8.21, 8.22, and 8.23 illustrate the electron density profiles for the various different versions of Jastrow factor; the corresponding energies are recorded in table 8.3.

The effect of including only the short-range two-body term (figure 8.20) is to reduce the energy and move the electron density away from the centre of the slab; as in the unbounded slab, the change in the electron density is small.

Even with the electrons confined to the slab, the plasmon two-body term com-

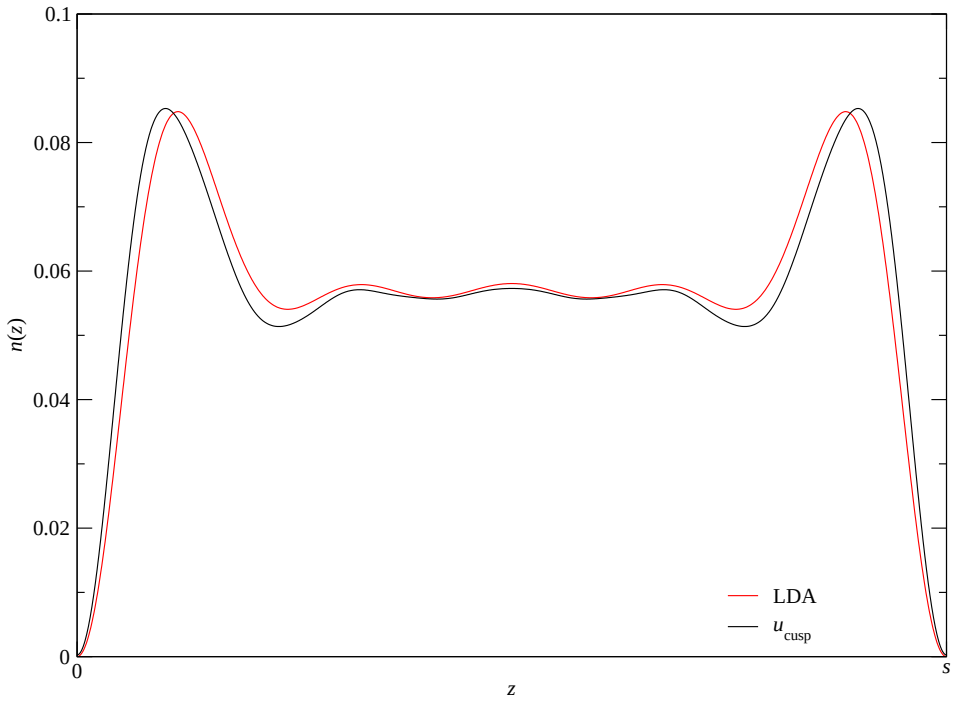


Figure 8.20: Electron density profile for the bounded slab, with a Jastrow factor containing only the short-ranged two-body term. The disturbance to the density is small; this is reflected in the VMC energy, which is significantly lower than when using no Jastrow factor.

Number of electrons in cell	Jastrow factor used	Energy per electron (mHa)
600	none	85.9 ± 0.4
	u_{cusp}	62.2 ± 0.3
	$u^{\mathrm{s}}_{\mathrm{pl}} + u_{\mathrm{cusp}}$	8870 ± 70
	short-ranged Jastrow	58.8 ± 0.2
	full plasmon Jastrow	55.1 ± 0.1
1600	none	87.9 ± 1.2
	short-ranged Jastrow	60.1 ± 0.9
	full plasmon Jastrow	53.3 ± 0.8

Table 8.3: The VMC energy per electron for the bounded jellium slab.

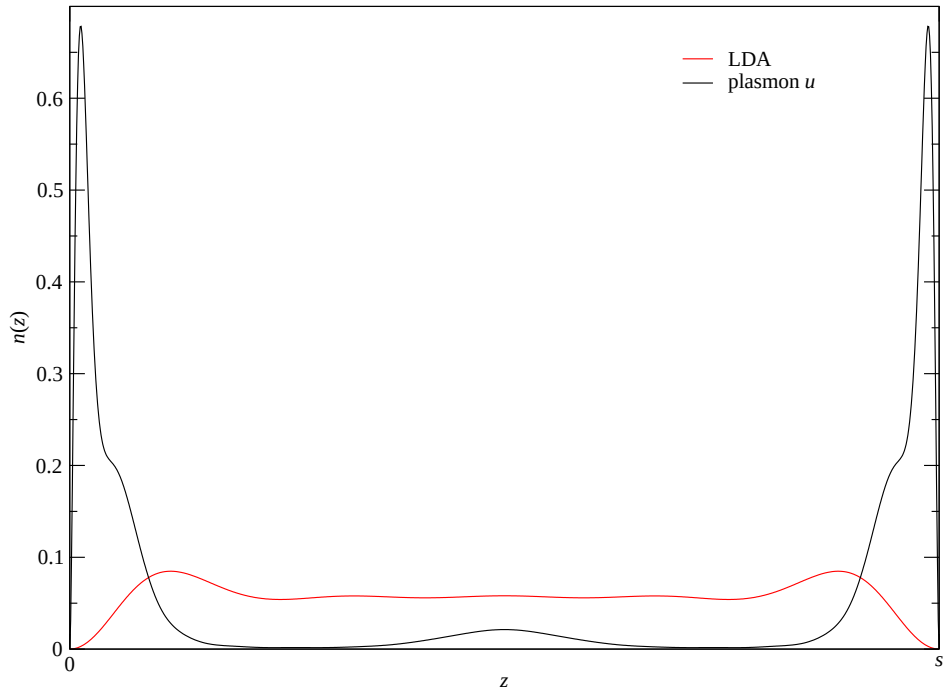


Figure 8.21: Electron density profile for the bounded slab, with the plasmon two-body term only. The long-ranged correlations alter the density drastically, as in the unbounded slab; electrons are pushed to the slab edges.

pletely changes the electron density when used without the corresponding one-body function; this is illustrated in figure 8.21. When the one-body function is applied, the density is very close to the LDA result, as can be seen in figure 8.22. The correction is better here than in the unbounded slab. Consequently, the energy is lower than that achieved with the full short-ranged Jastrow factor; the density for this configuration, shown in figure 8.23, is as good as that for the plasmon Jastrow factor, but the energy is higher.

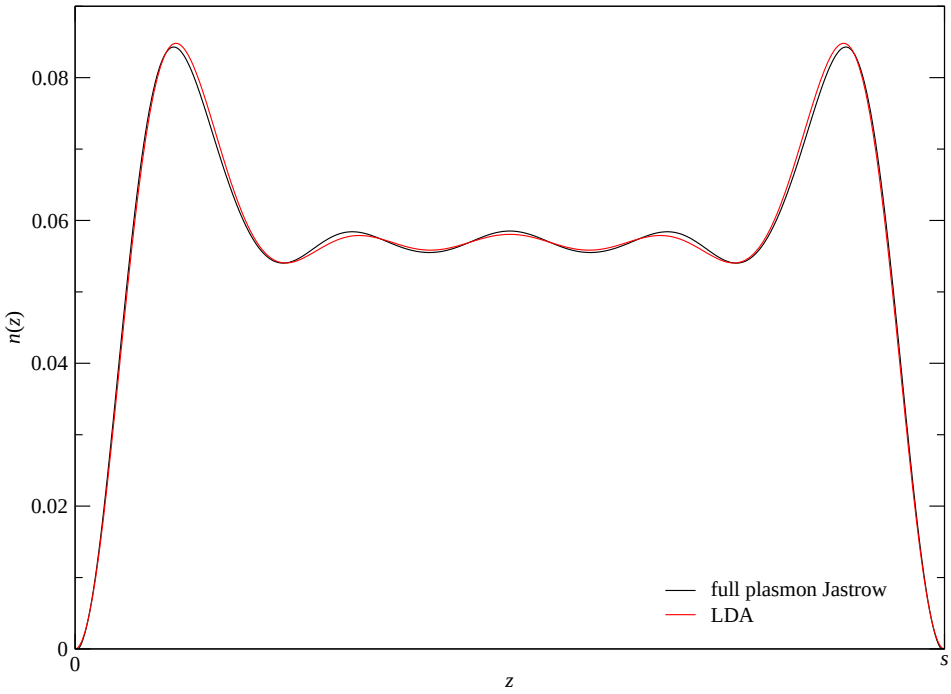


Figure 8.22: Electron density profile for the bounded slab, with the full plasmon Jastrow factor. The density is very close to the LDA form, and the energy is lower than when using only the short-ranged Jastrow.

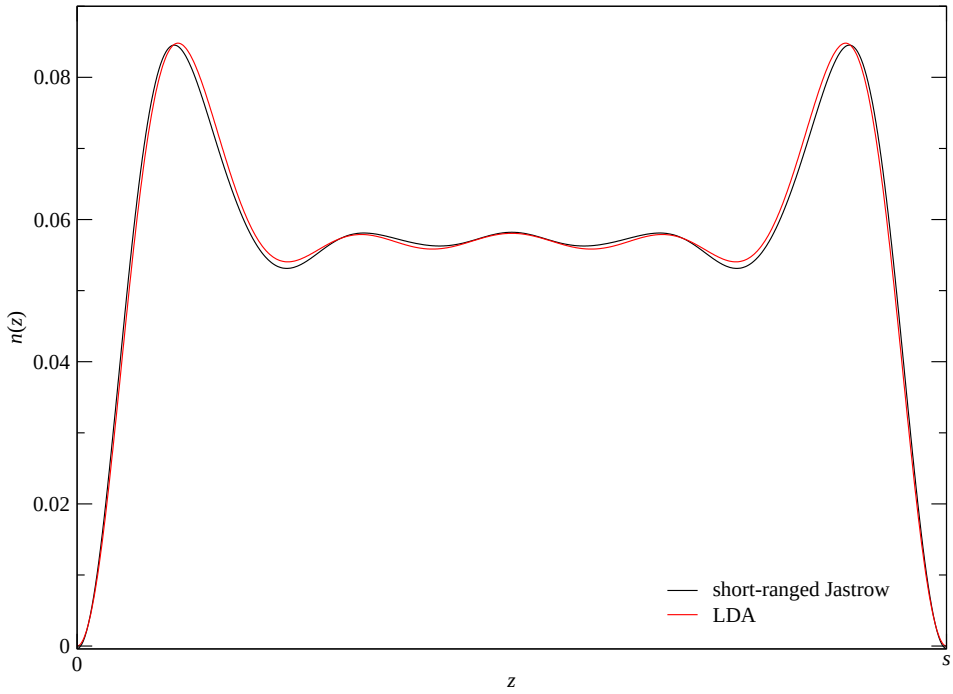


Figure 8.23: Electron density profile for the bounded slab, with the full short-ranged Jastrow factor. The LDA density is almost restored, and the energy is lower than when using only u_{cusp}, though not as low as that obtained with the full plasmon Jastrow factor.

Chapter 9

A new calculation of the jellium surface energy

The stated aim of this thesis is to investigate methods of improving the accuracy of surface calculations, with particular reference to quantum Monte Carlo simulations. The standard test system for surface calculations is the unbounded jellium slab, described in chapter 5 and referred to throughout this work; the disagreement between existing QMC results for this system and those obtained using other methods was discussed in chapter 5.

The techniques investigated here should make it possible to carry out a more accurate calculation of the jellium surface energy. The MPC interaction, while not reducing finite-size errors, makes simulations much more efficient; the plasmon normal mode theory, which produced a Jastrow factor more appropriate to the bounded jellium slab, nevertheless also led to the discovery of the efficient short-ranged two-body term with the accompanying analytic one-body term, helping to overcome the problems of trial wave function optimisation.

However, there are other avenues to explore in the attempt to achieve a more accurate simulation. Two of these will be described in this chapter; finally, a new calculation of the surface energy for one particular density will be presented.

9.1 Improved orbitals

The single-electron orbitals which make up the Slater determinants in the QMC trial wave function must be obtained from some prior calculation. Typically, this is an LDA calculation; the earlier QMC simulations of the jellium slab relied on LDA wave functions.

As described in chapter 2, a key component of any density-functional theory calculation is the exchange-correlation potential V_{XC}. It was shown by Lang and Kohn [48] that outside a metal surface, the correct asymptotic form of the potential is image-like:

$$V_{\mathrm{XC}}(z) = -\frac{1}{4(z - z_0)}. \tag{9.1}$$

Here z is the coordinate normal to the surface and z_0 is the position of the *image plane*. However, the LDA gives a potential which decays exponentially. Recently, Eguiluz and coworkers investigated the form of V_{XC} at a metal surface from first principles [19, 20]; their potential reproduced the asymptotic form given in equation (9.1), and matched the conventional LDA value inside the metal.

Having the correct image tail in the potential is important for studying several processes relevant to experiment: Eguiluz cites binding energies and lifetimes of image-potential-bound surface states, tunnelling currents in the scanning-tunnelling microscope, and resonant-tunnelling rates for ion-surface collisions as examples. It is not clear whether it will prove equally important when calculating the ground-state energy in QMC.

To investigate this, density-functional theory calculations were carried out using a version of V_{XC} containing the image potential, with two different positions[1] for the image plane. The resulting wave functions were tested in VMC simulations, and compared with the traditional LDA wave functions.

Figure 9.1 shows the original and modified forms of the potential. In the vacuum

[1]The two image-plane positions (relative to the slab edge) were $z_0 = 0.72$ and $z_0 = 1.49$; these values were obtained by Eguiluz, the first by fitting to the image tail, the second from the linear response.

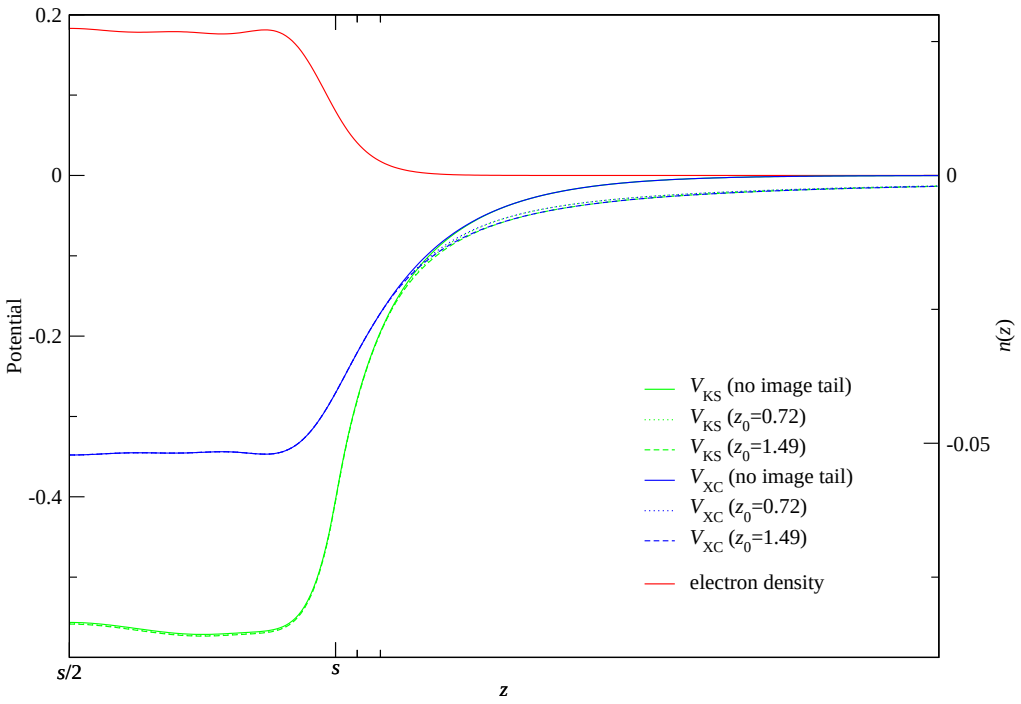


Figure 9.1: The effect of including the image tail in V_{XC}. The simple LDA form (shown for comparison) is matched to the image potential outside the slab. The test system has density parameter $r_s = 2.07$ and slab width $s = 17.64248$; the simulation cell contains 600 electrons. In addition to the exchange-correlation potential V_{XC}, the Kohn-Sham potential $V_{\mathrm{KS}} = V_{\mathrm{Ha}}+V_{\mathrm{B}}+V_{\mathrm{XC}}$ is also plotted. The positions of the image planes are indicated by ticks on the z-axis.

region, the LDA potential decays exponentially, while the modified potentials have the form described in equation (9.1). The electron density is plotted as a reference; on this scale, the difference between the three density curves corresponding to the different exchange-correlation potentials is not observable.

The effect on the single-electron orbitals is small; the greatest difference is observed in the higher sub-bands, because these extend further into the vacuum region. The highest occupied orbital for the current test system (the sixth sub-band) is plotted in figure 9.2. Note that there is a very small correction to the nodes.

The overall correction to the wave function is small, and this is reflected in the VMC energies listed in table 9.1. The wave functions calculated with the image plane at $z_0 = 1.49$ give a slightly higher energy, and will not be used in the QMC

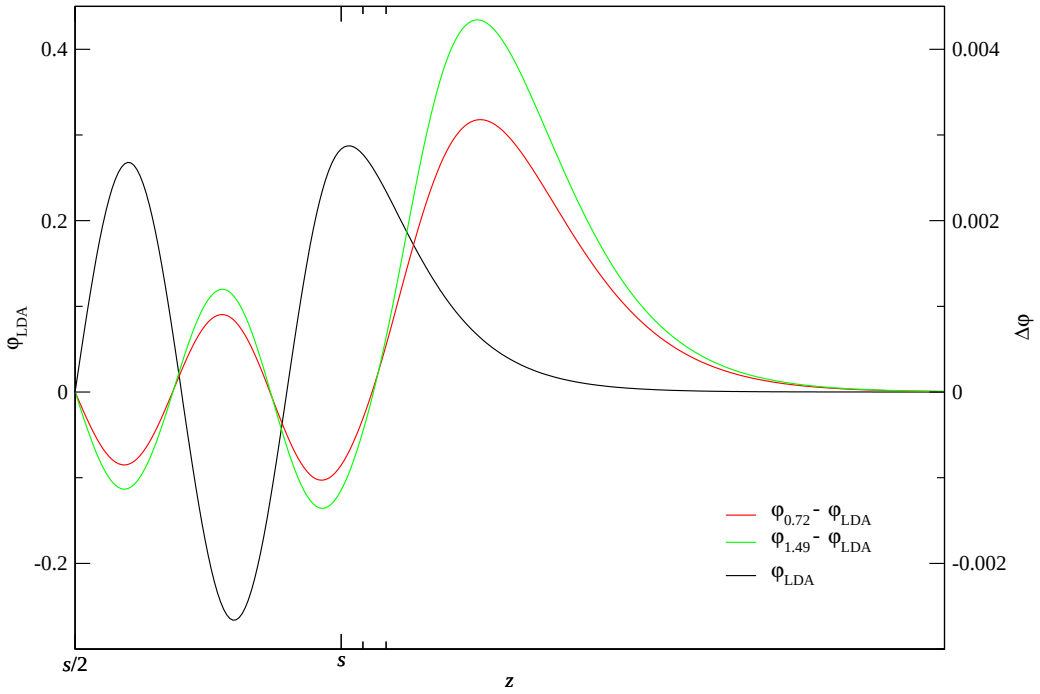


Figure 9.2: The effect of the image tail on the highest occupied single-electron sub-band. The original orbital is shown, along with the difference between this and the modified orbitals (magnified one hundred times - note the two scales).

simulations to follow. Those obtained with the image plane at $z = 0.72$ give the same VMC energy as the conventional LDA wave functions, although the variance appears to be slightly lower (less moves are required for a given accuracy); these wave functions will be used from now on.

9.2 Alternative k-point sampling

In all the simulations carried out so far, strictly periodic boundary conditions have been applied to the wave function in the xy-plane. The system is homogeneous in the xy-direction, which means that the DFT orbitals must have the following form:

$$\phi_{n\mathbf{k}_\parallel}(\mathbf{r}) = u_n(z)e^{i\mathbf{k}_\parallel \cdot \mathbf{r}_\parallel}. \tag{9.2}$$

Type of potential	VMC energy (mHa)
No image tail	32.54 ± 0.08
Image plane at $z_0 = 0.72$	32.54 ± 0.08
Image plane at $z_0 = 1.49$	32.70 ± 0.08

Table 9.1: Comparison of energies calculated with different versions of the exchange-correlation potential. The DFT energies are identical for all potentials. The trial wave function for the VMC calculations contained no Jastrow factor; the simulations used a simulation cell containing 600 electrons.

The simulation cell is square in the xy-plane, extending from $-\infty$ to ∞ in the z-direction. The boundary conditions therefore imply that

$$k_x = \frac{2\pi m_x}{L} \quad \text{where } m_x = 0, 1, 2, \ldots \tag{9.3}$$

with a corresponding relation for k_y.

However, this is not the only possible choice. The aim is to model an infinite slab, in which $\mathbf{k}_\parallel$ takes on a continuous range of values. Recent studies [74, 55] have used ideas from band-structure theory [6, 59] in QMC simulations to improve the extrapolation to the thermodynamic limit. The idea is to introduce some fixed phase shift across the simulation cell, so that

$$\phi(\mathbf{r} + \mathbf{L}_x) = e^{i\theta_x}\phi(\mathbf{r}) \tag{9.4}$$

where $\mathbf{L}_x$ is a vector of length L in the x-direction. Conventional periodic boundary conditions correspond to choosing $\theta_x = 0$. In the work of Lin et al. [55], an average over several different phase shifts is taken. Rajagopal and coworkers [74] found that using $\theta_x = \pi$ ('antiperiodic' boundary conditions) gave good results. Under these boundary conditions, equation (9.3) becomes

$$k_x = \frac{\pi}{L}(1 + 2m_x). \tag{9.5}$$

This demonstrates the advantage of using antiperiodic or periodic boundary conditions: it is always possible to construct real orbitals by taking pairs of $\mathbf{k}_\parallel$ and $-\mathbf{k}_\parallel$. This is not true for an arbitrary phase shift.

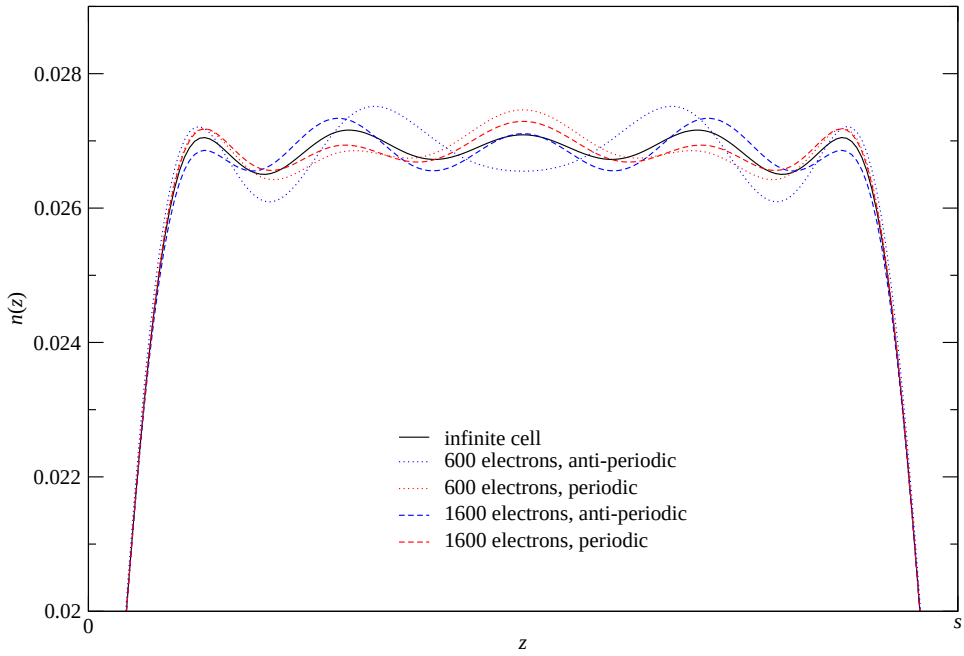


Figure 9.3: The electron density profile with periodic and antiperiodic boundary conditions for two different cell sizes. The infinite-cell limit is also shown (because this may be calculated in DFT).

In this section, the effect of imposing antiperiodic boundary conditions will be investigated.

Figure 9.3 shows that the electron density is altered quite dramatically by the new boundary conditions. The graph displays the densities for two different system sizes, illustrating that the effect of the boundary conditions is reduced as the cell becomes larger. Both finite-cell densities differ from the infinite-cell form by approximately the same amount, though in opposite directions.

Of all the sub-bands, only the lowest (u_0) is strongly affected by the altered boundary conditions; this is illustrated in figure 9.4. The change in the electron density is brought about by a significant change in the occupation of the sub-bands. The occupation numbers are listed in table 9.2; table 9.3 shows the fractional occupation for each system, along with the true value obtained in the infinite-cell limit. Despite the large change in the electron density, the effect on the ground-state energy in QMC is small. The VMC energies are listed in table 9.4. After the

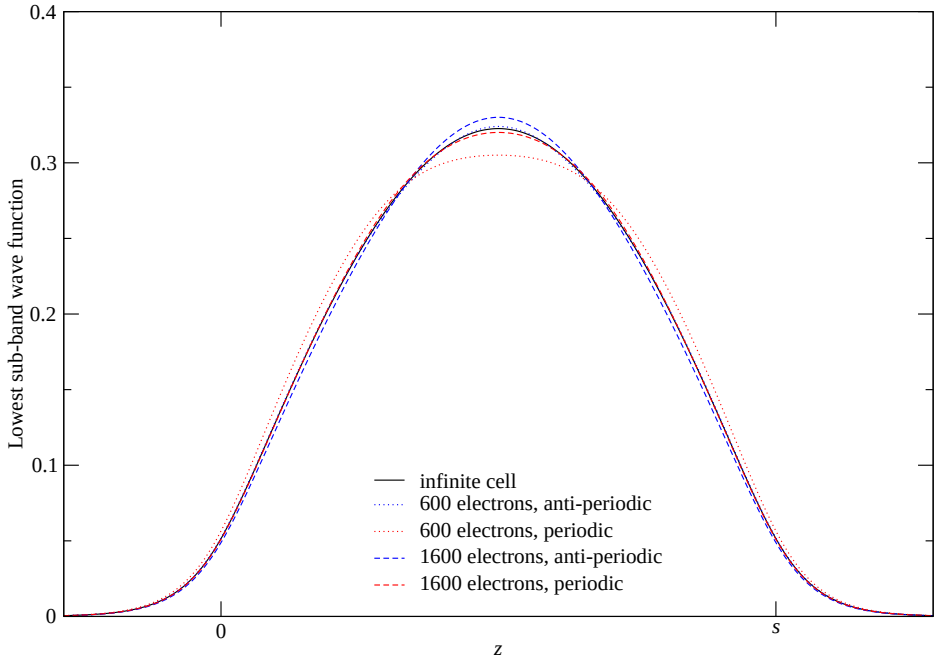


Figure 9.4: The lowest sub-band for different boundary conditions and cell sizes. The infinite-cell form is included for reference.

Number of electrons	Phase	Occupation of sub-band					
		1	2	3	4	5	6
600	0	178	138	134	90	50	10
600	π	168	152	120	104	48	8
1140	0	322	290	242	178	90	18
1140	π	312	296	240	176	104	12
1600	0	442	394	354	242	138	30
1600	π	432	416	328	240	152	32

Table 9.2: The occupation numbers of each sub-band for two cell sizes and periodic or anti-periodic boundary conditions.

Number of electrons	Phase	Fractional occupation of sub-band					
		1	2	3	4	5	6
600	0	0.297	0.23	0.223	0.15	0.083	0.017
600	π	0.28	0.253	0.2	0.173	0.08	0.013
1140	0	0.282	0.254	0.212	0.156	0.079	0.016
1140	π	0.274	0.260	0.211	0.154	0.091	0.011
1600	0	0.276	0.246	0.221	0.151	0.086	0.019
1600	π	0.27	0.26	0.205	0.15	0.095	0.02
∞	-	0.279	0.252	0.210	0.154	0.087	0.018

Table 9.3: The fractional occupation of each sub-band. This is simply the occupation number divided by the total number of electrons, and allows the results of finite-cell calculations to be compared with those obtained in the infinite-cell limit.

Number of electrons	Phase	VMC energy	DFT correction	Corrected energy
600	0	32.56 ± 0.11	-0.268	32.29
600	π	32.32 ± 0.11	-0.128	32.19
1140	0	32.75 ± 0.13	-0.117	32.63
1140	π	32.41 ± 0.14	0.210	32.62
1600	0	32.92 ± 0.14	-0.072	32.85
1600	π	32.73 ± 0.18	0.070	32.80

Table 9.4: Comparison of energies calculated with different versions of the exchange-correlation potential. All energies are in mHa. The trial wave function for the VMC calculations contained no Jastrow factor.

usual independent-particle finite-size error correction has been applied, the results are identical to within the statistical error: this is reassuring, but shows that using alternative boundary conditions cannot reduce the remaining finite-size error. The results are in agreement with the hole-squashing model proposed in chapter 6 to explain this error.

9.3 Surface energy calculations

On reviewing the earlier sections of this thesis, it is apparent that even with the many improvements introduced here, finite-size errors cannot be eliminated. In particular, the Coulomb finite-size error conjectured to be caused by 'squashing' of the exchange-correlation hole must still be dealt with. In order to minimise the effect of these errors on the calculation of the surface energy, great care must be taken with the extrapolation to infinite system size; in addition, it is unwise to compare results from slab and bulk systems.

This latter constraint means that simulations must be carried out for a range of slab widths; the bulk and surface energies may then be extracted by a fitting procedure, using equation (5.6), repeated here for convenience:

$$\epsilon_{\text{slab}} = \epsilon_{\text{bulk}} + \left(\frac{8\pi r_s^3 \sigma}{3}\right)\frac{1}{s}. \tag{9.6}$$

The motivation for this approach is that the errors arising in bulk and slab calculations may differ significantly; using only slab results should ensure a better cancellation of errors.

As was noted in chapter 5, the surface energy calculated using equation (9.6) displays oscillations of decreasing amplitude as the slab width is increased; the true surface energy is the limiting value of σ as $s \to \infty$. To minimise the errors introduced by these oscillations, it is possible to choose slab widths for which the LDA value of σ matches the extrapolated value: three such special slab widths are 11.7783, 15.1317 and 18.4851.

For each slab width, calculations must be performed for a range of cell sizes; the results can then be extrapolated to the limit of infinite cell size. By selecting sizes for which the LDA slab energy per electron corresponds to the infinite-cell value for the same slab width, the need to apply the independent-particle finite-size correction described in section 4.1.1 is obviated.

Each system to be simulated requires an optimised trial wave function. As usual, LDA orbitals (with the image-tail correction of section 9.1) make up the Slater

determinants; the form of the Jastrow factor was determined after consideration of the results of chapter 8. The short-ranged two-body term

$$u_{\text{cusp}}(\mathbf{x}_i, \mathbf{x}_j) = \frac{\alpha}{2(1+\delta_{\sigma_i\sigma_j})} e^{-r_{ij}/\alpha - r_{ij}^2/L_c^2} \tag{9.7}$$

has the advantage that the corresponding one-body function required to restore the original density is derived analytically: the one- and two-body terms may therefore be optimised by varying the single parameter α.[2] This speeds up the otherwise painfully slow manual optimisation process; the failure of automatic optimisation based on variance minimisation is documented in chapter 5. The evaluation of the Jastrow factor within a simulation is also inexpensive compared with the full plasmonic form.

9.3.1 Results

The first stage of the QMC simulations is the optimisation of the trial wave function. Short VMC calculations were carried out for different values of α; an example of the dependence of the VMC energy on α is displayed in figure 9.5. The optimal value of α, obtained by a quadratic fit to these curves, was found to increase slightly with the cell size L.

Having obtained optimised wave functions, the next step is to carry out longer VMC runs, the results of which are plotted in figure 9.6. Systems of size ranging from 300 up to 3000 electrons were simulated; the number of time steps required ranged from tens of thousands for the larger systems to around a million for the smaller ones. The MPC interaction was used for these calculations: testing at one slab width revealed that the quasi-2D Ewald and MPC interactions returned values which were in agreement with each other, confirming the results of chapter 6.

The two quasi-2D interactions are compared with the 3D Ewald sum in figure 9.7. The 3D Ewald sum gave slightly lower energies than the other interactions;

[2]Note that L_c is fixed by the requirement that $u_{\text{cusp}}(\mathbf{x}_i, \mathbf{x}_j)$ be effectively zero when $|\mathbf{x}_i - \mathbf{x}_j| \sim L$, and is not an optimiseable parameter.

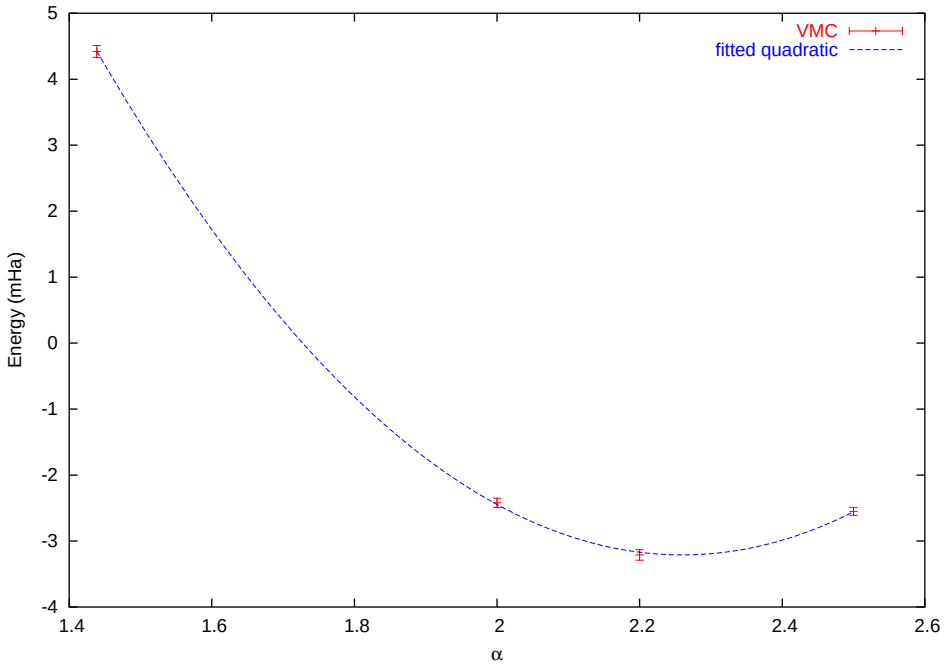


Figure 9.5: Optimising the parameter α. The VMC energies are shown along with the quadratic function fitted to them. These results are for a slab width of 11.7783 with a cell containing 588 electrons; the minimum is at $\alpha = 2.26$.

the difference has the correct sign to be associated with the additional Coulomb finite-size error arising from the periodicity in the z-direction.

More computationally-intensive DMC simulations were also carried out for some of the smaller systems: the results are listed in table 9.5.

9.3.2 Analysis

The best estimate of the jellium surface energy is obtained using figure 9.6. The graph illustrates that there is a significant finite-size effect: comparison with the DMC energies and the results of chapter 6 shows that this is due to the Jastrow factor. The Coulomb finite-size correction has a negative sign and is smaller in magnitude.

The optimised Jastrow factor becomes significantly better for larger system sizes; the axes of the graph have been chosen to demonstrate that the error scales approx-

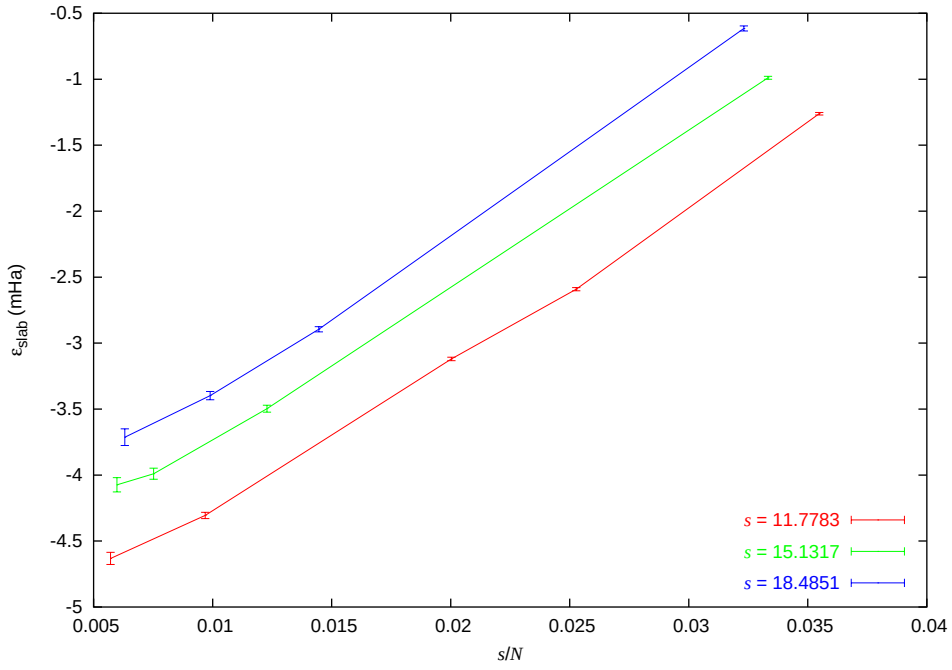


Figure 9.6: The slab energy per electron, plotted against s/N, where s is the slab width and N is the number of electrons. The abscissa is proportional to $1/L^2$, and was chosen to demonstrate the form of the finite-size errors.

imately as $1/L^2$. The reason that the Jastrow factor performs less well in small systems is almost certainly the parameter L_c. This parameter, which keeps the wave function free of unwanted cusps, depends on the cell size, and is not allowed to be optimised. It modifies the long-range behaviour of the two-body term; it does not appear to affect the electron density adversely (by rendering the derived one-body term less efficient).

Most importantly, this finite-size error is independent of the slab width: the evidence for this is the fact that the different curves in figure 9.6 remain equally-spaced. This allows an accurate calculation of the surface energy to be made.

First, the curves $s = 11.7783$ and $s = 15.1317$ are interpolated on the abscissae of the third curve, $s = 18.4851$. Working with the largest possible system size (and thus effectively eliminating the Coulomb finite-size error), the slab energies per electron at this point are (in mHa) -3.713 $\pm$ 0.063, -4.060 $\pm$ 0.051 and -4.593

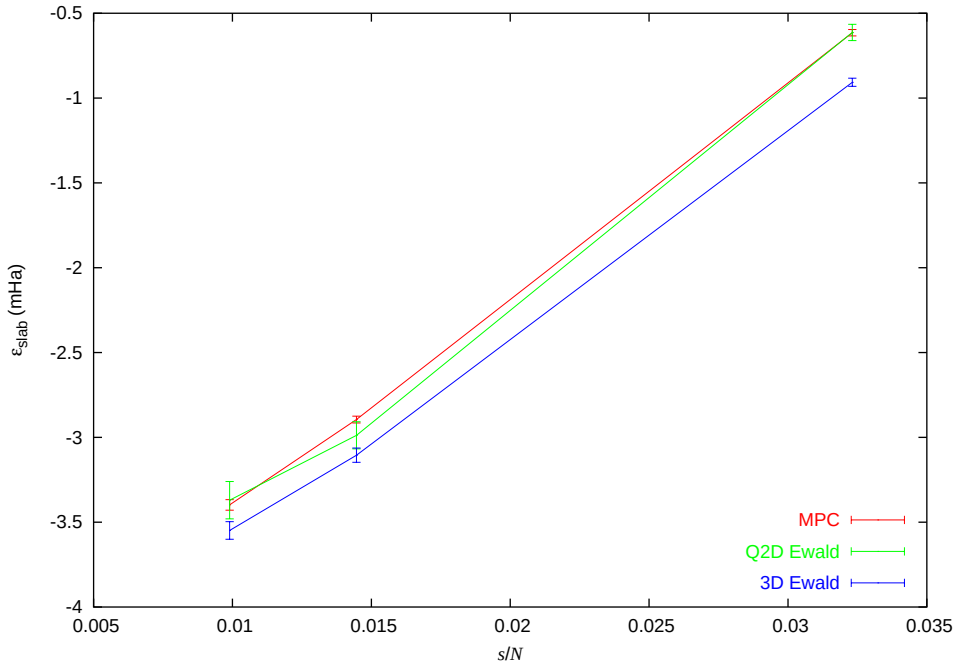


Figure 9.7: Comparing the MPC and Ewald interactions when $s = 18.4851$.

$\pm$ 0.042 respectively. From equation (9.6), these points should lie on a straight line when plotted against $1/s$; figure 9.8 shows that this is in fact the case. The gradient of the line passing through these points is -28.5 $\pm$ 2.3, which gives the surface energy of jellium as -0.384 $\pm$ 0.031 mHa bohr^{-2} or -600 $\pm$ 50 erg cm^{-2}. This figure brings the QMC result into line with the other calculations referred to in chapter 5.

In contrast, the values of σ quoted in table 9.5 were calculated using the method of Li et al. [53]: no Coulomb finite-size correction was applied, and the slab energy was simply combined with Ceperley and Alder's fixed-node value[3] for ϵ_{bulk} in equation (9.6). This demonstrates that, although the Coulomb finite-size error acts to decrease the energy (see the three calculations with $s = 11.7783$ in table 9.5), the greatest error is positive and comes from comparing the results of bulk and slab simulations. The quality of the nodal surface of the trial wave function is better in the homogeneous electron gas than in the jellium slab.

[3]The fixed-node bulk calculation of Ceperley and Alder referred to by Li is unpublished; the quoted energy per electron is $\epsilon_{\text{bulk}} = -0.2017$ eV per electron.

Slab width	Number of electrons	ϵ_{slab} (mHa)	σ (erg cm^{-2})
11.7783	332	-9.18 ± 0.13	-440 ± 30
11.7783	466	-8.901 ± 0.097	-370 ± 20
11.7783	588	-8.818 ± 0.088	-350 ± 20
15.1317	454	-8.486 ± 0.094	-340 ± 30
18.4851	572	-8.15 ± 0.15	-290 ± 60

Table 9.5: Slab energies per electron calculated in DMC, for various slab widths and system sizes. The surface energy estimates listed in the final column are included to demonstrate the errors introduced by combining the results of bulk and slab calculations without due care.

To carry out a more accurate DMC calculation of the surface energy using the fitting method applied above to the VMC results would require significant computational resources. The largest systems studied using VMC and used in the final surface energy calculations contained around 3000 electrons: in smaller systems, there remains the difficult problem of correcting the Coulomb finite-size error.

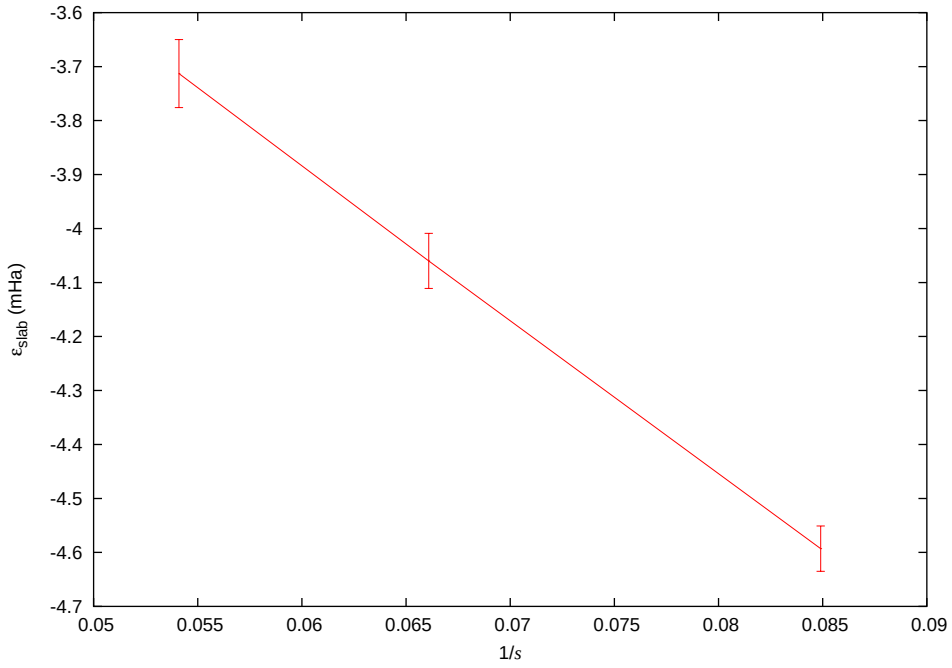


Figure 9.8: The slab energy per electron, with $L^2 = 1150$, as a function of the slab width.

Chapter 10

Conclusions

The original motivation for the work contained in this thesis was to resolve the controversy over the surface energy of jellium; the calculations contained in the final chapter suggest that this goal is close to being achieved. If the VMC results reported here are supported by DMC simulations, the extended-system QMC results will have been brought into line with those obtained using all other methods. This is a vital step if one is to have confidence in future QMC simulations of extended surface systems.

In working towards this goal, several techniques have been developed and investigated which may have applications in different areas. First of all, the MPC interaction has been introduced as a substitute for the Ewald interaction in quasi-2D simulations. Although it does not reduce the Coulomb finite-size errors in such systems, it is significantly faster than the Ewald interaction, and is therefore recommended for future work on slabs.

An explanation of the Coulomb finite-size error, and its resistance to the MPC, has also been provided. The model developed here leads to an error with the correct sign; the next step is to estimate the scaling of the finite-size error with system size, based on this model. An important lesson from this work is that the Coulomb finite-size errors in quasi-2D systems can be large, even for large cells, and need to be dealt with carefully.

A large part of this thesis is devoted to the connection between plasmons and

the electronic ground-state wave function. A new approach to this connection has been presented, based on physically intuitive ideas. The resulting prescription for the long-wavelength correlations in the wave function agrees with previous work; the value of the formalism presented here is that it is now much easier to incorporate prior knowledge about plasmon normal modes in any given system.

This idea has been tested on the jellium slab system, with some success: the plasmon-predicted Jastrow factor reduces the energy in VMC. However, the strong long-wavelength correlations disrupt the electron density significantly, and although the predicted one-electron term almost restores the original (and presumably correct) density, this restoration is not perfect. In contrast, the short-ranged correlation term developed here has a much weaker effect on the density, and the corresponding analytic one-electron term is able to restore the original density almost exactly. One important conclusion is that it appears to be more important to have the correct density than to include all the long-wavelength correlations, and this makes the short-ranged Jastrow factor more attractive.

Presumably, the best Jastrow factor would use both long- and short-ranged correlations, and take the predicted one-electron term as a starting point; adding some variational freedom to this term would then allow the correct density to be restored while maintaining all the correlations. However, the aim here was to minimise the need for optimisation, since this procedure is awkward for the jellium slab; therefore only the short-ranged correlation term (and the corresponding density-restoring function) were used in the calculations of the final chapter.

The density-restoring term corresponding to the short-ranged correlation function was derived analytically, using a relationship which has been postulated by other authors and which is automatically satisfied by the plasmon Jastrow factor. This suggests that the relationship may hold generally in inhomogeneous systems, and can provide a way to make the optimisation process much more efficient: a starting-point for the one-electron term is provided analytically, depending on the two-body term.

Two other techniques were shown to have little impact: alternative **k**-point sam-

pling and orbitals derived from a potential which includes the correct image-like asymptotic behaviour when an electron leaves the slab (in contrast to LDA orbitals). The fact that the finite-size error is not affected by alternative **k**-point sampling lends support to the 'hole-squashing' argument put forward to explain this error.

The final calculation of the jellium surface energy draws on all the advances made here. While this is only a VMC result, it suggests that earlier extended-system QMC simulations were flawed; the new result is in line with density-functional theory and finite-system QMC calculations. The earlier calculations did not take careful account of the Coulomb finite-size effect, and this has been shown to lead to a significant error. However, the most important error comes from comparing the results of slab and bulk calculations; it appears that bulk wave functions are significantly better than those available for slabs, which leads to an overestimate of the surface energy. This is an important consideration for any future surface energy calculations; the error is magnified in jellium, because the surface energy itself is so small.

Appendix A

The quasi-2D Ewald sum

The generalization of the Ewald summation to a system with periodic repeat in only two dimensions was first obtained by Parry [65, 66]. However, this original derivation, although it leads to the correct result, is difficult to follow. An alternative derivation is presented here in section A.1; in section A.2, the expansion of the result in the limit of small separation and large cell size is obtained.

A.1 Derivation

The problem is to find the potential due to a charge of unit magnitude at the origin and all its images in the plane. The charge distribution is therefore

$$\rho(\mathbf{r}) = \sum_{\mathbf{R}} \delta(\mathbf{r} - \mathbf{R}) \tag{A.1}$$

where $\mathbf{R}$ is a 2D lattice vector. In fact, as $|z| \to \infty$, the potential tends to $-\infty$ (relative to the potential at $z = 0$). However, the form of the potential in this limit can still be deduced, because the array of charges then resembles a uniform sheet; the potential therefore decreases linearly, and this defines the boundary conditions.

The Ewald method [78] is to rewrite the charge distribution, creating a smooth term which may be evaluated in reciprocal space and a rapidly-decaying term which converges quickly in real space. In the quasi-2D system, this is modified slightly,

and the charge distribution is in fact rewritten as

$$\begin{aligned}\rho(\mathbf{r}) &= \left(\sum_{\mathbf{R}}\left[\delta(\mathbf{r}-\mathbf{R}) - \frac{1}{\pi\sqrt{\pi}\sigma^3}e^{-(\mathbf{r}-\mathbf{R})^2/\sigma^2}\right]\right) \\ &\quad + \left(\sum_{\mathbf{R}}\frac{1}{\pi\sqrt{\pi}\sigma^3}e^{-(\mathbf{r}-\mathbf{R})^2/\sigma^2} - \frac{1}{\sqrt{\pi}\sigma A}e^{-z^2/\sigma^2}\right) \\ &\quad + \left(\frac{1}{\sqrt{\pi}\sigma A}e^{-z^2/\sigma^2}\right) \\ &= \rho_1(\mathbf{r}) + \rho_2(\mathbf{r}) + \rho_3(\mathbf{r}).\end{aligned} \tag{A.2}$$

Here, A is the area of the 2D cell defined by the primitive lattice vectors and σ is a parameter which may be adjusted to assure the rapid convergence of real and reciprocal space sums (without affecting the result). The three terms on the right-hand side of equation (A.2) will be dealt with separately,[1] starting with ρ_1.

The contribution to the potential from each term in the sum decays rapidly with $|\mathbf{r}-\mathbf{R}|$; the potential is therefore evaluated in real space, using Gauss' Law. Consider the charge distribution

$$\rho(\mathbf{r}) = \delta(\mathbf{r}) - \frac{1}{\pi\sqrt{\pi}\sigma^3}e^{-\mathbf{r}^2/\sigma^2}. \tag{A.3}$$

By Gauss' Law, the electric field generated by this potential has magnitude

$$E(r) = \frac{2e^{-r^2/\sigma^2}}{\sqrt{\pi}\sigma r} + \frac{1}{r^2}\operatorname{erfc}\left(\frac{r}{\sigma}\right) \tag{A.4}$$

in atomic units. Insisting that the potential must tend to zero as $r \to \infty$ gives

$$\begin{aligned}\phi(r) &= \int_r^\infty E(r')\,\mathrm{d}r' \\ &= \frac{1}{r}\operatorname{erfc}\left(\frac{r}{\sigma}\right)\end{aligned} \tag{A.5}$$

so that the contribution to the potential from the charge distribution ρ_1 is

$$\phi_1(\mathbf{r}) = \sum_{\mathbf{R}}\frac{1}{|\mathbf{r}-\mathbf{R}|}\operatorname{erfc}\left(\frac{|\mathbf{r}-\mathbf{R}|}{\sigma}\right). \tag{A.6}$$

[1] The first two terms are overall charge-neutral; far from the slab, the potential due to each of these terms must tend to zero. The linear dependence of the overall potential in this limit must be provided entirely by the third term.

The second contribution comes from a reciprocal space sum; the charge distribution is therefore rewritten as

$$\rho_2(\mathbf{r}) = e^{-z^2/\sigma^2} \sum_{\mathbf{k}} \rho_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{r}_\parallel} \tag{A.7}$$

where $\mathbf{k}$ is the set of 2D reciprocal lattice vectors and the $\rho_{\mathbf{k}}$ are Fourier coefficients. The sum excludes $\mathbf{k} = \mathbf{0}$, because $\rho_{\mathbf{k}=\mathbf{0}} = 0$ by design. The other coefficients are given by

$$\begin{aligned}
\rho_{\mathbf{k}} &= \frac{1}{A} \int_{\text{cell}} \sum_{\mathbf{R}} \frac{1}{\sqrt{\pi}\pi\sigma^3} e^{-(\mathbf{r}_\parallel - \mathbf{R})^2/\sigma^2} e^{-i\mathbf{k}\cdot\mathbf{r}_\parallel} \, \mathrm{d}\mathbf{r}_\parallel \\
&= \frac{1}{A\sqrt{\pi}\pi\sigma^3} \int_{\text{space}} e^{-\mathbf{r}_\parallel^2/\sigma^2 - i\mathbf{k}\cdot\mathbf{r}_\parallel} \, \mathrm{d}\mathbf{r}_\parallel \\
&= \frac{1}{A\sqrt{\pi}\sigma} e^{-k^2\sigma^2/4}.
\end{aligned} \tag{A.8}$$

The desired potential is expressed as a similar series:

$$\phi_2(\mathbf{r}) = \sum_{\mathbf{k}} \phi_{\mathbf{k}}(z) e^{i\mathbf{k}\cdot\mathbf{r}_\parallel}. \tag{A.9}$$

These expressions may then be substituted into Poisson's equation,

$$\nabla^2 \phi_2(\mathbf{r}) = -4\pi\rho_2(\mathbf{r}), \tag{A.10}$$

to give an equation for the coefficients $\phi_{\mathbf{k}}(z)$:

$$\left(\frac{d^2}{dz^2} - k^2\right) \phi_{\mathbf{k}}(z) = -\frac{4\sqrt{\pi}}{\sigma A} e^{-z^2/\sigma^2 - k^2\sigma^2/4}. \tag{A.11}$$

This may be solved with the Green's function

$$G_{\mathbf{k}}(z, z') = -\frac{1}{2k} e^{-k|z-z'|} \tag{A.12}$$

to give

$$\phi_{\mathbf{k}}(z) = -\int_{-\infty}^{\infty} \frac{1}{2k} e^{-k|z-z'|} \left(-\frac{4\sqrt{\pi}}{\sigma A} e^{-z'^2/\sigma^2 - k^2\sigma^2/4}\right) \mathrm{d}z', \tag{A.13}$$

which, after integration, yields the potential

$$\phi_2(\mathbf{r}) = \frac{\pi}{A} \sum_{\mathbf{k}} \frac{1}{k} \left[e^{-kz} \operatorname{erfc}\left(\frac{\sigma k}{2} - \frac{z}{\sigma}\right) + e^{kz} \operatorname{erfc}\left(\frac{\sigma k}{2} + \frac{z}{\sigma}\right) \right] e^{i\mathbf{k}\cdot\mathbf{r}_\parallel}. \tag{A.14}$$

Finally, the third charge distribution is

$$\rho_3(\mathbf{r}) = \frac{1}{A\sqrt{\pi}\sigma} e^{-z^2/\sigma^2}. \tag{A.15}$$

Because this function only depends on z, Poisson's equation reduces to a one-dimensional problem; the appropriate Green's function is

$$G(z, z') = \frac{1}{2}|z - z'|. \tag{A.16}$$

The potential is therefore given by

$$\begin{aligned} \phi_3(\mathbf{r}) &= -4\pi \int_{-\infty}^{\infty} \left(\frac{1}{2}|z - z'|\right)\left(\frac{1}{A\sqrt{\pi}\sigma} e^{-z'^2/\sigma^2}\right) \mathrm{d}z' \\ &= -\frac{2\pi}{A}\left(z \operatorname{erf}\left(\frac{z}{\sigma}\right) + \frac{\sigma}{\sqrt{\pi}} e^{-z^2/\sigma^2}\right). \end{aligned} \tag{A.17}$$

Combining the three previous results gives the following expression for the potential due to a charge at $\mathbf{r} = \mathbf{0}$ and the corresponding images:

$$\begin{aligned} v_{\mathrm{E}}(\mathbf{r}) = &\sum_{\mathbf{R}} \frac{1}{|\mathbf{r} - \mathbf{R}|} \operatorname{erfc}\left(\frac{|\mathbf{r} - \mathbf{R}|}{\sigma}\right) - \frac{2\pi}{A}\left[z \operatorname{erf}\left(\frac{z}{\sigma}\right) + \frac{\sigma}{\sqrt{\pi}} e^{-z^2/\sigma^2}\right] \\ &+ \sum_{\mathbf{k}} \frac{\pi}{kA}\left[e^{-kz} \operatorname{erfc}\left(\frac{\sigma k}{2} - \frac{z}{\sigma}\right) + e^{kz} \operatorname{erfc}\left(\frac{\sigma k}{2} + \frac{z}{\sigma}\right)\right] e^{i\mathbf{k}\cdot\mathbf{r}_\parallel}. \end{aligned} \tag{A.18}$$

The self-interaction energy is the energy associated with the interaction between this charge and its images:

$$\begin{aligned} \xi &= \lim_{\mathbf{r}\to\mathbf{0}} \left(v_{\mathrm{E}}(\mathbf{r}) - \frac{1}{r}\right) \\ &= \lim_{\mathbf{r}\to\mathbf{0}} \left(\frac{1}{r} \operatorname{erfc}\left(\frac{r}{\sigma}\right) - \frac{1}{r}\right) + \sum_{\mathbf{R}\neq\mathbf{0}} \frac{1}{R} \operatorname{erfc}\left(\frac{R}{\sigma}\right) - \frac{2\sigma\sqrt{\pi}}{A} + \sum_{\mathbf{k}\neq\mathbf{0}} \frac{2\pi}{kA} \operatorname{erfc}\left(\frac{\sigma k}{2}\right) \\ &= -\frac{2}{\sigma\sqrt{\pi}} + \sum_{\mathbf{R}\neq\mathbf{0}} \frac{1}{R} \operatorname{erfc}\left(\frac{R}{\sigma}\right) - \frac{2\sigma\sqrt{\pi}}{A} + \sum_{\mathbf{k}\neq\mathbf{0}} \frac{2\pi}{kA} \operatorname{erfc}\left(\frac{\sigma k}{2}\right). \end{aligned} \tag{A.19}$$

A.2 Expansion

The function appearing in the exchange-correlation energy ($U_{\mathrm{XC}}^{\mathrm{EW}}$ in equation (4.24)) is $v_{\mathrm{E}}(\mathbf{r}) - \xi$. The exchange-correlation hole is normally short-ranged; the extent to

which $v_{\mathrm{E}}(\mathbf{r}) - \xi$ deviates from $1/r$ at small r may be used to estimate the finite-size error associated with the Ewald interaction. The expansion of this function for small r and large lattice parameter (L) follows.

Combining equations (A.19) and (A.18) gives the function to be expanded:

$$\begin{aligned} v_{\mathrm{E}}(\mathbf{r}) - \xi = \frac{2}{\sigma\sqrt{\pi}} &- \sum_{\mathbf{R}\neq\mathbf{0}} \frac{1}{R} \operatorname{erfc}\left(\frac{R}{\sigma}\right) + \sum_{\mathbf{R}} \frac{1}{|\mathbf{r}-\mathbf{R}|} \operatorname{erfc}\left(\frac{|\mathbf{r}-\mathbf{R}|}{\sigma}\right) \\ &- \frac{2\pi}{A}\left[z \operatorname{erf}\left(\frac{z}{\sigma}\right) + \frac{\sigma}{\sqrt{\pi}}\left(e^{-z^2/\sigma^2} - 1\right)\right] \\ &+ \sum_{\mathbf{k}\neq\mathbf{0}} \frac{\pi}{kA}\left\{\left[e^{-kz}\operatorname{erfc}\left(\frac{\sigma k}{2} - \frac{z}{\sigma}\right) + e^{kz}\operatorname{erfc}\left(\frac{\sigma k}{2} + \frac{z}{\sigma}\right)\right]\cos\mathbf{k}\cdot\mathbf{r}_{\parallel}\right. \\ &\left. -2\operatorname{erfc}\left(\frac{\sigma k}{2}\right)\right\}. \end{aligned} \tag{A.20}$$

The first line of this expression reduces quickly to

$$\frac{1}{r} + \frac{2r^2}{3\sigma^3\sqrt{\pi}} + \mathcal{O}\left[r^4\right] + \mathcal{O}\left[e^{-L^2/\sigma^2}\right]. \tag{A.21}$$

The second line is also simply expanded, giving

$$-\frac{2\sqrt{\pi}z^2}{\sigma A} + \mathcal{O}\left[z^4\right]. \tag{A.22}$$

The sum in **k**-space is slightly more involved. To begin, we note that

$$\operatorname{erfc}(x_0 + x) = \operatorname{erfc}(x_0) + \frac{2}{\sqrt{\pi}}(x^2 x_0 - x)e^{-x_0^2} + \mathcal{O}\left[x^3\right]. \tag{A.23}$$

Applying this result,

$$\begin{aligned} e^{-kz}\operatorname{erfc}\left(\frac{\sigma k}{2} - \frac{z}{\sigma}\right) + e^{kz}\operatorname{erfc}\left(\frac{\sigma k}{2} + \frac{z}{\sigma}\right) &= \left(2 + k^2 z^2\right)\operatorname{erfc}\left(\frac{\sigma k}{2}\right) \\ &\quad - \frac{2z^2 k}{\sigma\sqrt{\pi}} e^{-(\sigma k/2)^2} + \mathcal{O}\left[z^4\right]. \end{aligned} \tag{A.24}$$

The error is of order z^4 rather than z^3 because any terms involving odd powers of z must cancel out. The next step is to expand the cosine to $\mathcal{O}\left[r^2\right]$; the **k**-space sum of equation (A.20) then becomes

$$\sum_{\mathbf{k}\neq\mathbf{0}}\left[\frac{\pi k}{A}\left(z^2 - \frac{(\mathbf{k}\cdot\mathbf{r}_{\parallel})^2}{k^2}\right)\operatorname{erfc}\left(\frac{\sigma k}{2}\right) - \frac{2\sqrt{\pi}z^2}{\sigma A}e^{-(\sigma k/2)^2}\right] + \mathcal{O}\left[r^4\right] \tag{A.25}$$

which reduces to

$$\sum_{\mathbf{k}\neq\mathbf{0}}\left[\frac{\pi k}{A}\left(z^2-\frac{k^2r_\parallel^2}{2k^2}\right)\operatorname{erfc}\left(\frac{\sigma k}{2}\right)-\frac{2\sqrt{\pi}z^2}{\sigma A}e^{-(\sigma k/2)^2}\right]+\mathcal{O}\left[r^4\right] \tag{A.26}$$

from a comparison of the contributions to the sum of all the **k**-vectors of a given magnitude.

To proceed further, we use the following two-dimensional Fourier series:

$$\sum_{\mathbf{R}}e^{-((\mathbf{r}-\mathbf{R})/\sigma)^2}=\sum_{\mathbf{k}}\frac{\pi\sigma^2}{A}e^{-(\sigma k/2)^2}e^{i\mathbf{k}\cdot\mathbf{r}} \tag{A.27}$$

with $\mathbf{r}=\mathbf{0}$. This gives

$$\begin{aligned}\sum_{\mathbf{k}\neq\mathbf{0}}e^{-(\sigma k/2)^2}&=\frac{A}{\pi\sigma^2}\sum_{\mathbf{R}}e^{-(R/\sigma)^2}-1\\&=\frac{A}{\pi\sigma^2}-1+\mathcal{O}\left[e^{-L^2/\sigma^2}\right]\end{aligned} \tag{A.28}$$

which, when substituted back into equation (A.20), leads to

$$v_{\mathrm{E}}(\mathbf{r})-\xi=\frac{1}{r}+\left(z^2-\frac{r_\parallel^2}{2}\right)\left(\sum_{\mathbf{k}\neq\mathbf{0}}\frac{\pi k}{A}\operatorname{erfc}\left(\frac{\sigma k}{2}\right)-\frac{4}{3\sigma^3\sqrt{\pi}}\right)+\mathcal{O}\left[r^4\right]+\mathcal{O}\left[e^{-L^2/\sigma^2}\right]. \tag{A.29}$$

The remaining **k**-space sum may written in terms of the new variable $\beta=1/\sigma^2$:

$$\begin{aligned}\sum_{\mathbf{k}\neq\mathbf{0}}\frac{\pi k}{A}\operatorname{erfc}\left(\frac{\sigma k}{2}\right)&=\sum_{\mathbf{k}}\frac{\pi k}{A}\operatorname{erfc}\left(\frac{k}{2\sqrt{\beta}}\right)\\&=S(\beta).\end{aligned} \tag{A.30}$$

Differentiating,

$$\begin{aligned}\frac{dS}{d\beta}&=\sum_{\mathbf{k}}\frac{\sqrt{\pi}}{2A\beta\sqrt{\beta}}k^2e^{-k^2/4\beta}\\&=\frac{2\sqrt{\pi}\sqrt{\beta}}{A}\frac{d}{d\beta}\left(\sum_{\mathbf{k}}e^{-k^2/4\beta}\right)\\&=\frac{2\sqrt{\pi}\sqrt{\beta}}{A}\frac{d}{d\beta}\left(\frac{A\beta}{\pi}\sum_{\mathbf{R}}e^{-R^2\beta}\right)\\&=\frac{2\sqrt{\beta}}{\sqrt{\pi}}\sum_{\mathbf{R}}(1-R^2\beta)e^{-R^2\beta}.\end{aligned} \tag{A.31}$$

Here, equation (A.27) has again been used to convert the reciprocal space sum to one in real space. To recover the original sum, we integrate from 0 to β:

$$\begin{aligned} S(\beta) - S(0) &= \int_0^\beta \frac{2\sqrt{\beta}}{\sqrt{\pi}} \sum_{\mathbf{R}} (1 - R^2\beta) e^{-R^2\beta} \, \mathrm{d}\beta \\ &= \frac{2}{\sqrt{\pi}} \sum_{\mathbf{R}} \int_0^{\sqrt{\beta}} 2t(1 - R^2t^2) e^{-R^2t^2} \, \mathrm{d}t \\ &= \frac{4}{\sqrt{\pi}} \sum_{\mathbf{R}\neq\mathbf{0}} (I_2 - R^4 I_4) + \frac{4}{\sqrt{\pi}} \int_0^{\sqrt{\beta}} t^2 \, \mathrm{d}t. \end{aligned} \tag{A.32}$$

The two remaining integrals are

$$\begin{aligned} I_4 &= \int_0^{\sqrt{\beta}} t^4 e^{-R^2t^2} \, \mathrm{d}t \\ &= -\frac{\beta\sqrt{\beta}}{2R^2} e^{-R^2\beta} + \frac{3}{2R^2} I_2 \end{aligned} \tag{A.33}$$

and

$$\begin{aligned} I_2 &= \int_0^{\sqrt{\beta}} t^2 e^{-R^2t^2} \, \mathrm{d}t \\ &= -\frac{\sqrt{\beta}}{2R^2} e^{-R^2\beta} + \frac{\sqrt{\pi}}{4R^3} \operatorname{erf}\left(R\sqrt{\beta}\right). \end{aligned} \tag{A.34}$$

Substituting these results into equation (A.32), and using the fact that $S(0) = 0$,

$$\begin{aligned} S(\beta) &= \frac{4}{\sqrt{\pi}} \sum_{\mathbf{R}\neq\mathbf{0}} \left[\left(\frac{\sqrt{\beta}}{4R^2} - \frac{\beta\sqrt{\beta}}{2} \right) e^{-R^2\beta} - \frac{\sqrt{\pi}}{8R^3} \operatorname{erf}\left(R\sqrt{\beta}\right) \right] + \frac{4\beta\sqrt{\beta}}{3\sqrt{\pi}} \\ &= \frac{4\beta\sqrt{\beta}}{3\sqrt{\pi}} - \frac{1}{2} \sum_{\mathbf{R}\neq\mathbf{0}} \frac{1}{R^3} \operatorname{erf}\left(R\sqrt{\beta}\right) + \mathcal{O}\left[e^{-L^2\beta}\right]. \end{aligned} \tag{A.35}$$

Returning to the original summation,

$$\begin{aligned} \sum_{\mathbf{k}\neq\mathbf{0}} \frac{\pi k}{A} \operatorname{erfc}\left(\frac{\sigma k}{2}\right) &= \frac{4}{3\sigma^3\sqrt{\pi}} - \frac{1}{2} \sum_{\mathbf{R}\neq\mathbf{0}} \frac{1}{R^3} \operatorname{erf}\left(\frac{R}{\sigma}\right) + \mathcal{O}\left[e^{-L^2/\sigma^2}\right] \\ &= \frac{4}{3\sigma^3\sqrt{\pi}} - \frac{C}{L^3} + \mathcal{O}\left[e^{-L^2/\sigma^2}\right] \end{aligned} \tag{A.36}$$

where the constant C is given by

$$C = \frac{1}{2} \sum_{m,n=-\infty}^{\infty}{}' (m^2 + n^2)^{-3/2} \tag{A.37}$$

and the prime here indicates that the term with $m = n = 0$ should be excluded from the sum. It may be shown [10] that this sum reduces to

$$4\beta(3/2)\zeta(3/2) \tag{A.38}$$

where β and ζ are the Dirichlet beta and Riemann zeta functions respectively, giving $C = 4.516810842$.

Combining equations (A.36) and (A.29) gives the final result:

$$v_{\mathrm{E}}(\mathbf{r}) - \xi = \frac{1}{r} - \frac{C}{L^3}(z^2 - \frac{r_{\parallel}^2}{2}) + \mathcal{O}\left[r^4\right] + \mathcal{O}\left[e^{-L^2/\sigma^2}\right]. \tag{A.39}$$

Appendix B

The cusp conditions

The true ground-state many-electron wave function satisfies the following Schrödinger equation:

$$\sum_i \left(-\frac{\hbar^2}{2m_e}\nabla_i^2 + V(\mathbf{r}_i) + \frac{1}{2}\sum_{j\neq i}\frac{e^2}{4\pi\epsilon_0 r_{ij}} \right) \Psi(\{\mathbf{r}_i\}) = E\Psi(\{\mathbf{r}_i\}). \tag{B.1}$$

The cusp conditions, originally derived by Kato [37], constrain the behaviour of the wave function as two electrons approach each other. To study this behaviour, it is useful to write the Hamiltonian operator as

$$\begin{aligned}\hat{H}(\{\mathbf{r}_i\}) = &\sum_{i>2}\left[-\frac{\hbar^2}{2m_e}\nabla_i^2 + V(\mathbf{r}_i) + \frac{1}{2}\sum_{j\neq i,\ j>2}\frac{e^2}{4\pi\epsilon_0 r_{ij}} \right] \\ &+ \frac{e^2}{4\pi\epsilon_0}\sum_{i>2}\left[\frac{1}{|\mathbf{R}-\mathbf{r}_j+\frac{1}{2}\mathbf{r}|} + \frac{1}{|\mathbf{R}-\mathbf{r}_j-\frac{1}{2}\mathbf{r}|} \right] \\ &+ \left[-\frac{\hbar^2}{4m_e}\nabla_{\mathbf{R}}^2 - \frac{\hbar^2}{m_e}\nabla_{\mathbf{r}}^2 + V\left(\mathbf{R}+\frac{\mathbf{r}}{2}\right) + V\left(\mathbf{R}-\frac{\mathbf{r}}{2}\right) + \frac{e^2}{4\pi\epsilon_0 r} \right]\end{aligned} \tag{B.2}$$

where the new coordinates are

$$\mathbf{R} = \frac{\mathbf{r}_1+\mathbf{r}_2}{2} \tag{B.3}$$

$$\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2. \tag{B.4}$$

The distance between electrons 1 and 2 is r. As $r \to 0$, the Coulomb interaction energy ($e^2/4\pi\epsilon_0 r$) of these electrons diverges. In this limit, the Hamiltonian operator

may be written as

$$\begin{aligned}\hat{H}(\{\mathbf{r}_i\}) = \sum_{i>2}\left[-\frac{\hbar^2}{2m_e}\nabla_i^2 + V(\mathbf{r}_i) + \frac{e^2}{4\pi\epsilon_0}\left(\frac{1}{2}\sum_{j\neq i,\ j>2}\frac{1}{r_{ij}} + \frac{2}{|\mathbf{R}-\mathbf{r}_i|}\right)\right] \\ + \left[-\frac{\hbar^2}{4m_e}\nabla_{\mathbf{R}}^2 + 2V(\mathbf{R})\right] + \left[-\frac{\hbar^2}{m_e}\nabla_{\mathbf{r}}^2 + \frac{e^2}{4\pi\epsilon_0 r} + \mathcal{O}[r^2]\right].\end{aligned} \tag{B.5}$$

For any eigenstate of $\hat{H}$, the divergence in the Coulomb energy must cancel out exactly; the only other term which can give rise to a matching divergence is

$$-\frac{\hbar^2}{m_e}\nabla_{\mathbf{r}}^2. \tag{B.6}$$

Therefore, when two electrons are close together, the Schrödinger equation is entirely dominated by these two terms and is approximately separable; the part of the wave function depending on $\mathbf{r}$ is a solution of the reduced Schrödinger equation

$$\left(-\frac{\hbar^2}{m_e}\nabla_{\mathbf{r}}^2 + \frac{e^2}{4\pi\epsilon_0 r}\right)\Psi_{\mathbf{r}}(\mathbf{r}) = E\Psi_{\mathbf{r}}(\mathbf{r}). \tag{B.7}$$

This hydrogenic equation has the well-known solution in spherical polar coordinates

$$\Psi_{\mathbf{r}}(\mathbf{r}) = \sum_{l,\,m} f_{lm}(r)Y_{lm}(\theta,\phi) \tag{B.8}$$

where the Y_{lm} are spherical harmonics and f_{lm} satisfies the radial equation

$$-\frac{\hbar^2}{m_e r^2}\frac{d}{dr}\left(r^2\frac{df_{lm}}{dr}\right) + \left(\frac{e^2}{4\pi\epsilon_0 r} + \frac{l(l+1)\hbar^2}{m_e r^2}\right)f_{lm} = Ef_{lm}. \tag{B.9}$$

Substituting the power series solution

$$f_{lm} = r^p\sum_{j=0}^{\infty} a_{lmj}r^j \tag{B.10}$$

into equation (B.9) gives

$$r^p\sum_{j=0}^{\infty} a_{lmj}\left(\frac{\hbar^2}{m_e}\big[l(l+1)-(p+j)(p+j+1)\big]r^{j-2} + \frac{e^2}{4\pi\epsilon_0}r^{j-1} - Er^j\right) = 0 \tag{B.11}$$

so that[1] $p = l$ and

$$a_{lm1} = a_{lm0}\left(\frac{m_e e^2}{4\pi\epsilon_0\hbar^2}\right)\frac{1}{2(l+1)}. \tag{B.12}$$

[1]The alternative result, $p = -l-1$, is rejected on physical grounds.

The higher coefficients (a_{lmj} with $j \geq 2$) depend on E and therefore on the positions of the other electrons. Combining equations (B.8), (B.10) and (B.12) gives the expansion of the wave function for small r:

$$\begin{aligned}\Psi_{\mathbf{r}}(\mathbf{r}) = &\left(1 + \frac{r}{2} \cdot \frac{m_e e^2}{4\pi\epsilon_0 \hbar^2}\right) a_{000} Y_{00} \\ &+ r \left(1 + \frac{r}{4} \cdot \frac{m_e e^2}{4\pi\epsilon_0 \hbar^2}\right) \sum_{m=-1}^{1} a_{1m0} Y_{1m}(\theta, \phi) \\ &+ \text{higher order terms.}\end{aligned} \tag{B.13}$$

The starting-point for the wave function used in quantum Monte Carlo simulations consists of a product of up- and down-spin determinants. Because the determinants are smooth functions of the electron coordinates, the wave function may be expanded about the point $r = 0$:

$$\Psi_D = D^{\uparrow} D^{\downarrow} = c_0 + \mathbf{c}_1 \cdot \mathbf{r} + \mathcal{O}[r^2]. \tag{B.14}$$

For electrons of parallel spin, the wave function has a node when $r = 0$, which means that $c_0 = 0$; for electrons of antiparallel spin, this is not the case. Rewriting the expansion in terms of spherical harmonics gives

$$\Psi_D = a_0 Y_{00} + r \left(a_{11} Y_{11} + a_{10} Y_{10} + a_{11} Y_{1(-1)}\right) + \mathcal{O}[r^2]. \tag{B.15}$$

If $a_0 = 0$ (parallel spins), a corrected wave function with the same cusp as derived in equation (B.13) is

$$\Psi_{\text{corr}}^{\uparrow\uparrow} = \left(1 + \frac{r}{4} \cdot \frac{m_e e^2}{4\pi\epsilon_0 \hbar^2}\right) \Psi_D. \tag{B.16}$$

The corrected version when $a_0 \neq 0$ (antiparallel spins) is

$$\Psi_{\text{corr}}^{\uparrow\downarrow} = \left(1 + \frac{r}{2} \cdot \frac{m_e e^2}{4\pi\epsilon_0 \hbar^2}\right) \Psi_D. \tag{B.17}$$

In each case, only the lowest-order term is corrected.

This behaviour may be conveniently incorporated into a Slater-Jastrow wave function of the form

$$\Psi = e^J \Psi_D \tag{B.18}$$

with

$$J = -\frac{1}{2}\sum_{i\neq j} u_{\text{cusp}}(\mathbf{x}_i, \mathbf{x}_j) = -\frac{1}{2}\sum_{i\neq j} \frac{m_e e^2}{4\pi\epsilon_0\hbar^2}\cdot\frac{1}{4}\left(\delta_{\sigma_i\sigma_j} - 2\right) r_{ij}. \tag{B.19}$$

However, while this form gives the correct divergence when two electrons are close together, it spoils the wave function when the separation is large. A more sensible choice for u_{cusp} is

$$u_{\text{cusp}}(\mathbf{x}_i, \mathbf{x}_j) = \left(\frac{m_e e^2}{4\pi\epsilon_0\hbar^2}\right)\beta_{\sigma_i\sigma_j} e^{-\alpha_{\sigma_i\sigma_j} r_{ij}} \tag{B.20}$$

with the parameters chosen so that

$$\alpha_{\sigma_i\sigma_j}\beta_{\sigma_i\sigma_j} = \begin{cases} 1/4 & \text{when } \sigma_i = \sigma_j \\ 1/2 & \text{otherwise.} \end{cases} \tag{B.21}$$

As $|\mathbf{r}_i - \mathbf{r}_j| \to \infty$, $u_{\text{cusp}}(\mathbf{r}_i, \mathbf{r}_j) \to 0$ and the original form of the wave function is restored; the rate at which u_{cusp} decays in this limit is determined by $\alpha_{\sigma_i\sigma_j}$.

Appendix C

Integrating the cusp function

In the course of determining the one-body function χ for the jellium slab, it is necessary to evaluate the following integral:

$$I = 2\pi \int_{z'=-\infty}^{\infty} \int_{\rho'=0}^{\infty} \bar{n}(z') e^{-k_c \sqrt{\rho'^2+(z-z')^2} - [\rho'^2+(z-z')^2]/(L_c^2)} \rho' \, d\rho' \, dz'. \tag{C.1}$$

With the substitution $q = \sqrt{\rho'^2 + (z-z')^2}$, this reduces to

$$\begin{aligned} I &= 2\pi \int_{z'=-\infty}^{\infty} \int_{q=|z-z'|}^{\infty} \bar{n}(z') e^{-k_c q - q^2/L_c^2} q \, dq \, dz' \\ &= 2\pi \int_{z'=-\infty}^{\infty} \int_{q=|z-z'|}^{\infty} \bar{n}(z') e^{-[q/L_c + k_c L_c/2]^2 + k_c^2 L_c^2/4} q \, dq \, dz'. \end{aligned} \tag{C.2}$$

The further substitution

$$p = \frac{q}{L_c} + \frac{k_c L_c}{2} \tag{C.3}$$

then gives

$$\begin{aligned} I &= 2\pi L_c^2 e^{k_c^2 L_c^2/4} \int_{z'=-\infty}^{\infty} \int_{p=|z-z'|/L_c + k_c L_c/2}^{\infty} \bar{n}(z') e^{-p^2} \left(p - \frac{k_c L_c}{2} \right) dp \, dz' \\ &= \pi L_c^2 e^{k_c^2 L_c^2/4} \int_{-\infty}^{\infty} \bar{n}(z') \left\{ -\frac{k_c L_c \sqrt{\pi}}{2} \operatorname{erfc}\left(\frac{|z-z'|}{L_c} + \frac{k_c L_c}{2} \right) \right. \\ &\quad \left. + \exp\left(-\frac{|z-z'|^2}{L_c^2} - k_c |z-z'| - \frac{k_c^2 L_c^2}{4} \right) \right\} dz'. \end{aligned} \tag{C.4}$$

To proceed further, it is necessary to know the density profile $\bar{n}(z')$. In deriving the

plasmon modes, the following approximation was used:

$$\bar{n}(z') = \begin{cases} n_0 & 0 < z' < s \\ 0 & \text{otherwise.} \end{cases} \tag{C.5}$$

Applying the same approximation here gives

$$\begin{aligned} I = n_0 \pi L_c^2 e^{k_c^2 L_c^2/4} \Bigg\{ & -\frac{k_c L_c \sqrt{\pi}}{2} \int_0^{a(z)} \operatorname{erfc}\left(\frac{z - z'}{L_c} + \frac{k_c L_c}{2} \right) \, dz' \\ & + \int_0^{a(z)} \exp\left(-\frac{(z - z')^2}{L_c^2} - k_c (z - z') - \frac{k_c^2 L_c^2}{4} \right) dz' \\ & - \frac{k_c L_c \sqrt{\pi}}{2} \int_{a(z)}^{s} \operatorname{erfc}\left(\frac{z' - z}{L_c} + \frac{k_c L_c}{2} \right) \, dz' \\ & + \int_{a(z)}^{s} \exp\left(-\frac{(z' - z)^2}{L_c^2} - k_c (z' - z) - \frac{k_c^2 L_c^2}{4} \right) dz' \Bigg\} \\ = n_0 \pi L_c^2 e^{k_c^2 L_c^2/4} & (I_1 + I_2 + I_3 + I_4) \end{aligned} \tag{C.6}$$

where the following function has been introduced to represent the limits:

$$a(z) = \begin{cases} 0 & \text{when } z < 0 \\ z & \text{when } 0 < z < s \\ s & \text{otherwise.} \end{cases} \tag{C.7}$$

The first integral may be solved with the substitution

$$w = \frac{z - z'}{L_c} + \frac{k_c L_c}{2}. \tag{C.8}$$

and by parts:

$$\begin{aligned}
I_1 &= -\frac{k_c L_c^2 \sqrt{\pi}}{2} \int_{[z-a(z)]/L_c+k_c L_c/2}^{z/L_c+k_c L_c/2} \operatorname{erfc}(w)\ dw \\
&= -\frac{k_c L_c^2 \sqrt{\pi}}{2} \left\{ \Big[w \operatorname{erfc}(w) \Big]_{w=[z-a(z)]/L_c+k_c L_c/2}^{w=z/L_c+k_c L_c/2} + \frac{2}{\sqrt{\pi}} \int_{[z-a(z)]/L_c+k_c L_c/2}^{z/L_c+k_c L_c/2} w e^{-w^2}\ dw \right\} \\
&= -\frac{k_c L_c^2 \sqrt{\pi}}{2} \left\{ \left(\frac{z}{L_c} + \frac{k_c L_c}{2} \right) \operatorname{erfc}\left(\frac{z}{L_c} + \frac{k_c L_c}{2} \right) \right. \\
&\quad - \left(\frac{z-a(z)}{L_c} + \frac{k_c L_c}{2} \right) \operatorname{erfc}\left(\frac{z-a(z)}{L_c} + \frac{k_c L_c}{2} \right) \\
&\quad \left. - \frac{1}{\sqrt{\pi}} e^{-(z/L_c+k_c L_c/2)^2} + \frac{1}{\sqrt{\pi}} e^{-([z-a(z)]/L_c+k_c L_c/2)^2} \right\}.
\end{aligned} \tag{C.9}$$

For the third integral, which is similar, the appropriate substitution is

$$w = \frac{z' - z}{L_c} + \frac{k_c L_c}{2}, \tag{C.10}$$

giving

$$\begin{aligned}
I_3 &= -\frac{k_c L_c^2 \sqrt{\pi}}{2} \int_{[a(z)-z]/L_c+k_c L_c/2}^{(s-z)/L_c+k_c L_c/2} \operatorname{erfc}(w)\ dw \\
&= -\frac{k_c L_c^2 \sqrt{\pi}}{2} \left\{ \Big[w \operatorname{erfc}(w) \Big]_{w=[a(z)-z]/L_c+k_c L_c/2}^{w=(s-z)/L_c+k_c L_c/2} + \frac{2}{\sqrt{\pi}} \int_{[a(z)-z]/L_c+k_c L_c/2}^{(s-z)/L_c+k_c L_c/2} w e^{-w^2}\ dw \right\} \\
&= -\frac{k_c L_c^2 \sqrt{\pi}}{2} \left\{ \left(\frac{s-z}{L_c} + \frac{k_c L_c}{2} \right) \operatorname{erfc}\left(\frac{s-z}{L_c} + \frac{k_c L_c}{2} \right) \right. \\
&\quad - \left(\frac{a(z)-z}{L_c} + \frac{k_c L_c}{2} \right) \operatorname{erfc}\left(\frac{a(z)-z}{L_c} + \frac{k_c L_c}{2} \right) \\
&\quad \left. - \frac{1}{\sqrt{\pi}} e^{-[(s-z)/L_c+k_c L_c/2]^2} + \frac{1}{\sqrt{\pi}} e^{-[(a(z)-z)/L_c+k_c L_c/2]^2} \right\}.
\end{aligned} \tag{C.11}$$

The second integral is

$$\begin{aligned}
I_2 &= \int_0^{a(z)} \exp\left(-\left[\frac{z-z'}{L_c} + \frac{k_c L_c}{2} \right]^2 \right) dz' \\
&= -L_c \int_{z/L_c+k_c L_c/2}^{[z-a(z)]/L_c+k_c L_c/2} e^{-w^2}\ dw
\end{aligned} \tag{C.12}$$

with the same substitution described in equation (C.8). This gives

$$I_2 = \frac{L_c\sqrt{\pi}}{2}\left[\operatorname{erfc}\left(\frac{z-a(z)}{L_c}+\frac{k_cL_c}{2}\right) - \operatorname{erfc}\left(\frac{z}{L_c}+\frac{k_cL_c}{2}\right)\right]. \tag{C.13}$$

Finally,

$$\begin{aligned} I_4 &= \int_{a(z)}^{s} \exp\left(-\left[\frac{z'-z}{L_c}+\frac{k_cL_c}{2}\right]^2\right) dz' \\ &= L_c \int_{[a(z)-z]/L_c+k_cL_c/2}^{(s-z)/L_c+k_cL_c/2} e^{-w^2}\, dw \end{aligned} \tag{C.14}$$

using the substitution defined in equation (C.10), which gives

$$I_4 = \frac{L_c\sqrt{\pi}}{2}\left[\operatorname{erfc}\left(\frac{b(z)-z}{L_c}+\frac{k_cL_c}{2}\right) - \operatorname{erfc}\left(\frac{s-z}{L_c}+\frac{k_cL_c}{2}\right)\right]. \tag{C.15}$$

Collating the separate sub-integrals gives the final result:

$$\begin{aligned} I = \frac{n_0k_cL_c^4\pi\sqrt{\pi}}{2}e^{k_c^2L_c^2/4}\Bigg\{ &-\left(\frac{z}{L_c}+\frac{k_cL_c}{2}+\frac{1}{k_cL_c}\right)\operatorname{erfc}\left(\frac{z}{L_c}+\frac{k_cL_c}{2}\right) \\ &-\left(\frac{s-z}{L_c}+\frac{k_cL_c}{2}+\frac{1}{k_cL_c}\right)\operatorname{erfc}\left(\frac{s-z}{L_c}+\frac{k_cL_c}{2}\right) \\ &+\left(\frac{z-a(z)}{L_c}+\frac{k_cL_c}{2}+\frac{1}{k_cL_c}\right)\operatorname{erfc}\left(\frac{z-a(z)}{L_c}+\frac{k_cL_c}{2}\right) \\ &+\left(\frac{a(z)-z}{L_c}+\frac{k_cL_c}{2}+\frac{1}{k_cL_c}\right)\operatorname{erfc}\left(\frac{a(z)-z}{L_c}+\frac{k_cL_c}{2}\right) \\ &+\frac{1}{\sqrt{\pi}}e^{-(z/L_c+k_cL_c/2)^2}+\frac{1}{\sqrt{\pi}}e^{-[(s-z)/L_c+k_cL_c/2]^2} \\ &-\frac{1}{\sqrt{\pi}}e^{-([z-a(z)]/L_c+k_cL_c/2)^2}-\frac{1}{\sqrt{\pi}}e^{-([a(z)-z]/L_c+k_cL_c/2)^2}\Bigg\}. \end{aligned} \tag{C.16}$$

Appendix D

Reconstructing a probability density function

The problem is this: given a set of points sampled from some distribution, how can one obtain a good approximation to the original distribution?

The solution must obviously address the issue of what constitutes a 'good' approximation: some possibilities are smoothness, correct incorporation of symmetry, and implementation of known boundary conditions.

Two ways of approaching this problem are presented here; the first method will be shown to be almost a special case of the second.

In both sections, the original probability density function is denoted $f(x)$; the approximate reconstructed function is $g(x)$, and the sampled points are $\{x_i : i = 1, 2, \ldots, N\}$.

D.1 Use of a weighting function

The first approach is to permit each sampled point to contribute to the value of the reconstructed function:

$$g(x) = \frac{1}{N} \sum_{i=1}^{N} w(x_i, x). \tag{D.1}$$

The weighting function $w(x_i, x)$ is not necessarily homogeneous (although in many cases, this is a sensible choice). Each sample point carries the same total weight; mathematically, this is expressed as

$$\int w(x_i, x)\, dx = 1. \tag{D.2}$$

The reason for choosing the total weight to be one rather than some other constant is to ensure that g has the normalisation appropriate to a probability density:

$$\begin{aligned} \int g(x)\, dx &= \frac{1}{N} \sum_{i=1}^{N} \int w(x_i, x)\, dx \\ &= 1. \end{aligned} \tag{D.3}$$

The expected value of the reconstructed function is

$$\big\langle g(x) \big\rangle_{\{x_i\}} = \big\langle w(x_i, x) \big\rangle_{x_i} = \int w(y, x) f(y)\, dy. \tag{D.4}$$

Evidently, the best approximation from this point of view is achieved when $w(y, x) = \delta(y - x)$. A good approximation is one in which $w(y, x)$ is localised, tending to zero quickly as $|y - x|$ increases; the extent of w should be smaller than the length scale on which changes in f occur.

However, this is not the whole story; the expected error in g must also be considered:

$$\begin{aligned} \left\langle \Big(g(x) - f(x) \Big)^2 \right\rangle_{\{x_i\}} &= \big\langle g(x)^2 \big\rangle_{\{x_i\}} - 2f(x) \big\langle g(x) \big\rangle_{\{x_i\}} + f(x)^2 \\ &= \frac{1}{N^2} \sum_{i,j} \big\langle w(x_i, x) w(x_j, x) \big\rangle_{x_i, x_j} - 2f(x) \big\langle w(x_i, x) \big\rangle_{x_i} \\ &\quad + f(x)^2. \end{aligned} \tag{D.5}$$

To proceed further, it is necessary to separate the terms with $i = j$, and to use the

fact that x_i and x_j are independent when $i \neq j$:

$$\begin{aligned}\left\langle \Big(g(x) - f(x)\Big)^2 \right\rangle_{\{x_i\}} &= \frac{N-1}{N}\big\langle w(x_i,x)\big\rangle_{x_i}^2 + \frac{1}{N}\big\langle w(x_i,x)^2\big\rangle_{x_i} \\ &\quad - 2f(x)\big\langle w(x_i,x)\big\rangle_{x_i} + f(x)^2 \\ &= \Big(\big\langle w(x_i,x)\big\rangle_{x_i} - f(x)\Big)^2 \\ &\quad + \frac{1}{N}\Big(\big\langle w(x_i,x)^2\big\rangle_{x_i} - \big\langle w(x_i,x)\big\rangle_{x_i}^2\Big) \\ &= \left(\int w(y,x)f(y)\,dy - f(x)\right)^2 \\ &\quad + \frac{1}{N}\int\Big(w(y,x) - \int w(z,x)f(z)\,dz\Big)^2 f(y)\,dy.\end{aligned} \tag{D.6}$$

When w is made too narrow and delta-function-like, the second term (the variance of w for a particular value of x) becomes large. The factor of $1/N$ means that with more sample points, w may be made narrower; the benefit of this is that more of the fine detail in f is then recovered.

The optimum w is therefore a compromise between a broad, slowly-decaying function which ensures that g is smooth but misses some of the features of f, and a narrow, quickly-decaying function which captures all the details but makes g very noisy.

Using an inhomogeneous weighting function, it is possible to incorporate various boundary conditions. For example, if it is known that $f(x) = 0$ for $x < a$, the weight may be adjusted so that $w(y, x < a) = 0$, while still satisfying the normalisation condition expressed in equation (D.2).

It is also possible to incorporate symmetry into this formalism: if f is known to be symmetric about $x = a$ then letting $w(y,x) = w(y, a-x)$ ensures that g has the same symmetry.

As an example, consider the homogeneous Gaussian weight function

$$w(y,x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(y-x)^2/2\sigma^2}, \tag{D.7}$$

which satisfies the normalisation condition (D.2). The expected value of the recon-

structed function is

$$\left\langle g(x)\right\rangle_{\{x_i\}} = \frac{1}{\sqrt{2\pi\sigma^2}} \int e^{-(y-x)^2/2\sigma^2} f(y)\, dy. \tag{D.8}$$

The integrand is only significant in the region close to $y = x$; the size of this region is determined by σ. Expanding $f(y)$ about this point gives

$$\begin{aligned}\left\langle g(x)\right\rangle_{\{x_i\}} &= \int \frac{e^{-(y-x)^2/2\sigma^2}}{\sqrt{2\pi\sigma^2}} \left(f(x) + (y-x)f'(x) + \frac{(y-x)^2}{2} f''(x) + \cdots \right) dy \\ &= f(x) + \frac{1}{2}\sigma^2 f''(x) + \mathcal{O}[\sigma^4].\end{aligned} \tag{D.9}$$

Thus the expected value of $g(x)$ differs from $f(x)$ by a term of order σ^2; this suggests that σ should be as small as possible.

In order to calculate the expected square deviation from $f(x)$, it is first necessary to evaluate the mean square weight:

$$\left\langle w(x_i, x)^2\right\rangle_{x_i} = \int \frac{e^{-(y-x)^2/\sigma^2}}{2\pi\sigma^2} f(y)\, dy. \tag{D.10}$$

Expanding $f(y)$ about $y = x$ as before gives

$$\left\langle w(x_i, x)^2\right\rangle_{x_i} = \frac{1}{2\sigma\sqrt{\pi}} f(x) + \frac{\sigma}{8\sqrt{\pi}} f''(x) + \mathcal{O}[\sigma^3]. \tag{D.11}$$

This, together with equation (D.9), can now be substituted into equation (D.6):

$$\begin{aligned}\left\langle \Big(g(x) - f(x)\Big)^2 \right\rangle_{\{x_i\}} &= \left(\frac{1}{2}\sigma^2 f''(x) + \mathcal{O}[\sigma^4] \right)^2 \\ &\quad + \frac{1}{N} \left\{ \left(\frac{1}{2\sigma\sqrt{\pi}} f(x) + \frac{\sigma}{8\sqrt{\pi}} f''(x) + \mathcal{O}[\sigma^3] \right) \right. \\ &\quad \left. - \left(f(x) + \frac{1}{2}\sigma^2 f''(x) + \mathcal{O}[\sigma^4] \right)^2 \right\}.\end{aligned} \tag{D.12}$$

When the number of samples N is finite, this function diverges as $\sigma \to 0$ and has a minimum at some non-zero value of σ. The optimum value of σ decreases as N increases.

D.2 Projection onto basis functions

The second approach to the problem proceeds by writing $f(x)$ in terms of some appropriate[1] basis:

$$f(x) = \sum_{i=1}^{\infty} c_i \chi_i(x) \tag{D.13}$$

where $\{\chi_i\}$ is a set of complete orthonormal functions[2] and

$$c_i = \int f(x)\chi_i(x)\, dx. \tag{D.14}$$

An approximation to c_i may be obtained from the set of sampled points $\{x_i\}$:

$$d_i = \frac{1}{N} \sum_{j=1}^{N} \chi_i(x_j). \tag{D.15}$$

The expectation value of d_i is

$$\left\langle d_i \right\rangle = \left\langle \chi_i(x_j) \right\rangle_{x_j} = c_i. \tag{D.16}$$

Using these approximate coefficients, an attempt at a reconstructed function is then

$$\begin{aligned} g(x) &= \sum_{i=1}^{N_\chi} d_i \chi_i(x) \\ &= \frac{1}{N} \sum_{i=1}^{N_\chi} \sum_{j=1}^{N} \chi_i(x_j)\chi_i(x). \end{aligned} \tag{D.17}$$

Only a finite number N_χ of basis functions has been used. Comparison with equation (D.1) shows that the technique discussed in the previous section is approximately a special case of this more general method, with

$$w(x_j, x) = \sum_{i=1}^{N_\chi} \chi_i(x_j)\chi_i(x). \tag{D.18}$$

[1] The choice of basis is of course influenced by whatever prior knowledge of f is available: if f has a certain symmetry, then only basis functions with the same symmetry need be considered. Likewise, if f satisfies certain boundary conditions, the basis functions can be chosen to satisfy the same constraints.

[2] The basis functions in fact need only be linearly independent, not strictly orthogonal. However, the analysis is much simpler in the case of an orthonormal basis.

However, there are some weight functions which cannot be reproduced as long as N is finite.

The expected value of the reconstructed function is

$$\begin{aligned}\big\langle g(x)\big\rangle_{\{x_j\}} &= \sum_{i=1}^{N_\chi}\big\langle \chi_i(x_j)\big\rangle_{x_j}\chi_i(x)\\ &= \sum_{i=1}^{N_\chi} c_i\chi_i(x).\end{aligned} \tag{D.19}$$

This is a good approximation if $c_i = 0$ for $i > N_\chi$.

Finally, the expected square deviation from the original function is

$$\begin{aligned}\left\langle \Big(g(x)-f(x)\Big)^2\right\rangle_{\{x_j\}} &= \frac{1}{N^2}\sum_{i=1}^{N_\chi}\sum_{j=1}^{N}\sum_{k=1}^{N_\chi}\sum_{l=1}^{N}\chi_i(x)\chi_k(x)\big\langle\chi_i(x_j)\chi_k(x_l)\big\rangle_{x_j,x_l}\\ &\quad - 2f(x)\big\langle g(x)\big\rangle_{\{x_i\}} + f(x)^2\\ &= \frac{1}{N}\sum_{i=1}^{N_\chi}\sum_{k=1}^{N_\chi}\chi_i(x)\chi_k(x)\big\langle\chi_i(x_j)\chi_k(x_j)\big\rangle_{x_j}\\ &\quad + \frac{N-1}{N}\sum_{i=1}^{N_\chi}\sum_{k=1}^{N_\chi}\chi_i(x)\chi_k(x)\big\langle\chi_i(x_j)\big\rangle_{x_j}\big\langle\chi_k(x_j)\big\rangle_{x_j}\\ &\quad - 2f(x)\big\langle g(x)\big\rangle_{\{x_i\}} + f(x)^2\\ &= \Big(\big\langle g(x)\big\rangle_{\{x_i\}} - f(x)\Big)^2 + \frac{1}{N}\sum_{i=1}^{N_\chi}\sum_{k=1}^{N_\chi}\chi_i(x)\chi_k(x)\\ &\quad \times\Big(\big\langle\chi_i(x_j)\chi_k(x_j)\big\rangle_{x_j} - \big\langle\chi_i(x_j)\big\rangle_{x_j}\big\langle\chi_k(x_j)\big\rangle_{x_j}\Big).\end{aligned} \tag{D.20}$$

This is analogous to equation (D.6). The first term is large when an insufficient number of basis functions are used (compare equation (D.19)). However, the second term becomes large when *too many* functions are used. To see this, consider what happens as $N_\chi \to \infty$. The second term becomes

$$\begin{aligned}&\frac{1}{N}\int f(y)\sum_{i=1}^{\infty}\chi_i(x)\chi_i(y)\left\{\sum_{k=1}^{\infty}\chi_k(x)\chi_k(y) - \int f(z)\sum_{k=1}^{\infty}\chi_k(x)\chi_k(z)\,dz\right\}dy\\ &\qquad = \frac{1}{N}\int f(y)\delta(x-y)\Big(\delta(x-y) - \int f(z)\delta(x-z)\,dz\Big)\,dy,\end{aligned} \tag{D.21}$$

which is infinite.

In general, there is an optimum number of functions to use; when more points have been sampled, more functions should be used. Ideally the coefficients for the neglected functions (c_i, where $i = N_\chi + 1, \ldots, \infty$) should be very small; for this reason, it is important to choose a sensible basis. The best basis is one in which all the coefficients are zero above some cut-off.

Bibliography

[1] P. H. Acioli and D. M. Ceperley. Diffusion Monte Carlo study of jellium surfaces: Electronic densities and pair correlation functions. *Physical Review B*, 54(23):17199, 1996.

[2] S. A. Alexander, R. L. Coldwell, H. J. Monkhorst, and J. D. Morgan III. Monte Carlo eigenvalue and variance estimates from several functional optimizations. *Journal of Chemical Physics*, 95(9):6622, 1991.

[3] L. M. Almeida, J. P. Perdew, and C. Fiolhais. Surface and curvature energies from jellium spheres: Density functional hierarchy and quantum Monte Carlo. *Physical Review B*, 66:075115, 2002.

[4] N. W. Ashcroft and N. D. Mermin. *Solid State Physics*. Holt, Reinhardt and Winston, New York, 1976.

[5] G. B. Bachelet and M. Schlüter. Relativistic norm-conserving pseudopotentials. *Physical Review B*, 25(4):2103, 1982.

[6] A. Baldereschi. Mean-value point in the brillouin zone. *Physical Review B*, 7(12):5212, 1973.

[7] J. H. Bartlett, Jr., J. J. Gibbons, Jr., and C. G. Dunn. The normal helium atom. *Physical Review*, 47:679, 1935.

[8] A. D. Becke. Density-functional exchange-energy approximation with correct asymptotic behavior. *Physical Review A*, 38(6):3098, 1988.

[9] D. Bohm and D. Pines. A collective description of electron interactions: III. Coulomb interactions in a degenerate electron gas. *Physical Review*, 92(3):609, 1953.

[10] J. M. Borwein and P. B. Borwein. *Pi and the AGM: A Study in Analytic Number Theory and Computational Complexity*. Canadian Mathematical Society series of monographs and advanced texts. Wiley, New York, 1987.

[11] D. Bressanini, G. Morosi, and M. Mella. Robust wave function optimization procedures in quantum Monte Carlo methods. *Journal of Chemical Physics*, 116(13):5345, 2002.

[12] D. M. Ceperley and B. J. Alder. Ground state of the electron gas by a stochastic method. *Physical Review Letters*, 45(7):566, 1980.

[13] R. J. Coldwell. *International Journal of Quantum Chemistry. Quantum Chemistry Symposium*, 11:215, 1977.

[14] H. Conroy. Molecular Schrödinger equation. II. Calculation of ground-state energies by extrapolation. *Journal of Chemical Physics*, 41(5):1336, 1964.

[15] H. Conroy. Molecular Schrödinger equation. II. Monte Carlo evaluation of integrals. *Journal of Chemical Physics*, 41(5):1331, 1964.

[16] H. Conroy. Molecular Schrödinger equation. II. Properties of the energy variance function: The estimation of energy eigenvalues. *Journal of Chemical Physics*, 47(3):930, 1967.

[17] S. Doniach and E. H. Sondheimer. *Green's functions for solid state physicists*. Imperial College, London, 1998.

[18] R. M. Dreizler and E. K. U. Gross. *Density Functional Theory: An Approach to the Quantum Many-Body Problem*. Springer-Verlag, 1990.

[19] A. G. Eguiluz and W. Hanke. Evaluation of the exchange-correlation potential at a metal surface from many-body perturbation theory. *Physical Review B*, 39(14):10433, 1989.

[20] A. G. Eguiluz, M. Heinrichsmeier, A. Fleszar, and W. Hanke. First-principles evaluation of the surface barrier for a kohn-sham electron at a metal surface. *Physical Review Letters*, 68(9):1359, 1992.

[21] S. Fahy, X. W. Wang, and S. G. Louie. Variational quantum Monte Carlo nonlocal pseudopotential approach to solids: Formulation and application to diamond, graphite, and silicon. *Physical Review B*, 42:3503–3522, 1990.

[22] C. Filippi and C. J. Umrigar. Multiconfiguration wave functions for quantum Monte Carlo calculations of first-row diatomic molecules. *Journal of Chemical Physics*, 105(1):213, 1996.

[23] W. M. C. Foulkes, R. Q. Hood, and R. J. Needs. Symmetry constraints and variational principles in diffusion quantum Monte Carlo calculations of excited-state energies. *Physical Review B*, 60(7):4558, 1999.

[24] W. M. C. Foulkes, L. Mitas, R. J. Needs, and G. Rajagopal. Quantum Monte Carlo simulations of solids. *Reviews of Modern Physics*, 73(1):33, 2001.

[25] L. M. Fraser, W. M. C. Foulkes, G. Rajagopal, R. J. Needs, S. D. Kenny, and A. Williamson. Finite-size effects and Coulomb interactions in quantum Monte Carlo calculations for homogeneous systems with periodic boundary conditions. *Physical Review B*, 53(4):1814, 1996.

[26] P. Garcia-Gonzalez. Program for density-functional theory calculations of jellium slabs (**dftsurf3**), 2001.

[27] P. Garcia-Gonzalez and R. W. Godby. Self-consistent calculation of total energies of the electron gas using many-body perturbation theory. *Physical Review B*, 63:075112, 2001.

[28] R. Gaudoin, M. Nekovee, W. M. C. Foulkes, R. J. Needs, and G. Rajagopal. Inhomogeneous random-phase approximation and many-electron trial wave functions. *Physical Review B*, 63:115115–115130, 2001.

[29] O. Gunnarsson and B. I. Lundqvist. Exchange and correlation in atoms, molecules, and solids by the spin-density-functional formalism. *Physical Review B*, 13(10):4274–4298, 1976.

[30] B. L. Hammond, W. A. Lester, Jr., and P. J. Reynolds. *Monte Carlo Methods in Ab Initio Quantum Chemistry*. World Scientific, Singapore, 1994.

[31] A. Harju, B. Barbiellini, S. Siljamäki, R. M. Nieminen, and G. Ortiz. Stochastic gradient approximation: An efficient method to optimize many-body wave functions. *Physical Review Letters*, 79(7):1173, 1997.

[32] J. E. Harriman. Orthonormal orbitals for the representation of an arbitrary density. *Physical Review A*, 24(2):680, 1981.

[33] D. R. Hartree. *Proceedings of the Cambridge Philosophical Society*, 24:89, 1927.

[34] D. R. Hartree. Approximate wave functions and atomic field for mercury. *Physical Review*, 46:738–743, 1934.

[35] P. Hohenberg and W. Kohn. Inhomogeneous electron gas. *Physical Review*, 136:B864, 1964.

[36] C.-J. Huang, C. J. Umrigar, and M. P. Nightingale. Accuracy of electronic wave functions in quantum Monte Carlo: The effect of high-order correlations. *Journal of Chemical Physics*, 107(8):3007, 1997.

[37] T. Kato. *Communications on Pure and Applied Mathematics*, 10:151, 1957.

[38] S. D. Kenny, G. Rajagopal, R. J. Needs, W.-K. Leung, M. J. Godfrey, A. J. Williamson, and W. M. C. Foulkes. Quantum Monte Carlo calculations of the energy of the relativistic homogeneous electron gas. *Physical Review Letters*, 77(6):1099, 1996.

[39] P. R. C. Kent, R. Q. Hood, A. Williamson, R. J. Needs, W. M. C. Foulkes, and G. Rajagopal. Finite-size errors in quantum many-body simulations of extended systems. *Physical Review B*, 59(3):1917, 1999.

[40] P. R. C. Kent, R. J. Needs, and G. Rajagopal. Monte Carlo energy and variance-minimization techniques for optimizing many-body functions. *Physical Review B*, 59(19):12344, 1999.

[41] W. Kohn and L. Sham. Self-consistent equations including exchange and correlation effects. *Physical Review*, 140:A1133, 1965.

[42] E. Krotscheck. Theory of inhomogeneous quantum systems. III. Variational wave functions for fermi fluids. *Physical Review B*, 31(7):4267, 1985.

[43] E. Krotscheck, W. Kohn, and G.-X. Qian. Theory of inhomogeneous quantum systems. IV. Variational calculations of metal surfaces. *Physical Review B*, 32(9):5693, 1985.

[44] S. Kurth and J. P. Perdew. Density-functional correction of random-phase-approximation correlation with results for jellium surface energies. *Physical Review B*, 59:10461, 1999.

[45] Y. Kwon, D. M. Ceperley, and R. M. Martin. Effects of three-body and backflow correlations in the two-dimensional electron gas. *Physical Review B*, 48(16):12037, 1993.

[46] Y. Kwon, D. M. Ceperley, and R. M. Martin. Effects of backflow correlation in the three-dimensional electron gas: Quantum Monte Carlo study. *Physical Review B*, 58(11):6800, 1998.

[47] N. D. Lang and W. Kohn. Theory of metal surfaces: Charge density and surface energy. *Physical Review B*, 1(12):4555, 1970.

[48] N. D. Lang and W. Kohn. Theory of metal surfaces: Induced surface charge and image potential. *Physical Review B*, 7(8):3541, 1973.

[49] D. C. Langreth and M. J. Mehl. Easily implementable nonlocal exchange-correlation functional. *Physical Review Letters*, 47(6):446, 1981.

[50] D. C. Langreth and M. J. Mehl. Beyond the local-density approximation in calculations of ground-state electronic properties. *Physical Review B*, 28(4):1809, 1983.

[51] M. Levy. Universal variational functionals of electron densities, first-order density matrices, and natural spin-orbitals and solution of the v-representability problem. *Proceedings of the National Academy of Sciences of the United States of America*, 76:6062–6065, 1979.

[52] M. Levy. Electron densities in search of Hamiltonians. *Physical Review A*, 26:1200–1208, 1982.

[53] X. P. Li, R. J. Needs, R. M. Martin, and D. M. Ceperley. Green's-function quantum Monte Carlo study of a jellium surface. *Physical Review B*, 45(11):6124, 1992.

[54] E. H. Lieb. Density functionals for Coulomb systems. *International Journal of Quantum Chemistry*, 24(3):243–277, 1983.

[55] C. Lin, F.-H. Zong, and D. M. Ceperley. Twist-averaged boundary conditions in continuum quantum Monte Carlo. *Physical Review E*, 64:16702, 2001.

[56] A. Malatesta, S. Fahy, and G. B. Bachelet. Variational quantum Monte Carlo calculation of the cohesive properties of cubic boron nitride. *Physical Review B*, 56:12201–12210, 1997.

[57] A. Messiah. *Quantum Mechanics*, volume II. North-Holland, Amsterdam, 1961.

[58] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. Equation of state calculations by fast computing machines. *Journal of Chemical Physics*, 21:1087, 1953.

[59] H. J. Monkhorst and J. D. Pack. Special points for brillouin-zone integrations. *Physical Review B*, 13(12):5188, 1976.

[60] J. W. Moskowitz, K. E. Schmidt, M. A. Lee, and M. H. Kalos. A new look at correlation energy in atomic and molecular systems. II. The application of the Green's function Monte Carlo method to Li. *Journal of Chemical Physics*, 77(1):5188, 1976.

[61] R. J. Needs, M. D. Towler, N. D. Drummond, and P. R. C. Kent. *CASINO Version 1.6.2 User Manual*. University of Cambridge, 2003.

[62] E. Oran Brigham. *The Fast Fourier Transform and its Applications*. Prentice-Hall, New Jersey, 1988.

[63] G. Ortiz and P. Ballone. Correlation energy, structure factor, radial distribution function, and momentum distribution of the spin-polarized uniform electron gas. *Physical Review B*, 50(3):1391, 1994.

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

[65] D. E. Parry. The electrostatic potential in the surface region of an ionic crystal. *Surface Science*, 49:433, 1975.

[66] D. E. Parry. The electrostatic potential in the surface region of an ionic crystal: Erratum. *Surface Science*, 54:195, 1976.

[67] J. P. Perdew, K. Burke, and M. Ernzerhof. Generalized gradient approximation made simple. *Physical Review Letters*, 77(18):3865, 1996.

[68] J. P. Perdew, S. Kurth, A. Zupan, and P. Blaha. Accurate density functional with correct formal properties: A step beyond the generalized gradient approximation. *Physical Review Letters*, 82(12):2544, 1999.

[69] J. P. Perdew and Y. Wang. Accurate and simple representation of the electron-gas correlation energy. *Physical Review B*, 45(23):13244, 1992.

[70] J. P. Perdew and A. Zunger. Self-interaction correction to density-functional approximations for many-electron systems. *Physical Review B*, 23:5048, 1981.

[71] J. M. Pitarke. Comment on 'Diffusion Monte Carlo study of jellium surfaces: Electronic densities and pair correlation functions'. *Physical Review B*, 70:087401, 2004.

[72] J. M. Pitarke and A. G. Eguiluz. Jellium surface energy beyond the local-density approximation: Self-consistent-field calculations. *Physical Review B*, 63:045116, 2001.

[73] S. Raimes. *Many-electron Theory*. North-Holland, Amsterdam-London, 1972.

[74] G. Rajagopal, R. J. Needs, S. Kenny, W. M. C. Foulkes, and A. James. Quantum Monte Carlo calculations for solids using special k points methods. *Physical Review Letters*, 73(14):1959, 1994.

[75] P. J. Reynolds, D. M. Ceperley, B. J. Alder, and W. A. Lester, Jr. Fixed-node quantum Monte Carlo for molecules. *Journal of Chemical Physics*, 77(11):5593, 1982.

[76] K. E. Schmidt and J. W. Moskowitz. Correlated Monte Carlo wave functions for the atoms He through Ne. *Journal of Chemical Physics*, 93(6):4172, 1990.

[77] F. Sottile and P. Ballone. Fixed-node diffusion Monte Carlo computations for closed-shell jellium spheres. *Physical Review B*, 64:045105, 2001.

[78] M. P. Tosi. *Solid State Physics: Advances in Research and Applications*, volume 16. Academic, London, 1965.

[79] C. J. Umrigar, M. P. Nightingale, and K. J. Runge. A diffusion Monte Carlo algorithm with very small time-step errors. *Journal of Chemical Physics*, 99(4):2865, 1993.

[80] C. J. Umrigar, K. G. Wilson, and J. W. Wilkins. Optimized trial wave functions for Quantum Monte Carlo calculations. *Physical Review Letters*, 60(17):1719, 1988.

[81] S. V. Vonsovsky and M. I. Katsnelson. *Quantum Solid-State Physics*, volume 73 of *Solid-State Sciences*. Springer-Verlag, Berlin, 1989.

[82] D. H. Weinstein. Modified Ritz method. *Proceedings of the National Academy of Sciences of the United States of America*, 20(9):529, 1934.

[83] A. Williamson, S. D. Kenny, G. Rajagopal, A. J. James, R. J. Needs, L. M. Fraser, W. M. C. Foulkes, and P. Maccullum. Optimized wave functions for quantum Monte Carlo studies of atoms and solids. *Physical Review B*, 53(15):9640, 1996.

[84] A. Williamson, G. Rajagopal, R. J. Needs, L. M. Fraser, W. M. C. Foulkes, Y. Wang, and M. Y. Chou. Elimination of Coulomb finite-size effects in quantum many-body simulations. *Physical Review B*, 55(8):R4851, 1997.

[85] Z. Yan, J. P. Perdew, S. Kurth, C. Fiolhais, and L. Almeida. Density-functional versus wave-function methods: Towards a benchmark for the jellium surface energy. *Physical Review B*, 61(4):2595, 2000.