New *ab initio* formulation of electron correlation and spin resonance

by

Gábor Csányi

Submitted to the Department of Physics
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Physics

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2001

Author ....
Department of Physics
August 6, 2001

Certified by ....
Tomás A. Arias
Professor, Cornell University
Thesis Supervisor

Certified by ....
John D. Joannopoulos
Francis D. Wright Professor of Physics
Thesis Supervisor

Accepted by ....
Thomas J. Greytak
Associate Department Head for Education

New *ab initio* formulation of electron correlation and spin resonance

by

Gábor Csányi

Submitted to the Department of Physics
on August 6, 2001, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Physics

Abstract

We explore a new class of computationally feasible approximations of the two–body density matrix as a finite sum of tensor products of single–particle operators. Physical symmetries then uniquely determine the two–body matrix in terms of the one–body matrix. Representing dynamical correlation alone as a single tensor product results in a theory which predicts near zero dynamical correlation in the homogeneous electron gas at moderate to high densities. But, representing both dynamical and statistical correlation effects together as a tensor product leads to the recently proposed "natural orbital functional." We find that this latter theory has some asymptotic properties consistent with established many–body theory but is no more accurate than Hartree–Fock in describing the homogeneous electron gas for the range of densities typically found in the valence regions of solids. We present a new density–matrix functional within this tensor–product framework which performs well both for the homogeneous electron gas as well as atoms. For the electron gas, it is very accurate at high densities, and comparable to Hartree–Fock at metallic valence densities. For atoms it is on par with previous density–matrix functionals. We also present analytic results for the correlation energy in the low density limit of the free electron gas for a broad class of such functionals.

We investigate the ways in which first principles computational methods can be used to interpret Electron Spin Resonance experiments. Based on *ab initio* calculation, we propose a new structure for the fundamental excitation of the reconstructed 30° partial dislocation in silicon. This soliton has a rare structure involving a five–fold coordinated atom near the dislocation core. The unique electronic structure of this defect is consistent with the electron spin resonance signature of the hitherto enigmatic thermally stable R center of plastically deformed silicon. We present the first *ab initio* determination of the free energy of the soliton, which is also in agreement with the experimental observation. This identification suggests the possibility of an experimental determination of the density of solitons, a key defect in understanding the plastic flow of the material.

The recently introduced Projector Augmented Wave (PAW) method is used to compute hyperfine coupling constants for the azafullerene $C_{59}N$. The results are in much better agreement with experiment than previous all–electron calculations. We emphasize the utility of the PAW method as a *post processing* step used in conjunction with traditional plane wave pseudopotential density functional methods. We also find that the previously employed method of obtaining hyperfine coupling constants by rescaling the pseudo density is not sufficiently accurate in this case and the full PAW projector formalism has to be used. We demonstrate the numerical difficulties in the case of hydrogenated pyridine.

In the last chapter we collect the ingredients from the literature for the calculation of the g tensor in extended systems. We show how the the calculation can be carried out within the context of the PAW formalism and test the required approximations by calculating the spin–orbit splitting in bulk silicon and germanium.

Thesis Supervisor: Tomás A. Arias
Title: Professor, Cornell University

Thesis Supervisor: John D. Joannopoulos
Title: Francis D. Wright Professor of Physics

Acknowledgments

The author would like to extend his sincerest thanks to Professor Tomas A. Arias, who guided him in his endeavors throughout the last five years. The help and support of Sohrab Ismail–Beigi was invaluable as a colleague, office mate and friend, while the camaraderie of fellow students Torkel Engeness, Dicle Yesilleten, Darren Segall and Jason Cline was exceptional. Thanks are due to the Cornell crew Ivan Daykov, Hande Ustunel, Daniel Freedman, Sahak Petrosyan and Anjali Gopalakrishnan who made the last frantic months actually enjoyable. Michelle Povinelli was my staunch ally concerning all matters in the Physics Department, ranging from the never materializing "student couch" to the much more tangible establishment of the CMT Journal Club. The number of friends, both in Boston and around the world, who helped keep me sane and not loose sight of the yellow brick road, is too great to list them all here, but they cannot go unacknowledged, at least collectively. And last but not least, none of this would have been possible without the companionship of my beloved wife, Ágnes Borszéki.

Contents

List of Figures

List of Tables

Introduction

This thesis falls into the broad field of Computational Physics. Although in most circles the subject has gained solid acceptance in recent years alongside more traditional branches of physics, namely Experimental and Theoretical, it is felt that some insight might be gained by briefly reexamining its role. Afterall, if the computer is used merely to save time and paper in solving the same equations that a traditional Theoretician might solve, the notion of a separate field is hardly justified.

A separate field emerges when not just radically different *types* of problems can be attacked with new methods, but when the types of *questions* asked are different. In traditional theoretical physics, mathematical models are built and their predictions tested against the experiments. One model is better than another, if it is simpler, while reproducing the same results. The notion of "simpler" here means fewer parameters and variables. Understanding and insight are gained by figuring out which models work to within the desired accuracy. This then shows which (few) parameters and variables are the important ones. In contrast, computational physics often takes a different approach. The models there tend to be "simple" another way: as close as possible to the elementary interactions of the components of the physical system, but with many variables.

The best example is from statistical physics. If we are to regard a collection of interacting spins, the theoretical approach is to define order parameters, and use the renormalization method to say something about the macroscopic quantities of interest. In the computational approach, the model consists of the variables corresponding to the individual spins and their interactions. The Lagrange equations could then be integrated with various methods, and macroscopic quantities *measured*. Of course this measurement is not to be confused with the measurement one makes in an experiment. The computational approach still uses a *model*, which is simpler than the real system. In a ferromagnetic material for example, the real magnetic interactions are very complicated, the full quantum mechanical Hamiltonian is by no means just a spin–spin interaction. So in this sense, the computational model is in between the real system and the macroscopic (theoretical) model. Its main utility is that often, approximating a physical system with elementary microscopic interactions is much more straightforward than coming up with macroscopic models for it that work well.

The question what insight is gained by the use of the computational model is a crucial one. The answer is twofold. First, just as with any model, the validation of the model as a "simpler system" gives physical insight. This means figuring out which kind of elementary interactions are necessary to predict a certain experiment. Secondly, once we are satisfied that the necessary physics is captured in our computational model, making the connection to macroscopic quantities is much easier than in a real experiment. In the computer, the system under study can be isolated, measurements can be made without interference from other effects that are necessarily part of a real experiment. Thus connections can be made in both directions. A computational measurement can help to identify plausible microscopic underpinnings of real experimental results, and also be able to test macroscopic, theoretical models much more efficiently.

In the last fifty years, enormous progress has been made in developing microscopic models of materials. In light of the preceding discussion, particular importance is attached to so called *ab initio* methods. Computational models in this class aim to use, as their basic microscopic model, the full quantum mechanical interactions that are present in material systems. Of course, this is still "just" a model with respect to real materials, but a very good one. Usually interactions below a certain length scale are not modeled explicitly, for example, to figure out the crystal structure of a material, it is widely believed that the super–string nature of the elementary particles can be altogether ignored. Given a level of quantum mechanical description, say, electrons and the Coulomb potential of the nuclei, the microscopic model is fixed. It then remains to extract the desired observables from this model.

In practice, numerical approximations must be made in the above approach. Clearly, no computer is large enough to hold the information that is needed to describe all the electrons and nuclei in a macroscopic material, not to mention the processing requirements to extract observables. In fact, the amount of information that would be required to model the full many–body quantum mechanics of even a tiny amount of material (say a few thousand atoms) is incomprehensibly enormous. Fortunately, this is not necessary. The terms in the many–body Hamiltonian are at most two–body interactions, and the formalism can be exactly reduced to deal only with two–body wavefunctions and operators. Furthermore, in arguably the seminal papers in this field, Hohenberg, Kohn and Sham showed that ground state properties can be extracted from a one–body description. The resulting *Density Functional Theory* forms the basis of modern quantum mechanical calculations.

At any point in the development of a scientific field, two major pathways are open to investigation. One is to apply the present models to new problems, the other is to try to refine and improve upon the models themselves. Of course, in the long run, these endeavours go hand in hand, the refinements becoming necessary when the old models are not good enough for a new application. Accordingly, this thesis is in two parts. The first two chapters outline a new method for calculating total energies using a *density–matrix functional.* Instead of reducing the two–body Hamiltonian problem to a minimization over physical charge densities, as in the Hohenberg–Kohn–Sham theory, these methods use the one–body density matrix as the basic variable. There has been a recent surge of interest in such methods, and here we provide a total energy density–matrix functional that competes with the best available traditional density functionals. We also outline a formalism and strategy to construct such functionals in general, and show how previous proposals all fit into this formalism.

The second part of the thesis, chapters three through five, explores how current density functional theory can be used to predict parameters measured in Electron Spin Resonance (ESR) experiments. In chapter three, we use the symmetry of the electronic state associated with an unpaired electron to identify and elucidate the structure of a point defect in plastically deformed silicon, which had hitherto not been understood. In chapter four, we turn to experimentally directly observable parameters, the hyperfine splitting of the ESR spectrum. The application here is the $C_{59}N$ molecule. The plane wave pseudopotential method, in conjunction with the recently introduced Projector Augmented Wave (PAW) method used as a *post processing* step is shown to give excellent results for the Fermi contact interaction, enabling an identification of the hyperfine lines with particular atoms in the molecule. In the last chapter, we collect from the literature the ingredients for a first principles calculation of the g tensor. Computing the matrix elements of the spin–orbit Hamiltonian is the first step towards calculating the g tensor, and we show how this can be accomplished within the PAW formalism. We test the necessary approximations by computing spin–orbit splittings in silicon and germanium. We then outline, for the first time, the necessary formalism to complete an *ab initio* calculation of the g tensor in extended systems.

Chapter 1

Tensor product expansions for correlation in quantum many–body systems

The fundamental difficulties associated with *ab initio* solutions of the quantum mechanical many–body problem stem directly from the large dimensionality of the wave function. Rather than dealing directly with the many–body wave function, large scale electronic structure calculations determine the ground–state energy by minimizing an energy functional over the much more manageable space of single–particle orbitals. Hohenberg–Kohn–Sham theory [1, 2] places this approach on a firm theoretical footing by establishing that exact electronic ground-state energies can be determined in this manner by minimizing a universal functional of such orbitals.

Although the requisite universal energy functional is not known exactly, the local density approximation[2] and various improvements thereon [3, 4, 5, 6, 7] are sufficient to resolve bond energies to a little better than one–tenth of an electron Volt. With this accuracy, these functionals make possible highly predictive first principles studies in diverse areas such as the study of surfaces, point defects, plastic deformation and chemical reactions. (For a review see[8].) Although sufficient for many studies, this error is too large to allow accurate prediction of the rates of microscopic processes at room temperature, thus limiting the ultimate predictive power of the *ab initio* density–functional approach. Whereas further improvement of energy functionals for single–particle theories is an active area of research, much less work has been done to construct energy functionals of the one–body density matrix[3, 9]. Such functionals are promising because the density matrix contains more explicit information than do the Kohn–Sham orbitals, and an accurate energy functional thus is likely to be simpler in form. The exact kinetic energy functional, for instance, is known for the density matrix, but not the Kohn–Sham orbitals.

To date, the only extensively used density–matrix energy functional is the Hartree–Fock approximation. One may view Hartree–Fock theory as *approximating* the two–body matrix as a sum of Hartree and exchange terms, each of which is a tensor product of the one–body density matrix. Hartree–Fock then uses the known, *exact* energy functional of the two–body density matrix to produce an energy functional of the one–body density matrix. In this work, we build upon this perspective and consider tensor product approximations to the two–body density matrix which not only serve as generators of energy functionals of the one–body density matrix but also provide estimates of the two–body matrix in terms of the one–body matrix and thereby shed light on the nature of correlation in quantum many–body systems.

1.1 Notation

For a system of N electrons, the exact energy functional of the two–body density matrix $\gamma(\mathbf{x}_1\mathbf{x}_2, \mathbf{x}'_1\mathbf{x}'_2) \equiv \langle\Psi|\hat{\psi}^\dagger(\mathbf{x}'_1)\hat{\psi}^\dagger(\mathbf{x}'_2)\hat{\psi}(\mathbf{x}_2)\hat{\psi}(\mathbf{x}_1)|\Psi\rangle/2$, is

$$\begin{aligned} E &= \int d\mathbf{x}_1 \left(\left[-\frac{1}{2}\nabla^2_{\mathbf{r}_1} + U(\mathbf{r}_1) \right] n(\mathbf{x}_1, \mathbf{x}'_1) \right) \Big|_{\mathbf{x}'_1=\mathbf{x}_1} \\ &\quad + \int d\mathbf{x}_1 d\mathbf{x}_2 \, \frac{\gamma(\mathbf{x}_1\mathbf{x}_2, \mathbf{x}_1\mathbf{x}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|}. \end{aligned} \tag{1.1}$$

Here and throughout, we work in atomic units, and the $\mathbf{x}_i$ are compound coordinates representing both position $\mathbf{r}_i$ and spin s_i so that integration over $\mathbf{x}_i$ represents integration over space and summation over spin channels. The external potential and one–body density matrix are $U(\mathbf{r})$ and $n(\mathbf{x}_1, \mathbf{x}'_1) \equiv \langle\Psi|\hat{\psi}^\dagger(\mathbf{x}'_1)\hat{\psi}(\mathbf{x}_1)|\Psi\rangle$, respectively. Finally, the one–body density matrix comes directly from γ through the sum–rule

$$\left(\frac{N-1}{2}\right) n(\mathbf{x}_1, \mathbf{x}'_1) = \int d\mathbf{x}_2 \, \gamma(\mathbf{x}_1\mathbf{x}_2, \mathbf{x}'_1\mathbf{x}_2). \tag{1.2}$$

1.2 Computational considerations

Although (1.1) is exact, the two–body density matrix γ is a function of four variables, and direct computation with such functions is unfeasible. Viable techniques, however, exist for dealing directly with two–variable functions such as the one–body density matrix[10, 11, 12, 13, 14, 15]. Accordingly, we expand γ in terms of two–variable functions,

$$\hat{\gamma} = \sum_i \hat{g}_i \otimes \hat{h}_i, \tag{1.3}$$

which, by separation of variables, is always possible. Here, $\hat{g}_i$ and $\hat{h}_i$ are one–body operators (functions of two variables), we view the two–body density matrix as a two–body operator (a function of four variables), and the tensor product denotes one of the three possible choices for separating four variables,

$$\begin{aligned}
\hat{g}\underset{\text{I}}{\otimes}\hat{h} &= g(\mathbf{x}_1,\mathbf{x}'_1)h(\mathbf{x}_2,\mathbf{x}'_2) \\
\hat{g}\underset{\text{II}}{\otimes}\hat{h} &= g(\mathbf{x}_1,\mathbf{x}_2)h(\mathbf{x}'_2,\mathbf{x}'_1) \\
\hat{g}\underset{\text{III}}{\otimes}\hat{h} &= g(\mathbf{x}_1,\mathbf{x}'_2)h(\mathbf{x}_2,\mathbf{x}'_1).
\end{aligned}$$

1.3 Quantum statistical considerations

We wish to maximize the physical content of a truncation of the expansion (1.3) to a limited number of terms. Physically, $n(\mathbf{x}_1,\mathbf{x}'_1)$ is the quantum amplitude for the insertion of a hole at position $\mathbf{x}_1$ and its *instantaneous* removal from $\mathbf{x}'_1$, whereas $\gamma(\mathbf{x}_1\mathbf{x}_2,\mathbf{x}'_1\mathbf{x}'_2)$ is the quantum amplitude for the insertion of a pair of holes at $\mathbf{x}_1,\mathbf{x}_2$ and their removal from $\mathbf{x}'_1,\mathbf{x}'_2$. Approximating the latter events as independent gives $\hat{\gamma} \approx \hat{\gamma}_H \equiv (\hat{n}\underset{\text{I}}{\otimes}\hat{n})/2$, the familiar Hartree approximation, where the factor of two maintains the normalization implicit in (1.2). We may then refine this mean–field behavior and define $\hat{\gamma} \equiv \hat{\gamma}_H + \hat{\gamma}_{xc}$, where $\hat{\gamma}_{xc}$ represents exchange and correlation effects which we then expand, without loss of generality, according to (1.3).

The most significant drawback of the Hartree approximation is that γ_H is not properly antisymmetric with respect to particle exchange $(\mathbf{x}'_1 \leftrightarrow \mathbf{x}'_2)$, so that, potentially, many terms will be required in $\hat{\gamma}_{xc}$ to restore this symmetry. Alternately, we may explicitly ensure antisymmetry and take $\hat{\gamma} = \hat{\gamma}_{HF} + \hat{\gamma}_c \equiv (\hat{n}\underset{\text{I}}{\otimes}\hat{n} - \hat{n}\underset{\text{III}}{\otimes}\hat{n})/2 + \hat{\gamma}_c$. In this form, $\hat{\gamma}_{HF}$ is precisely the familiar Hartree–Fock approximation. Again, we may expand the unknown $\hat{\gamma}_c$ as in (1.3).

Finally, we note that a fundamental difficulty exists when using a finite number of terms of Type II. Because of the instantaneous nature of the quantum event which γ represents, we expect for normal systems that $\gamma \to 0$ as the insertion $\mathbf{x}'_1\mathbf{x}'_2$ and removal $\mathbf{x}_1\mathbf{x}_2$ locations are placed at ever further distances from one another. In an extended system with the expansion for γ containing a finite sum of terms of Type II, taking this limit while keeping the insertion points near one another and the removal points near one another will violate this asymptotic condition unless each Type II term is identically zero. Accordingly, we focus below on approximating $\hat{\gamma}_{xc}$ and $\hat{\gamma}_c$ with terms of Type I and III.

1.4 Symmetry

Physical symmetries of $\hat{\gamma}$ significantly restrict the allowable forms for the tensor products appearing in expansions of the form (1.3) for $\hat{\gamma}_{xc}$ and $\hat{\gamma}_c$. We first consider Hermiticity $\gamma(\mathbf{x}_1\mathbf{x}_2, \mathbf{x}_1'\mathbf{x}_2') = \gamma^*(\mathbf{x}_1'\mathbf{x}_2', \mathbf{x}_1\mathbf{x}_2)$ and particle permutation symmetry $\gamma(\mathbf{x}_1\mathbf{x}_2, \mathbf{x}_1'\mathbf{x}_2') = \gamma(\mathbf{x}_2\mathbf{x}_1, \mathbf{x}_2'\mathbf{x}_1')$ for each of the three types of tensor product and defer discussion of Fermionic antisymmetry until after the sum–rule immediately below. Because both γ_H and γ_{HF} respect these symmetries, so must tensor products representing $\hat{\gamma}_{xc}$ and $\hat{\gamma}_c$. These symmetries imply relations with arbitrary constants of proportionality between the one–body operators of such products which may be absorbed into the one–body operators. Table (1.1) *summarizes* the final result of such considerations for all three tensor products. We note that in all cases symmetry restricts the tensor product to involve only a *single* one–body operator which remains to be determined.

Remarkably, when a single term satisfying Hermiticity and particle permutation is used to represent $\hat{\gamma}_{xc}$ or $\hat{\gamma}_c$, we can "invert" the sum–rule (1.2) and determine the two–body density matrix γ in terms of the one–body matrix n. Inserting the resulting form for γ into (1.1) then generates an energy functional of the one–body matrix which implicitly satisfies the sum–rule, an important property for obtaining good ground state energies.

When terms of Type I represent $\hat{\gamma}_{xc}$ or $\hat{\gamma}_c$, the sum rule becomes $\hat{g}(\mathrm{Tr}\,\hat{g}) = \hat{o}$, where $\hat{o} = \hat{n}$ or $\hat{n}(1-\hat{n})$, respectively. The solution of this equation is $\hat{g} = \pm\hat{o}/\sqrt{\mathrm{Tr}\,\hat{o}}$. In either case, the extensivity of $\mathrm{Tr}\,\hat{o}$ makes the resulting solution irrelevant in the thermodynamic limit $N \to \infty$. For products of Type III, the sum–rule combined with appropriate symmetries gives $\hat{g}^2 = \hat{o}$, so that $\hat{g} = \sqrt{\hat{o}}$ with $\hat{o}$ defined as above.

Fermionic antisymmetry, $\gamma(\mathbf{x}_1\mathbf{x}_2, \mathbf{x}_1'\mathbf{x}_2') = -\gamma(\mathbf{x}_1\mathbf{x}_2, \mathbf{x}_2'\mathbf{x}_1')$, is a stronger condition than the particle permutation symmetry condition considered above and imposes more complex constraints. Two–body density matrices γ which satisfy this condition may always be written with Type I and III products appearing in corresponding pairs of the form $\hat{g}\underset{\mathrm{I}}{\otimes}\hat{g} - \hat{g}\underset{\mathrm{III}}{\otimes}\hat{g}$ and with Type II products (were we to consider such) whose one–body terms are separately antisymmetric.

The two viable functionals remaining after the above considerations involve a Type III representation for either $\hat{\gamma}_{xc}$ or $\hat{\gamma}_c$. Unfortunately these forms do not consist of symmetric pairs of Type I and Type III products. To satisfy antisymmetry, one could take $\hat{\gamma}_{xc} = -(\hat{n}\underset{\mathrm{III}}{\otimes}\hat{n})/2$, so that $\hat{\gamma} = (\hat{n}\underset{\mathrm{I}}{\otimes}\hat{n} - \hat{n}\underset{\mathrm{III}}{\otimes}\hat{n})/2$, which is simply the Hartree–Fock approximation. To go beyond this, antisymmetry requires representing $\hat{\gamma}_c$ as a pair of terms so that $\hat{\gamma} = (\hat{n}\underset{\mathrm{I}}{\otimes}\hat{n} - \hat{n}\underset{\mathrm{III}}{\otimes}\hat{n} + \hat{g}\underset{\mathrm{I}}{\otimes}\hat{g} - \hat{g}\underset{\mathrm{III}}{\otimes}\hat{g})/2$.

The sum–rule for this latter extension on Hartree–Fock is $\hat{g}^2 - \hat{g}(\mathrm{Tr}\,\hat{g}) = \hat{n}(1 - \hat{n})$, which has no solution for $\hat{g}$ in extended systems unless $\mathrm{Tr}\,\hat{g}$ vanishes in the thermodynamic limit. Under this condition, we have the solution $\hat{g} = \pm\sqrt{\hat{n}(1-\hat{n})}$, where the signs of the eigenvalues in the square root must be chosen to ensure $\mathrm{Tr}\,\hat{g} = 0$. For the paramagnetic phase we consider below, the natural choice is to take the upper sign for the spin–up block of the density matrix and the lower sign for the spin–down block. The structure of the energy functional (1.1), however, is such that for this choice the resulting energy functional of $\hat{n}$ is equal to that generated by representing $\hat{\gamma}_c$ as a single Type III product.

1.5 Representation

The above discussion leaves the energy of two representations of the two–body matrix to explore: corrected Hartree theory, $\hat{\gamma} = \hat{\gamma}_H - (\sqrt{\hat{n}} \underset{\text{III}}{\otimes} \sqrt{\hat{n}})/2$, and corrected Hartree–Fock theory, $\hat{\gamma} = \hat{\gamma}_{HF} - (\sqrt{\hat{n}(1-\hat{n})} \underset{\text{III}}{\otimes} \sqrt{\hat{n}(1-\hat{n})})/2$. The corresponding energy functionals may be represented either directly in terms of the one–body density matrix for use with direct density-matrix methods[10, 11, 12, 13, 14, 15], or in terms of the spectral ("natural orbital") representation of the density matrix.

The energy functional (1.1) contains one–body terms which may be evaluated directly in terms of the one–body density matrix. The remaining two–body term may be computed in terms of the two–point density $c(\mathbf{r}_1, \mathbf{r}_2) \equiv \sum_{s_1 s_2} \gamma(\mathbf{x}_1\mathbf{x}_2, \mathbf{x}_1\mathbf{x}_2)$. When the above functionals represent paramagnetic states, this density has the following form in terms of the eigenvectors (natural orbitals) ϕ_i and corresponding eigenvalues (occupancies) f_i of the one–body density matrix $\hat{n}$,

$$\begin{aligned} c(\mathbf{r}_1, \mathbf{r}_2) \;=\; & \frac{1}{2}\Bigg[4\sum_{ij} f_i f_j \phi_i(\mathbf{r}_1)\phi_i^*(\mathbf{r}_1)\phi_j(\mathbf{r}_2)\phi_j^*(\mathbf{r}_2) \\ & -2\sum_{ij} g_{ij}\phi_i(\mathbf{r}_1)\phi_i^*(\mathbf{r}_2)\phi_j^*(\mathbf{r}_1)\phi_j(\mathbf{r}_2)\Bigg], \end{aligned} \tag{1.4}$$

where for corrected Hartree and Hartree–Fock theories, we take $g_{ij} = \sqrt{f_i f_j}$ and $g_{ij} = f_i f_j + \sqrt{f_i(1-f_i)f_j(1-f_j)}$, respectively.

The first of the above forms corresponds to the natural orbital functional proposed by Goedecker and Umrigar [3] as an *ansatz* among the many possible choices for the g_{ij} which satisfy the sum–rule. Here, we see that this form is the unique Type III correction to Hartree theory which satisfies basic symmetries and the sum–rule. This theory has been studied analytically in the case of the homogeneous electron gas in the low density regime ($r_s > 5.77$)[19]. Here, we present results for both this theory and our new corrected Hartree-Fock theory over a range of densities which includes both this low density regime and the regime more physically relevant in the valence region of solids, $r_s < 5.77$.

1.6 Homogeneous electron gas

Translational invariance in the homogeneous gas ensures that the natural orbitals are plane waves. The remaining freedom in $\hat{n}$ lies in the occupancy eigenvalues $f_{\mathbf{k}}$. For the paramagnetic phase, the energy (1.1) then takes the following form,

$$E = 2\int \frac{V d\mathbf{k}}{(2\pi)^3}\frac{\mathbf{k}^2}{2} f_{\mathbf{k}} - \frac{1}{V}\int \frac{V^2 d\mathbf{k} d\mathbf{k}'}{(2\pi)^6} g_{\mathbf{k}\mathbf{k}'} \frac{4\pi}{|\mathbf{k}' - \mathbf{k}|^2}, \tag{1.5}$$

where V is the volume of the system and the factor of two in the kinetic energy arises from the sum over spin.

The minimum of this energy functional occurs when $\delta(E - \mu N)/\delta f_{\mathbf{k}} = 0$. For both corrected Hartree and corrected Hartree–Fock, this gives to leading order as $|k| \to \infty$,

$$\delta(E - \mu N)/\delta f_{\mathbf{k}} \sim k^4 - C/\sqrt{f_{\mathbf{k}}},$$

where C is a constant. Hence, for large k, the momentum density scales as $f_{\mathbf{k}} \sim |\mathbf{k}|^{-8}$, in agreement with the random phase approximation and other many–body calculations[16]. This analysis holds for the present two functionals at all particle number densities and, in fact, for all theories of the form of (1.5) in which $g_{\mathbf{k}\mathbf{k}'} \to \sqrt{f_{\mathbf{k}} f_{\mathbf{k}'}}$ as $k \to \infty$ and $f_{\mathbf{k}} \to 0$.

1.7 Numerical Results

To determine the total energy of the homogeneous electron gas within the corrected Hartree and Hartree–Fock approximations, we exploit spherical symmetry and reduce $f_{\mathbf{k}}$ to a single variable function $f(k)$, which we represent on a radial mesh of variable spacing. We then minimize (1.5) with conjugate–gradients techniques subject to both the constraint that $f(k)$ corresponds to a total of N electrons and the Fermi constraint that $0 \leq f(k) \leq 1$. The latter constraint we implement by defining $f(k) \equiv (e^{-x}+1)^{-2}$, where x is then free to range over the real line.

To evaluate the energy functional in terms of the fillings, we integrate numerically the kinetic and exchange–correlation terms in (1.5). A key step in producing accurate results is to avoid numerical integration across the familiar logarithmic singularity at $k = k'$. For corrected Hartree theory, we avoided this by integrating the exchange–correlation terms once by parts analytically and then evaluating the result numerically. For the corrected Hartree–Fock case, we found it more efficient to approximate the Coulomb potential by a Yukawa potential, $e^{-\kappa r}/r$, and numerically extrapolate the results to $\kappa = 0$. For these latter calculations, we found it more efficacious to work in the grand canonical ensemble (holding the chemical potential fixed while minimizing $E - \mu N$) than to impose the particle–number constraint explicitly.

The above procedures introduce four sources of numerical error: the finite spacing of the radial mesh, the finite extent of the mesh in k–space, the error in the extrapolation in the case of using the Yukawa potential, and the finite number of iterations involved in our searches for the minima. Through analysis and numerical experiments, we determine these errors to be less than 10^{-3}, 10^{-4}, 10^{-3} and 10^{-6} Hartree per particle, respectively, over the range of densities which we explored $(0.01 < r_s < 30)$.

We found the minimization of the above functionals problematic for two reasons.

First and foremost, the discretized representation of the one–dimensional function $f(k)$ on the radial mesh artificially stabilizes any discontinuities in $f(k)$ which may arise during the minimization process. This may be managed by taking care to ensure that the radial mesh provides sufficient resolution to describe any large gradients which arise in $f(k)$. An additional factor making the calculations difficult is that both the corrected Hartree and corrected Hartree–Fock functionals suffer from a very large condition number (in excess of 10^7).

Figure (1-1) compares our results for the correlation energy per particle of the electron gas with established many–body predictions. We find that at very low densities, both new theories begin to approach one another as $r_s \to \infty$. For even moderately low densities ($6 < r_s$), both theories perform significantly worse than Hartree–Fock, over-estimating the correlation energy by factors from two to four and higher. For the densities experienced in the valence regions of solids ($1 < r_s < 6$), corrected Hartree–Fock theory (diamonds) crosses over from over–correlation to under–correlation, whereas corrected Hartree theory (circles) over–correlates by roughly the same amount by which traditional Hartree–Fock theory under–correlates. For higher densities more representative of atomic cores ($r_s < 1$), the corrected Hartree–Fock theory predicts near zero correlation. At these same densities, corrected Hartree theory remains over–correlated and gives a somewhat better estimate of the correlation energy than does traditional Hartree–Fock. The asymptotic form of the corrected Hartree results in the high density limit appears to be a straight line in our plot, which corresponds to the form of the leading order terms of the Gell–Mann–Brueckner expansion[20]. However, the prefactors which we find by a least squares fit, $E \approx 0.0570 \log(r_s) - 0.0714$, are only correct to within a little better than a factor of two: the well–known analytic values for these coefficients are 0.0311 and -0.0480, respectively. This suggests that the favorable reports for atoms for the natural orbital functional[3] may be in part a consequence of the high densities experienced in atomic cores, the limited particle

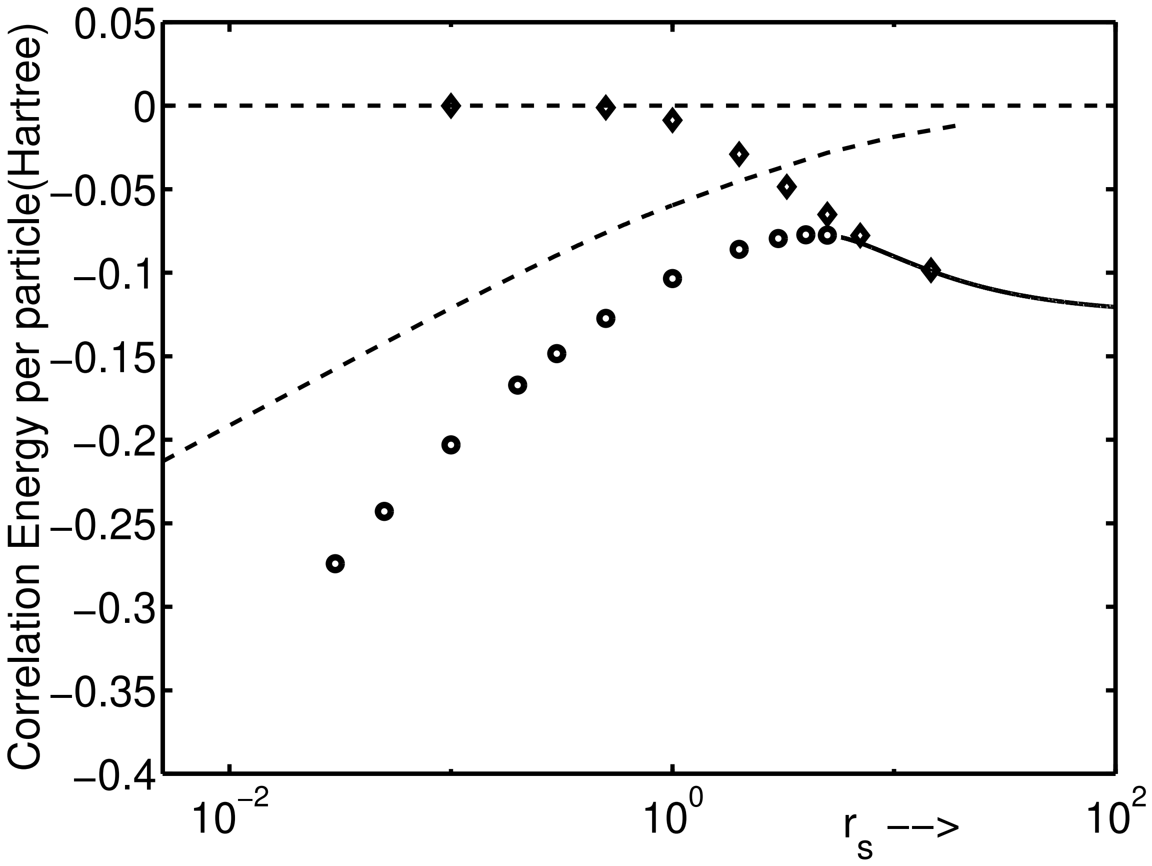

Figure 1-1: Correlation energy per particle of the uniform electron gas. Hartree–Fock (horizontal dashed line); exact many-body results[17, 18] (dashed curve), corrected Hartree theory (open circles), corrected Hartree–Fock theory (diamonds), analytic results[19] for corrected Hartree theory in the low density limit (solid curve).

number in the low density regions, and the self–interaction corrected nature of those calculations.

Figure (1-2) shows momentum distributions for $r_s = 7$ and $r_s = 1$, which are representative of what we find at low and high densities, respectively. We note that the numerical results of both theories exhibit the correct k^{-8} scaling at large momenta and that we have very good agreement with the analytic results for the natural orbital functional[19] (applicable for $5.77 < r_s$).

We observe that at low densities the momentum distributions of both corrected Hartree theory and corrected Hartree–Fock theory resemble one another quite closely and, in fact, approach one another as r_s approaches infinity. This occurs because in this limit both theories predict very low occupations $f_i \ll 1$ so that the central quantities in the two theories approach one another, $f_i f_j + \sqrt{f_i(1-f_i)f_j(1-f_j)} \rightarrow \sqrt{f_i f_j}$.

The next panel of the figure illustrates that for densities outside the range of the applicability of the analytic solution ($r_s < 5.77$), this solution predicts $f > 1$ for sufficiently low momenta. We find that for such densities, the momentum distributions of corrected Hartree and corrected Hartree-Fock theory exhibit different behaviors at low momenta. Our numerical results indicate (to better than 1 part in 10^5) that the momentum distribution of corrected Hartree theory saturates to $f = 1$ near the Fermi level. (See inset.) Such behavior is in direct contradiction of analytic and numerical many–body results for the electron gas[16, 18] and is contrary to the previously reported experience with atoms for this functional[3]. In contrast, the momentum distribution of corrected Hartree–Fock theory never saturates the Fermi occupation constraint, but rather approaches a value at $k = 0$ which remains less than unity. We find that in the high density limit the corrected Hartree–Fock momentum distribution in fact approaches the standard Hartree–Fock result, ultimately leading to zero correlation energy for this theory in the high density limit.

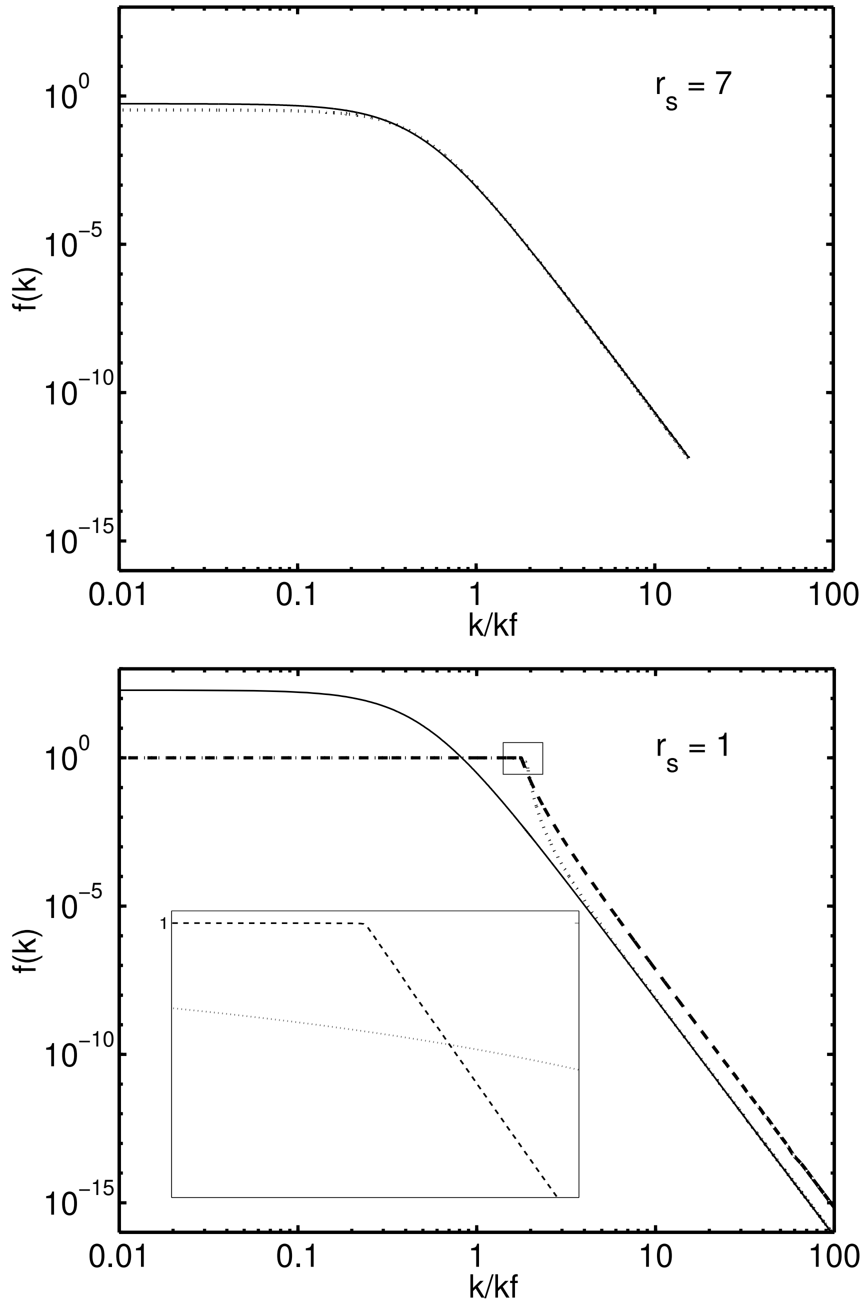

Figure 1-2: Predictions of momentum distribution of the uniform electron gas for $r_s = 7$ and $r_s = 1$: corrected Hartree theory (dashed curves), corrected Hartree–Fock theory (dotted curves), analytic results[19] for corrected Hartree theory in the low density limit (solid curve). Inset shows magnified view.

1.8 Conclusion

In conclusion, we find that when the two–body density matrix is expanded beyond traditional Hartree and Hartree–Fock theories in terms of tensor products of one–body operators, fundamental symmetries allow the exchange–correlation hole sum–rule to be inverted to give estimates for the two–body matrix, and thus correlations, in terms of the one–body matrix. By approximating the two–body density matrix in terms of the one–body density matrix, this approach also leads to new energy functionals of the one–body density matrix.

The addition of the first separable term beyond Hartree theory shows some promising features when applied to the homogeneous electron gas, such as giving the correct scaling behavior of the momentum distribution in the high momentum limit and the correct scaling behavior of the correlation energy in the high density limit. However, the electronic states predicted by this functional are significantly *over*–correlated, and attempts to improve upon it by the addition of further terms without additional constraints would only add more freedom in the minimization and lead to further over–correlation. Hence, before adding a second term, it is important to use antisymmetry to constrain the first such term to be the traditional Fock term. This gives the second functional which we considered, corrected Hartree–Fock theory. This second theory under–correlates the electron gas in the high density limit but still over–correlates at lower densities. This indicates that the addition of further terms will only hold the promise of producing a more accurate functional if additional, appropriate constraints are imposed.

	Hermiticity	Particle Permutation
$\hat{g} \overset{\otimes}{_{\mathrm{I}}} \hat{h}$	$\hat{g} = \hat{g}^\dagger$, $\hat{h} = \hat{h}^\dagger$	$\hat{g} = \hat{h}$
$\hat{g} \overset{\otimes}{_{\mathrm{II}}} \hat{h}$	$\hat{g} = \hat{h}^\dagger$	$\hat{g} = \pm\hat{g}^{\mathrm{T}}$, $\hat{h} = \pm\hat{h}^{\mathrm{T}}$
$\hat{g} \overset{\otimes}{_{\mathrm{III}}} \hat{h}$	$\hat{g} = \pm\hat{h}^\dagger$	$\hat{g} = \hat{h}$

Table 1.1: Implications of Hermiticity and particle permutation symmetries for tensor product expansions of γ.

Chapter 2

Improved tensor–product expansions for the two–particle density matrix

For the Goedecker–Umrigar functional [3], it has been shown analytically that the *per particle* correlation energy in the low density limit approaches a non-zero constant[19]. Furthermore, our numerical studies demonstrated that density matrix functionals, to date, give very poor results for correlation energies at intermediate and high density. Building on our tensor–product formalism introduced in the previous chapter, below we extend this proof, and demonstrate this shortcoming for a broad class of functionals. As regions of low densities contribute relatively little to the total energies of solids and molecules, we focus our attention on improving the results for the high and intermediate density regimes. We present a new density–matrix functional which is more than an order of magnitude more accurate in the high–density limit than those proposed previously and also represents a significant improvement in the intermediate regime. In atoms, the new functional performs similarly to previously explored density–matrix functionals and the generalized gradient approximation (GGA)[23] to density functional theory.

2.1 Notation

As in the previous chapter, we expand γ into Hartree and exchange–correlation contributions,

$$\hat{\gamma} = \hat{\gamma}_{\mathrm{H}} + \hat{\gamma}_{\mathrm{XC}}. \tag{2.1}$$

By separation of variables, it is always possible to expand the unknown exchange–correlation part as a (potentially infinite) sum of tensor products,

$$\hat{\gamma}_{\mathrm{XC}} = \sum_i \hat{u}_i \otimes \hat{v}_i. \tag{2.2}$$

Note that throughout this chapter we employ the twisted tensor product (denoted as "Type III" product in the previous chapter),

$$[u \otimes v](\mathbf{x}_1\mathbf{x}_1', \mathbf{x}_2\mathbf{x}_2') \equiv u(\mathbf{x}_1, \mathbf{x}_2')v(\mathbf{x}_1', \mathbf{x}_2),$$

to combine the one–body operators.

It has been point out, that given such an expansion truncated to finite order, the combination of the symmetry constraints on γ with the sum–rule is sometimes sufficient to determine the full four–variable function γ directly in terms of the two–variable density matrix n. Under these conditions, one can then evaluate the energy as an explicit functional of the density matrix using (1.1). Minimizing the energy over all physical density matrices then yields the ground–state of the system under this approximation.

2.2 New Functional

Considerable freedom remains in the terms for the expansion (2.2). Previously, we explored two choices, "Corrected Hartree–Fock" (CHF), and the Goedecker–Umrigar functional (GU)[3],

$$
\begin{aligned}
\hat{\gamma}_{\mathrm{XC}}^{\mathrm{CHF}} &= -\frac{1}{2}(\hat{n} \otimes \hat{n}) - \frac{1}{2}(\sqrt{\hat{n}(1-\hat{n})} \otimes \sqrt{\hat{n}(1-\hat{n})}), \\
\hat{\gamma}_{\mathrm{XC}}^{\mathrm{GU}} &= -\frac{1}{2}(\sqrt{\hat{n}} \otimes \sqrt{\hat{n}}),
\end{aligned}
$$

respectively. One way to view these functionals is in terms of the coefficient of the Fock term $\hat{n} \otimes \hat{n}$, which is $-1/2$ in the Corrected Hartree–Fock functional and is zero in the Goedecker–Umrigar functional. Under truncation of the expansion to one additional term, the value of this coefficient, through the sum–rule, then uniquely determines the form of the remaining tensor product.

2.3 Homogeneous electron gas

Figure (2-1) shows the correlation energies of the resulting functionals for the spin–unpolarized homogeneous electron gas. Corrected Hartree–Fock leads to no improvement in the correlation energy over Hartree–Fock, and approaches zero for high electron density. The Goedecker–Umrigar functional is a little better, and represents an improvement by a factor of two over Hartree–Fock at high densities, where it is overcorrelated. The functional is also overcorrelated at low densities, where it performs badly, with its per particle correlation energy approaching a *non–zero* constant. This property of the Goedecker–Umrigar functional has been demonstrated analytically in [19].

Having seen the behaviour of the above two functionals in the high density regime,

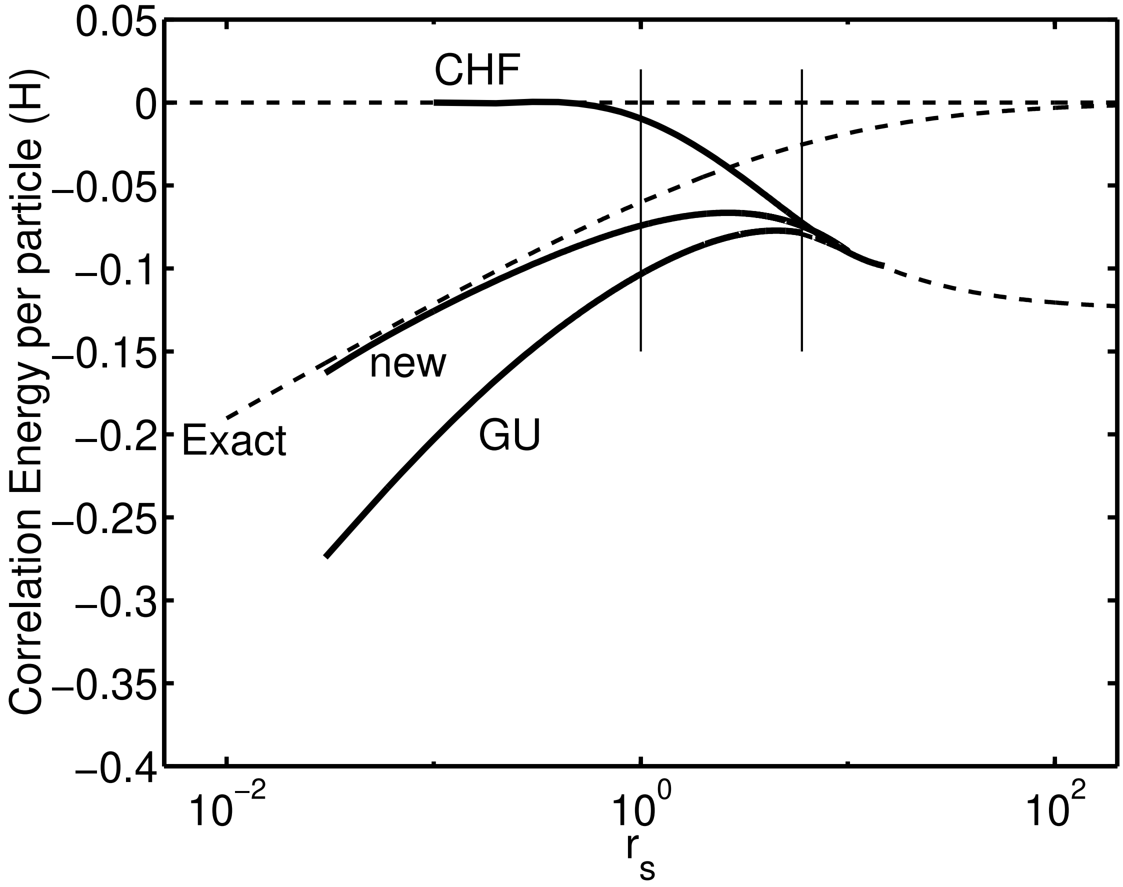

Figure 2-1: Correlation energies of the three functionals in the text and the exact result [17, 18, 20], versus the mean particle separation. The solid lines are numerical calculations, the dashed lines are exact results. The vertical lines indicate the range of typical metallic valence densities.

it is natural to look for a form with a Fock coefficient intermediate to those above. A coefficient of $-1/4$ results in the following functional,

$$\hat{\gamma}_{\text{XC}} = -\frac{1}{4}\left(\hat{n} \otimes \hat{n} + \sqrt{\hat{n}(2-\hat{n})} \otimes \sqrt{\hat{n}(2-\hat{n})}\right). \tag{2.3}$$

Figure (2-1) shows that this new functional results in a dramatic improvement at high densities (low r_s) and significant improvement at metallic valence densities ($1 < r_s < 6$). At low density, the behaviour of *all three* functional becomes the same, resulting in unrealistically large correlation energies *per particle*. Note that because this happens at low density, the contribution of these regions to the total energy is likely to be small. We now explore the reasons for this behaviour to ascertain the feasibility of making improvements in this regime as well.

2.4 Low density limit

To understand the approach of all of these functionals to one another in the low density limit, we recall that Cioslowski and Pernal[19] have determined the predictions of the Goedecker–Umrigar functional analytically for $r_s \gtrsim 5.769$. Their analysis applies, provided that $\hat{\gamma}_{\text{XC}}$ is represented as a single tensor product of a specific homogeneous form. As an illustration of the anomalous correlation behaviour which these functionals exhibit in the low density limit, Figure (2-2) compares the known many–body momentum distribution with numerical results for our new functional and the analytic distribution from [19] derived for the Goedecker–Umrigar functional for $r_s = 10$. Both density–matrix functionals have very low occupation numbers, and lose the characteristic signature of the sharp drop near the Fermi wave vector evident in the exact curve.

We now show that, as a consequence of the sum–rule, all of the above functionals

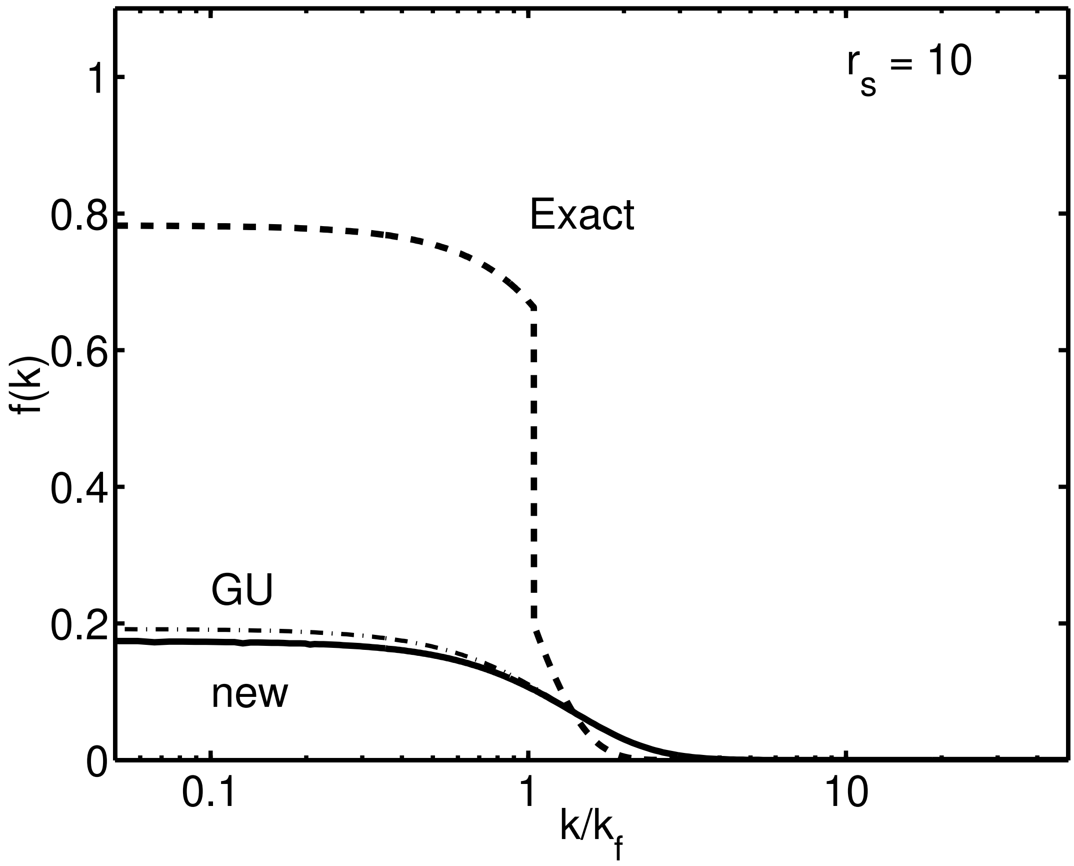

Figure 2-2: The momentum distribution $f(k)$ as a function of the wave vector k: exact [16] (dashed line), Goedecker–Umrigar functional [19] (dash–dotted line) and our new functional (solid line) at $r_s = 10$. The Fermi wave vector is denoted k_f

reduce to this particular form in the limit of high r_s. Recall from the previous chapter, that Fourier transforming (1.1), the total energy in the homogeneous electron gas becomes

$$E = 2\int \frac{Vd\mathbf{k}}{(2\pi)^3}\frac{\mathbf{k}^2}{2}f_{\mathbf{k}} - \frac{1}{V}\int \frac{V^2 d\mathbf{k}d\mathbf{k}'}{(2\pi)^6} g_{\mathbf{k}\mathbf{k}'}\frac{4\pi}{|\mathbf{k}'-\mathbf{k}|^2}, \tag{2.4}$$

where $g_{\mathbf{k}\mathbf{k}'}$ is the momentum space representation of $\hat{\gamma}_{\mathrm{XC}}$. In the case that $\hat{\gamma}_{\mathrm{XC}}$ is expanded as a sum of tensor products of single particle operators, its momentum space representation will be a sum of tensor products of the momentum space representation of those operators. Because our one–body operators are explicit functions of the density matrix, which for the homogeneous electron gas is diagonal in the momentum representation, their momentum space representations are also diagonal. Thus, their tensor products become simple products, and we have

$$g_{\mathbf{k}\mathbf{k}'} = \Omega_1[n(\mathbf{k}), n(\mathbf{k}')] + \Omega_2[n(\mathbf{k}), n(\mathbf{k}')] + \ldots, \tag{2.5}$$

where the Ω are the same functions of the density matrix which appear in the tensor products expressing $\hat{\gamma}_{\mathrm{XC}}$ but with the tensor–product operator replaced with simple multiplication. For example, for the Goedecker–Umrigar functional $g_{\mathbf{k}\mathbf{k}'} = -\frac{1}{2}\sqrt{n(\mathbf{k})n(\mathbf{k}')}$.

We now show that, in the low–density limit, the analytic solution presented in [19] for the case when $g_{\mathbf{k}\mathbf{k}'}$ consists of a *single* Ω function of the homogeneous form

$$\Omega[\lambda n(\mathbf{k}), \lambda n(\mathbf{k})] = \lambda^{\beta}\Omega[n(\mathbf{k}), n(\mathbf{k}')], \tag{2.6}$$

applies more generally to cases with multiple tensor products. The analysis of [19] begins with the scaling *ansatz*

$$n(\mathbf{k}) = \rho^{\Delta}\eta(\rho^{\nu}\mathbf{k}),$$

where ρ is the total electron density per spin channel, ν is a constant, η is a universal function that does not depend on the density explicitly, and $\Delta = 1 - 3\nu$ due to the constraint of the total particle density. Note that as $\rho \to 0$, $n(\mathbf{k}) \to 0$ uniformly (provided that $\Delta > 0$). This behaviour, evident in Figure (2-2), forms the crux of our argument; we verify it self–consistently below.

In the above limit, all of the Ω_i in the above functionals approach homogeneous functions with their own exponents β_i. Substituting $n(\mathbf{k})$ into (2.5), we then find that a single homogeneous function dominates $g_{\mathbf{kk'}}$, the one with the smallest β_i ($\equiv \beta_{\text{min}}$). Therefore, all the conditions for the analysis of [19] apply, in particular that

$$\nu \quad = \quad \frac{1-\beta_{\text{min}}}{3\beta_{\text{min}}-2}.$$

One can show that not only the above functionals, but any functional satisfying the sum–rule and dominated by a *single* homogeneous term in the low–density limit, will exhibit $\beta_{\text{min}} = 1$ in that limit. If indeed a single tensor product term dominates $\hat{\gamma}$, then the symmetry arguments presented in the previous chapter imply that $\hat{\gamma}_{\text{XC}}$ must take the form

$$\hat{\gamma}_{\text{XC}} = \hat{u}[\hat{n}] \otimes \hat{u}[\hat{n}]. \tag{2.7}$$

The sum–rule (1.2) then implies

$$\int d\mathbf{x}_2' \hat{\gamma}_{\text{XC}} = \hat{n}, \tag{2.8}$$

which, using (2.7), can be written as an operator equation

$$(u[\hat{n}])^2 = \hat{n}, \tag{2.9}$$

which implies that $u[\ ]$ must be a homogeneous function with exponent $1/2$ and therefore that $\hat{\gamma}_{\mathrm{XC}}$ is dominated by a term of homogeneous exponent $\beta_{\mathrm{min}} = 1$. Therefore $\nu = 0$ and $\Delta = 1$; the fact that $\Delta > 0$ confirms the self–consistency and the validity of the above analysis.

Ref. [19] analyzes the case $\beta = 1$ in detail, finding for the correlation energy per particle the result

$$\epsilon_{\mathrm{c}} \quad \rightarrow -\tfrac{1}{8},$$

as $\rho \rightarrow 0$. Figure (2-1) verifies that, in the low density limit, indeed the correlation energy of all three functionals approaches one another and, in particular, that they all approach a non–zero constant correlation energy of $-1/8$ Hartree per particle.

2.5 Atoms

The electrons in solids exhibit both free–electron like and localized atomic–orbital behaviour. Thus, another important limiting case for assessing new functionals is their behavior in atoms. Table (2.1) presents the results of our calculations[1] with these functionals for light atoms and their ions. For comparison, we included atomic results within density functional theory, both for the local density approximation (LDA), and the Perdew–Burke–Ernzerhof (PBE) version of the generalized gradient approximation[23], as well as results for the Goedecker–Umrigar functional, implemented *without* self–interaction correction.

The table reveals that, on the whole, the new functional performs for atoms quite comparably to both the Goedecker-Umrigar functional and the GGA. In particular,

[1]The atomic calculations were carried out by our collaborator, Stefan Goedecker

for the two–electron series, the new functional is somewhat worse, but for the four–electron series, it is somewhat better. Note that the Goedecker–Umrigar functional always overestimates the correlation energy, which is not the case for the new functional. Both density–matrix functionals are generally much better than LDA. For the neutral atoms, the density–matrix functionals are a little bit worse than GGA, but for the ions, both are better, or in some cases, similar to GGA. Thus, we feel that the density–matrix functionals provide an adequate description of atom–like systems.

2.6 Conclusions

We have introduced a new density–matrix functional, which is represent a major improvement over existing functionals of the density matrix. It is very accurate in the high density regime of the homogeneous electron gas, significantly improves the correlation energies at typical valence densities , and it is comparable to the previously proposed Goedecker–Umrigar functional and the generalized gradient approximation in atoms. We have shown also that in our tensor–product expansion of the two–body density matrix little further improvement can be expected at the present truncation of two terms. Our test systems span the range of environments encountered in solids and molecules, so this new functional is a good candidate to be used in calculations of real physical systems.

Atom	Present	GU	LDA	GGA
He	0.01	0.01	0.07	0.01
Be	-0.07	0.1	0.2	0.04
Ne	0.1	0.07	0.7	0.07
Be^{2+}	-0.02	0.005	0.2	0.04
C^{4+}	-0.02	0.003	0.4	0.06
C^{2+}	0.1	0.2	0.4	0.07
O^{4+}	0.1	0.2	0.6	0.1

Table 2.1: Error in the total energies for the new functional in atoms, in Hartrees. For comparison, results are included for the Goedecker–Umrigar functional and various approximations within density functional theory (see text.)

Chapter 3

Paramagnetic structure for the soliton of the 30° partial dislocation in silicon

Dislocations are central to the understanding of the mechanical response of materials. The mechanical behaviour of any crystalline material is determined by a hierarchy of crystalline defects of successively lower dimension. Grain boundaries are two dimensional defects that control the evolution of the microstructure of the material. The creation and motion of dislocations, which are one dimensional extended topological defects of the lattice, mediate the plastic response of a crystal to external stress. In silicon, which has a bipartite lattice, the primary mobile dislocations are the screw and the 60° dislocation, which belong to the glide set. They dissociate into partial dislocations bounding stacking faults. The 30° degree partial is less mobile than the 90° degree partial, so the simultaneous motion of both types of dislocations in the glide set is controlled by the 30° degree partial. The mobility of dislocations in high Peierls barrier materials such as silicon is effected by the motion of kinks, zero dimensional defects. The low energy kinks along the 30° partial have been shown to involve a composite structure where kinks bind with soliton excitations in the reconstructed ground state of the dislocation core. The solitons arise because the reconstruction involves pairing of the core atoms, which leads to two degenerate ground state configurations. The domains of the two configurations are separated by the soliton. These soliton excitations are also known as "anti-phase defects" (APDs)[24] or "phase switching defects". For a general review on dislocations in semiconductors, see [25, 26], or more recently [27, 28].

We have carried out an *ab initio* study exploring the lattice and electronic structures, excitation energy, and the density of these APDs which are the simplest, lowest energy, fundamental excitations of the dislocations in the hierarchy ultimately leading to the macroscopic mechanical behaviour of the solid. Using a special multiscale sampling technique, we are able to present, for the first time, an *ab initio* calculation for the free energy of formation for this defect. The soliton is associated with an atom in the dislocation core which is not part of a reconstructed dimer. In the simple, conventional picture, this atom (henceforth to be referred to as the "soliton atom") only has three bonds and therefore an unpaired electron. This simple model, however, does not lead to predictions consistent with of any of the observed ESR signals associated with plastically deformed silicon.

3.1 New Structure

We propose a new theory for the structure of the soliton. We find that the ground state of the soliton has an unexpected structure with electronic states which *are* consistent with the most stable ESR center in plastically deformed silicon, the only one which remains after careful annealing. The reason why the natural connection between this center and the lowest energy excitation of the dislocation core has not been made previously is that the observed ESR center has a highly unusual symmetry. In support of our theory for the structure of the soliton, we gather here several pieces of evidence from both reports of ESR results and our own *ab initio* calculations. The final combined *ab initio*–experimental identification which we make allows for the possibility that future more precise measurements of the ESR signal strength could be used to determine experimentally the soliton density in plastically deformed silicon, a parameter which is important for understanding the plastic response of the material.

Figure (3-1) reviews the basic geometry of the 30° partial dislocation studied in this work. The dislocation is the one–dimensional boundary defining the edge of a half–planar (111) stacking fault. Atoms in the central core of the dislocation (shaded gray in the figure) are connected to the bulk with only three bonds per atom. The dislocation undergoes a reconstruction whereby the core atoms pair up in dimers forming intra–core bonds and thus become four–fold coordinated. This reconstruction breaks the original translational symmetry along the core and doubles the primitive repeat distance along the core axis. Associated with this broken translational symmetry is a low energy soliton defect where, by creating a single unpaired core atom, the system may change along the line from one of the two symmetry related degenerate ground state phases to the other. Because the soliton atom is expected to have a dangling bond, it is natural to look for an ESR signal for this defect. The relatively low energy we expect for such an excitation leads us to expect a relatively large equilibrium population at silicon annealing temperatures ($\sim$ 900K) and therefore that the ESR signal would not anneal out as quickly as other signals associated with the formation and motion of dislocations. We further would expect this signal to be detected in all systems which contain 30° partial dislocations.

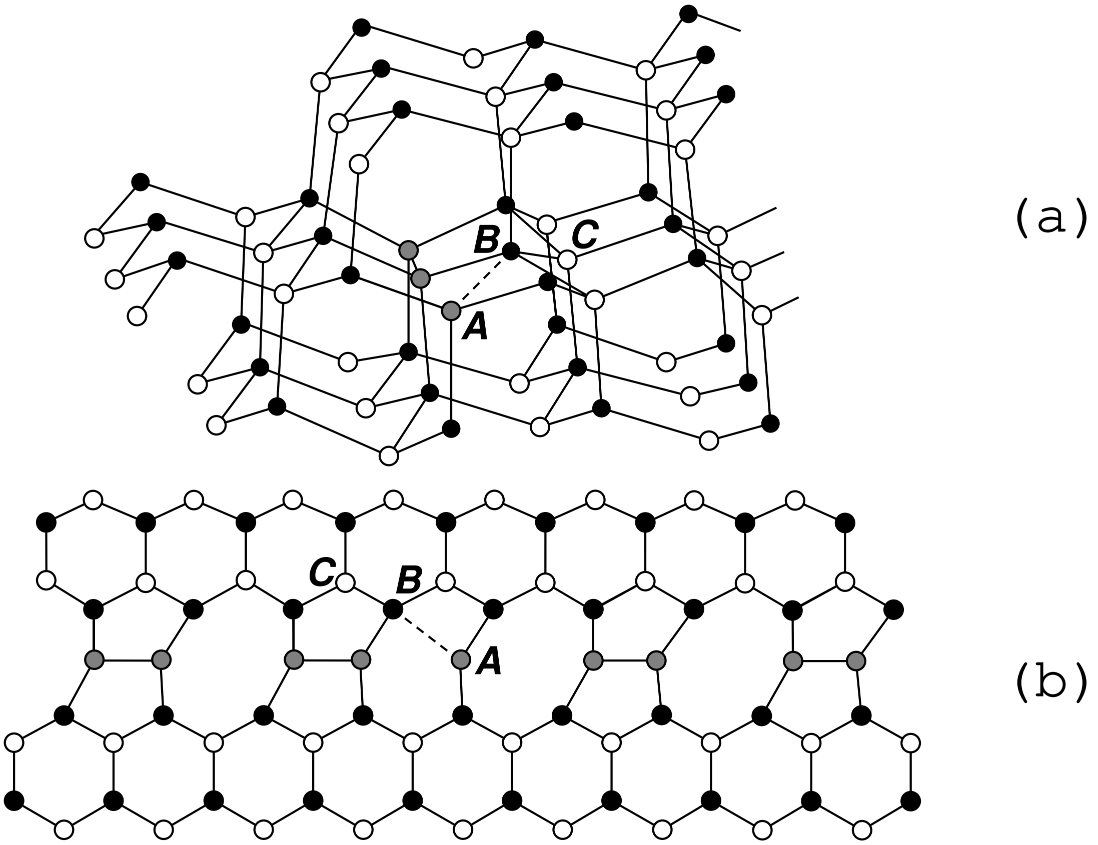

Figure 3-1: (a) Three dimensional view of a 30° partial dislocation, with a soliton domain wall, (b) schematic drawing of the same in the (111) plane. The grey circles represent the atoms in the central core of the dislocation.

3.2 Electron Spin Resonance

Indeed, it was discovered over thirty years ago that plastically deformed silicon gives a wide variety of ESR signals [29, 30, 31, 32, 33, 34]. Out of the dozens of ESR centers, four have been identified as associated with the 30° dislocation core. The literature refers to these centers as Si–K1, Si–K2, Si–Y and Si–R. The K2 and K1 defects have been identified to be electronic excitations of the same structural defect. Kisielowski–Kemmerich [27] made the currently accepted identification of the K and Y defects. It is well known that the first three of the aforementioned ESR signals anneal out over the (temperature dependent) time–scale of about an hour [25]. Only one signal remains, the one labeled R [33, 35, 36]. This center is "thermally stable" (it does not anneal out) and is observed even at high deformation temperatures ($> 900°$ K) where the other signals anneal out too quickly to be observed. The R center was shown to be very similar to Y in its near isotropy and large width and also similar to the ESR signals obtained from amorphous silicon [35]. Kisielowski–Kemmerich *et al.* [35] mentioned the possibility that R is the residue of Y after annealing. However, there is no direct evidence that the Y center and the R center are due to the same structural defect. The intensities of both centers are proportional to the dislocation density, but the Y intensity is also proportional to the area swept out by the glide of the dislocations during the deformation.

The nature of the ESR centers and the soliton excitation energy of the 30° partial dislocation have yet to be addressed with modern *ab initio* techniques. The APD was studied previously by Heggie and Jones [37, 38, 39]. Nunes *et al.* [40] reported results of a tight binding study including the energetics of kinks and solitons for the 90° partial dislocation. More recently, Bennetto *et al.* [41] proposed a new kind of period–doubling reconstruction for the 90° partial using calculations based on density functional theory. Spence *et al.* [42] investigated kink motion. Bulatov *et al.* [24]

have carried out a comprehensive study of the 30° partial dislocation and its defects but using only the Stillinger–Weber (SW) inter–atomic potential[43], which provides no electronic structure information. In order to investigate the electronic structure of low energy excitations of the 30° partial dislocation core, we embarked upon a density functional study of the system.

3.3 Computation

To prepare approximately relaxed initial ionic configurations with the correct bonding topology, we first relaxed lattices containing dislocation cores using the SW potential. While doing this, we discovered that the soliton atom moves out of line with respect to the dislocation core. To probe this interesting feature further, we carried out calculations within the plane wave total energy density functional approach[44]. To describe the electron–electron interactions we used the Perdew–Zunger[45] parametrization of the Ceperly–Alder[46] exchange–correlation energy of the uniform electron gas. To describe the electron–ion interactions we used a non–local pseudopotential of the Kleinmann–Bylander form[47]. The electronic wave functions were expanded in a plane wave basis up to a cutoff of 8 Ry.

All supercells used in this study have the same size in the plane perpendicular to the (110) dislocation axis. Two partial dislocations of equal but opposite Burgers vectors at a separation of 14 Å cut through this plane. Following Bigger *et al.* [48], the lattice vectors are arranged so that the periodic dislocation array has a quadrupolar arrangement. Each cell contains forty-eight atoms per core atom in the dislocation core. To calculate the excitation energy of the soliton, it is also necessary to calculate the energy of the perfectly reconstructed dislocation. However, the smallest supercell which is commensurate with both structures contains six bilayers (288 atoms). It is possible, however, to reduce the computational time by using two different supercells.

For the reconstructed case, the supercell contains two bilayers along the dislocation line, while the soliton structure contains three, (in total, 96 and 144 atoms, respectively). The lattice vectors were obtained by relaxing a completely reconstructed dislocation within the SW model in the 96 atom cell. The three bilayer cell was then obtained from this by multiplying the lattice vector that points along the dislocation axis by $\frac{3}{2}$. To minimize numerical errors, including those from k-point sampling, basis set truncation and supercell effects, we compute differences of energy differences. For each supercell, we generated a completely unreconstructed configuration where all the core atoms have only three bonds. These structures can be realized in both supercells, and thus serve as the reference point. The final excitation energy is the difference between the deviations in the energy from the unreconstructed structure in each cell. By keeping the lattice vectors fixed throughout the calculations, we simulate the strain field which widely separated solitons would experience along a reconstructed dislocation line.

To ensure maximum transferabilty of results between the two cells, the calculations employed k–point sets which give identical sampling of the Brillouin zone for the two supercells: $\{(0, 0, \pm 1/4)\}$ for the 144 atom cell and $\{(0, 0, \pm 1/6), (0, 0, 1/2)\}$ for the 96 atom cell. To find relaxed structures, we moved the ions along the Hellmann–Feynman forces until the ionic forces were less than 0.02 eV/Å. Typically, this was accomplished in 40 ionic steps, where between ionic steps we made 10–15 electronic relaxation steps using the analytically continued functional approach[49].

3.4 Results: structure

Figure (3-2) shows the projection in the (110) plane of our *ab initio* results for the structure of the soliton. The structure on the left is our prediction for the ground state. In this configuration, there is a five–fold coordinated bulk atom (B) in the

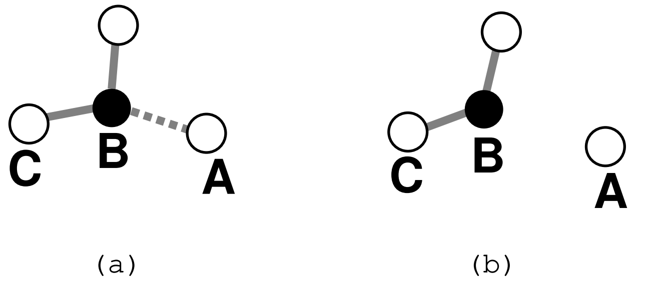

Figure 3-2: Coplanar atoms near the soliton: (a) in the proposed ground state for the soliton, (b) conventional structure with the soliton atom in line with the remaining atoms in the dislocation core. The notation for the atoms is the same as in Fig. (3-1): atom A is the soliton atom, B is the five–fold coordinated bulk atom (see text), C is a bulk atom bonded to B opposite from A.

immediate neighbouring row to the dislocation core. Its new, fifth neighbour is the soliton atom (A). Our *ab initio* results show that the conventional structure on the right (generated by keeping the soliton atom collinear with the dislocation core) is not only higher in energy but also spontaneously decays into the ground state on the left.

The *ab initio* excitation energy of the soliton is $0.65 \pm (\approx\ 0.2)$ eV, where we attribute most of the uncertainty to the uncontrolled local density approximation and supercell effects. If we are to estimate the equilibrium density of the APDs at 900 K, the entropy of the system must be taken into account. Specifically, we need the change in entropy associated with the creation of an APD.

3.5 Free energy

To determine this, we used a new technique to carry out *ab initio* free energy calculations[1] for the fully reconstructed dislocation, and the completely unreconstructed

[1]The calculations were carried out by Torkel D. Engeness

dislocation, whose core can be viewed as a row of APDs. The entropy change between these two states gives an estimate of the entropy change of creating a pair of APDs on the dislocation core. Traditional techniques for ab initio evaluation of the entropy is at present unfeasible, hence we applied, for the first time in an *ab initio* calculation, the multiscale sampling approach[50]. The change in free energy is given by the adiabatic work of a single degree of freedom in going from the reconstructed state ($\lambda = 0$) to the unreconstructed state ($\lambda = 1$).

$$\Delta F = \int_0^1 d\lambda \langle \partial E / \partial \lambda \rangle$$

The essence of multiscale sampling is that the phase space of the integral is explored in a crude, simple model (in this case, Stillinger–Weber) using a very large number of samples, resulting in a much smaller number of statistically independent points. These are evaluated within an accurate model (*ab initio*) and a corrective Boltzmann factor is used to establish the true Boltzmann distribution of the accurate model. This method is related to correlated sampling, or biased sampling. The key here is that the samples are obtained from an atomistic, higher level description of the same physical system, but the exact *ab initio* thermal ensemble average is obtained at a dramatically accelerated rate (See [50] for details.) The sampling was carried out with the standard Monte Carlo method using the Metropolis algorithm.

Figure (3-3) shows the free energy as a function of λ for the SW model, a tight binding model and the *ab initio* model, the latter two obtained using multiscale sampling from the SW model. Table (3.1) shows the entropy calculated from the free energy. Although the free energy values are quite different for the three models (due to the subtleties of the quantum mechanics of the bond being broken), the entropy variation (mostly changes in the vibrational modes of the surrounding lattice) is much smaller. Taking $T\Delta S$ to be about 0.1–0.2 eV, the free energy difference associated

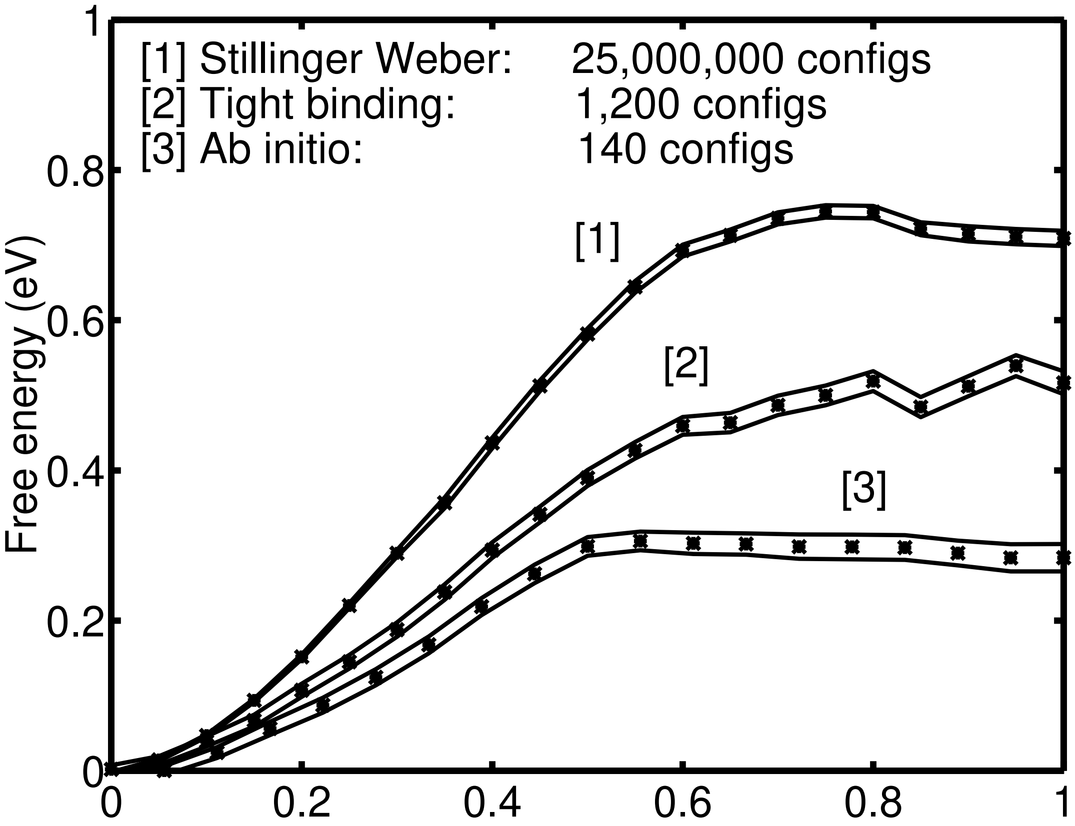

Figure 3-3: Free energy of reconstruction in three different models using the multiscale sampling approach (see text). The free energy is evaluated by integrating the work done in moving the system along the coordinate λ from the reconstructed state to the completely unreconstructed state. The lines above and below the data points represent statistical error bars resulting from the sampling.

with creating an APD can be estimated to be about 0.3–0.6 eV. This corresponds to a density

$$\rho = \tfrac{1}{2} e^{-\Delta F/kT}$$

in the range of 4×10^{-4} to 10^{-2} solitons per core atom at 900 K, which is consistent with the observed ESR signal strengths; the R center is estimated to have a density of about 0.001 per core atom [36, 51]. (The factor of one half in ρ comes from the fact that in a particular given phase of the ground state, only half of the core atoms represent possible soliton sites.) Given the exponential sensitivity of the density, we find this agreement encouraging, with the caveat that the exact temperature at which the density of solitons is frozen during the quenching after the annealing is at present unclear.

3.6 Results: symmetry

A great advantage of the *ab initio* calculations, beyond their accuracy, is that they also yield the electronic states, in particular giving information about their spatial symmetry. In Figure (3-4) we plot the angular momentum decomposition of the local density of states as obtained from the Kleinmann–Bylander projections of the electronic eigenstates. Part (a) shows, for an atom far from the core, the familiar concentration of s–like states at the bottom of the valence band and p–like states at the top. To explore the nature of the soliton state, we compare this to the local densities of states for the soliton atom in the proposed (3-4b) and conventional (3-4c) configurations. We also plot the local density of states for the quasi–fivefold coordinated atom (3-4d). The appearance of the peak near the top of the valence band in the s channel of the soliton atom in its ground state (3-4b) shows that the state associated with the soliton is much less anisotropic than the simple dangling

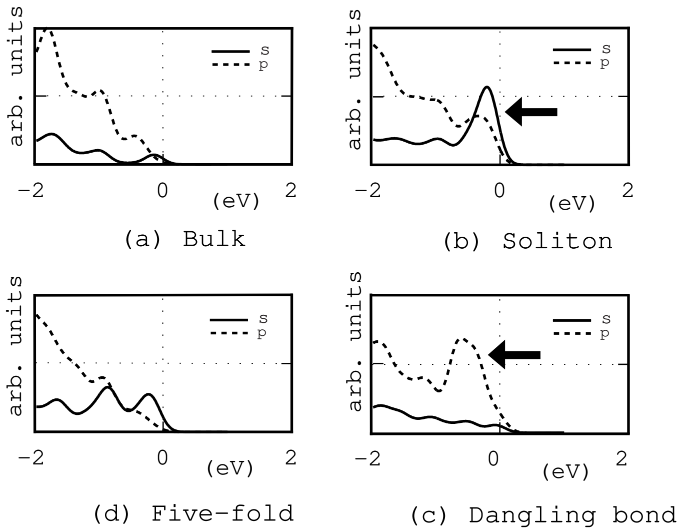

Figure 3-4: Local density of states, calculated by acting on the filled bands with the Kleinmann–Bylander projectors centered at (a) an atom deep in the bulk , (b) the soliton atom in its proposed ground state, (c) a conventional soliton atom with a dangling bond, (d) the quasi five–fold coordinated atom. The horizontal axis is the energy (eV), the scale of the vertical axis is arbitrary but the same for all four panels. Solid and dashed lines represent densities in the s and p channels respectively.

p–like bond on the soliton atom in the conventional picture (3-4c). Note also that the density in the p channel of the soliton atom is also correspondingly diminished relative to that in the conventional state. We further note an enhancement at the same energy in the s channel of the quasi–fivefold coordinated atom (3-4d), which indicates that the unpaired electron is shared between this atom and the soliton atom. Defects in the dislocation core therefore need not be associated with strongly directional electronic states, as has been previously assumed in identifications of ESR centers. Similar mechanisms can plausibly play a role in the decrease in anisotropy of other point defects.

The connection between the symmetry of the electronic state and the corresponding ESR signal is via the effective g tensor,

$$g_{ij} \approx g_0 \delta_{ij} - 2\Lambda \sum_n \frac{\langle \phi_0 | \mathbf{L}_i | \phi_e \rangle \langle \phi_e | \mathbf{L}_j | \phi_0 \rangle}{E_e - E_0},$$

where ϕ_0 is the unpaired state and ϕ_e are the excited states and Λ is the atomic spin–orbit coupling constant. In general ϕ_0 can be broken into angular momentum components (as in Figure (3-4)), of which the s wave component makes no contribution to the off–diagonal matrix elements, so we expect the anisotropy of g to be proportional to the population of the p channel. (Higher angular momentum components are negligible for filled states in silicon.) This population, the area under the peak associated with the unpaired electron in the p channel, drops by about a factor of two as the soliton moves from its symmetrical dangling bond configuration (3-4c) to our proposed state (3-4b). The literature contains qualitative observations of the decreased anisotropy of the R signal and one quantitative comparison which comes from measurements of the Y signal, of which the R is presumed to be the residual after annealing. In [27], this anisotropy is compared directly with that of the K1,2 centers, which have typical dangling bonds, and is shown to be less by about a factor of two, in agreement with our electronic structure results.

To explore the nature of the bonding near the soliton atom, we plot the total valence charge density in Figure (3-5), which shows that the soliton atom (A) makes a weak bond with the neighbouring five–fold coordinated atom in the bulk (B). The new bond (A–B) of the bulk atom is very similar to a now weakened but previously existing bond in the spatially opposite direction (B–C). Such five–fold coordinated structures have been considered previously in silicon by Pantelides [52] and more recently by Duesbery et al.[53]. Through *ab initio* studies it was demonstrated that such five–fold defects in amorphous silicon should show similar anomalies in the angular momentum decomposition of the local density of states[54]. In amorphous silicon, however, the

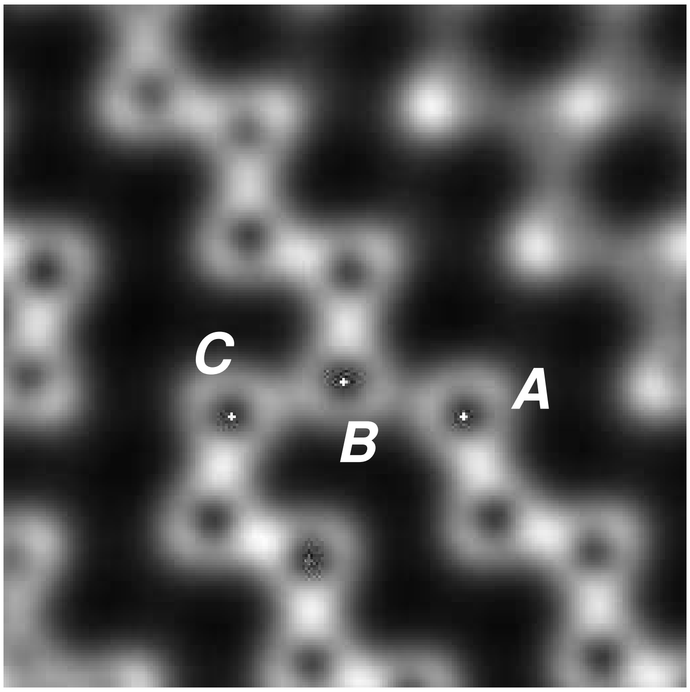

Figure 3-5: Two dimensional slice of the total charge density from our *ab initio* calculation, through the plane containing the atoms A, B, C, as labelled in the previous figures. The new A–B bond is nearly as pronounced as the bulk B–C bond, which is clearly weakened compared to the other (vertical) bulk bond of atom B seen in the figure.

shift from the p channel to the s channel is more pronounced with the peak in the p channel disappearing completely.

3.7 Conclusion

We have presented *ab initio* results indicating that the ground state of the soliton has an unusual structure involving a five–fold coordinated atom and a correspondingly unusual electronic structure. The excitation energy we calculate with the first *ab initio* application of the multiscale approach for this new defect corresponds to a thermal equilibrium density which is compatible with the observed signal strength of the R center, which is the only thermally stable paramagnetic center associated with the 30° partial dislocation. In line with our notion of the soliton being the fundamental excitation of the reconstructed dislocation core, the R center is observed independent of the method of deformation and in proportion to the dislocation density. Our calculations show that the soliton has an enhanced isotropy compared to that of a simple dangling bond, which correlates well with the puzzling, nearly isotropic signal of the R center. In the ground state structure which we propose, the soliton atom makes a weak bond with a neighbouring bulk atom and thus gives rise to an amorphous–like bonding arrangement. This could explain in part the similarity of the ESR signature of the R center to that of amorphous silicon. Based on the above arguments and results, we propose that the domain walls in the reconstruction of the 30° partial dislocation and the R centers observed in ESR experiments are one and the same. Any viable competing theory which does not identify the R signal with the soliton must both predict a more plausible microscopic structure for the R center and explain why the unpaired electron of the low energy soliton does not exhibit an ESR signal.

It must be reemphasized that the above results and indentification relate to ther-

Model	ΔE (eV)	ΔF (eV)	$T\Delta S$ (eV)
SW	0.81	0.71 ± 0.01	0.10 ± 0.01
Tight binding	0.71	0.53 ± 0.02	0.18 ± 0.02
Ab initio	0.42	0.27 ± 0.03	0.15 ± 0.03

Table 3.1: Free energy and entropy of reconstruction. The statistical uncertainty from the multiscale sampling method is shown. The reconstruction energy is displayed in the first column for information.

mally equilibrated paramagnetic centers. It is not clear at present, what is the precise connection between the solitons and the Y center. Furthermore, in light of this new proposal for the native defect of the system, it might be time to reexamine the problem of the C line in the DLTS experiments on plastically deformed silicon. This, so far unidentified defect is also thermally stable, so it would be natural to investigate its relationship to the soliton of the 30° degree partial.

Chapter 4

Hyperfine parameters for azafullerene $C_{59}N$

In the last chapter, the connection between the *ab initio* calculation and the experimental observables was rather indirect. It would be immensely useful if one could compute the observables directly. The principal quantities of interest in an Electron Spin Resonance (ESR) experiment are the g tensor and the hyperfine coupling tensor. It is the purpose of this and the next chapter to show that both of these can be obtained from a traditional plane wave pseudopotential method. In this chapter, we are going to focus on hyperfine couplings.

The chapter is organized as follows. First, we introduce the azafullerene molecule, and recall how plane wave pseudopotential density functional theory can be used to calculate the hyperfine couplings for various nuclei. The hyperfine couplings are predicted to be very susceptible to numerical errors, so we undertake a study of this difficulty by presenting the different ways to compute the relevant observables in a much smaller molecule, hydrogenated pyridine, where the unpaired electron has a qualitatively similar orbital to that in $C_{59}N$. Finally we compute the hyperfine couplings for $C_{59}N$, where our results are in much better agreement with the experiment than previously published calculations.

4.1 $C_{59}N$

Azafullerene, the nitrogen substituted version of the famous Buckyball is one of the most interesting modifications of the basic C_{60} molecule. Since the discovery of the new form of carbon in 1985, there has been an enormous interest in fullerene and nanotube structures, with potential applications ranging from new types of displays to extremely strong threads and even drugs[55]. The singly substituted versions of C_{60} could be used as dopants in making a semiconducting or even metallic fullerene crystal. Attempts are also being made to put $C_{59}N$ inside carbon nanotubes, where it would alter the electrical properties of the tube, for example, making it metallic. For a general review of heterofullerenes, see [58]. The structural and electronic properties of $C_{59}N$ monomer have been studied before[57] and it was found that there is significant delocalization of the unpaired electron over the molecular cage. However, until recently, $C_{59}N$ was only available in dimerized form. It is now possible to make solid solutions of the monomer in conventional C_{60} in macroscopic quantities[56], and thus a direct comparison with the experiment can be made. The degree of delocalization over the $C_{59}N$ cage has a significant bearing on the electrical properties of the alloy.

A calculation of the electronic structure of $C_{59}N$ was published along with the experimental details in [56]. The relative values of the hyperfine coupling were showing the same trends as the experimental data, but the absolute values were off by more than a factor of two. The calculated hyperfine couplings enable to identification of the measured values with particular sites on the cage, and thus give the degree of delocalization of the unpaired electron.

The ESR spectrum of the system has a rich hyperfine structure with three main lines corresponding to the splitting caused by ^{14}N and several satellites due to ^{13}C. To identify these lines with particular atoms, it is necessary to carry out a calculation of the hyperfine splittings expected from the individual ions on the molecule. We used the traditional density functional plane wave pseudopotential method and the local spin density approximation to relax the molecular geometry[1], then applied the Projector Augmented Wave (PAW) method[59] described below, to calculate all–electron spin densities at the nuclear sites. The calculation was carried out in vacuum in a 24 Å cubic cell, the plane wave cutoff was 20 Hartrees.

4.2 Fermi contact interaction

In the electron spin resonance experiment an external magnetic field is applied which breaks the Kramers degeneracy, thus separating the energy levels occupied by mostly upspin and mostly downspin electrons [2]. The system absorbs microwaves if the energy of the photons equals the energy separation of the Kramers conjugate states, thus creating the conditions for resonance. The properties of this basic resonance line can be altered by chemical the environment, which makes ESR an invaluable tool for learning about the microstructure of materials. If there is a paramagnetic nucleus, such as ^{13}C or ^{14}N, in the vicinity of the resonating electron, the microscopic magnetic field arising from this nucleus acts in addition to the external field. The upshot of this is that the resonance line is split, for spin 1/2 nuclei such as ^{13}C, into two lines, or into three lines in the case of ^{14}N which has spin 1. The split spectrum is said to have a "hyperfine structure" and the amount of splitting is the "hyperfine coupling". The interaction of the electron and nuclear spins is well described by the Dirac equation, and can be reduced to an isotropic term, and dipole term. In many cases, the isotropic

[1] Other details of the method are as in the previous chapter

[2] After taking account of the spin–orbit interaction, s_z is no longer a good quantum number

term, also called the Fermi contact interaction, dominates, and can be expressed by the Hamiltonian

$$H_{\mathrm{hf}} = A_{\mathrm{hf}} \mathbf{S}_I \cdot \mathbf{S}_e \delta(\mathbf{r} - \mathbf{r}_I),$$

where the subscript I refers to the nucleus, and A_{hf} is the isotropic hyperfine coupling constant and $\mathbf{S}$ is the spin operator. Because of the Dirac delta function, the hyperfine splitting for a given nucleus only depends then on the spin density at the position of the nucleus.

The isotropic hyperfine coupling is thus very sensitive to the values of the wavefunctions near the atomic cores. This presents a difficulty when using pseudopotentials because the pseudo wavefunctions have completely different structure near the cores. In fact, this different structure is what makes pseudopotentials useful. The real wavefunctions are highly oscillatory near the cores, but this oscillation arises purely by the necessity to be orthogonal to the lower lying core states, not the chemical environment. Pseudopotential theory works with smooth wavefunctions that approximate very well the real wavefunctions in the inter–atomic regions. Thus when we are interested in observables that depend on core region properties, we need a way to *reconstruct* the oscillatory behaviour.

Before looking at how this might be done rigorously, it is important to point out that, intuitively, this reconstruction should be possible. After all, if the oscillatory behaviour is purely due to the core states, which are assumed not to change too much in various chemical environments, the very same oscillations are going to be present in the fully reconstructed valence wavefunctions.

In the next section, we briefly recall the proper mathematical formalism, introduced by Blöchl in the mid–nineties, that achieves this reconstruction rigorously. The electronic structure community has embraced the method as a self–contained one, whereby the total energy is calculated as an all–electron observable from the reconstructed wavefunctions.

In this thesis, I wish to emphasize and explore the utility of the pseudo operators and the reconstruction as a *post processing* step in conjunction with existing plane wave pseudopotential methods. Many such codes exist, and a lot of new research can be done by extracting core–region information from them. To do this, we take the pseudo valence wavefunctions from the existing code and reconstruct them according to the prescription of Blöchl to evaluate operators in the core regions. This is the approach we take below and in the next chapter. Here, we investigate the various different ways in which the post processing can be carried out in the case of hyperfine coupling.

4.3 Projector Augmented Waves

The reconstruction idea was given the proper mathematical treatment by Blöchl. We introduce angular momentum projector functions $|p\rangle$ for every atomic species, that have the property

$$\langle p_i|\phi_j^{\mathrm{ps}}\rangle = \delta_{ij},$$

where $\phi_j^{\mathrm{ps}}(\mathbf{r})$ are atomic pseudo wavefunctions. The projectors $|p\rangle$ are localized near the nucleus and they vanish outside the cutoff radius of the pseudopotential, the usual indices n, l, m have been condensed for clarity. The reconstruction formula is

$$|\Psi^{\mathrm{rec}}\rangle = |\Psi^{\mathrm{ps}}\rangle + \sum_R \sum_i \left(|\phi_{R,i}^{\mathrm{ae}}\rangle - |\phi_{R,i}^{\mathrm{ps}}\rangle\right) \langle p_{R,i}|\Psi^{\mathrm{ps}}\rangle, \tag{4.1}$$

where $|\Psi^{\mathrm{ps}}\rangle$ is an extended pseudo wavefunction, ϕ^{rmae} are the all–electron atomic orbitals and R runs through the position of the nuclei. Note that because the atomic pseudo wavefunctions are identical with the all–electron atomic wavefunctions outside the cutoff radius, the reconstructed wavefunction $|\Psi^{\mathrm{rec}}\rangle$ is identical to the pseudo wavefunction $|\Psi^{\rm ps}\rangle$ in the inter–atomic region. The reconstruction is exact within the frozen core approximation if we take a complete set of angular momentum projectors. In practice, it is often enough to take just two, or even just one projector for every pair l and m.

In the few years since the first introduction of the above reconstruction, Blöchl extended the idea into a completely self–contained method for generating pseudopotentials and doing *ab initio* calculations using the reconstructed all–electron wavefunctions. This new, so–called Projector Augmented Wave (PAW) method, has begun to be tested on real systems by some other groups as well, but it already seems to encompass and supersede the various different pseudopotential approaches that have been used up until now. If the calculation is done self–consistently, the only uncontrolled approximations are those of density functional theory itself through the chosen exchange–correlation functional and the frozen core approximation. The frozen core approximation can be controlled by choosing which states to include in the core. The truncation of the above angular momentum expansion can be extended by using multiple projectors until the desired convergence is achieved.

For our purposes here, we also need the notion of a *pseudo operator*, which arises naturally within the PAW formalism. Let us compute the matrix elements of an operator $\mathcal{O}$ between reconstructed wavefunctions. Using the definition (4.1),

$$\begin{aligned}\langle\Psi_1^{\rm rec}|\mathcal{O}|\Psi_2^{\rm rec}\rangle &= \langle\Psi_1^{\rm ps}|\mathcal{O}|\Psi_2^{\rm ps}\rangle + \\ &+ \sum_{RR',ij}\langle\Psi_1^{\rm ps}|p_{R',j}\rangle\left(\langle\phi^{\rm ae}_{R',j}|\mathcal{O}|\phi^{\rm ae}_{R,i}\rangle - \langle\phi^{\rm ps}_{R',j}|\mathcal{O}|\phi^{\rm ps}_{R,i}\rangle\right)\langle p_{R,i}|\Psi_2^{\rm ps}\rangle.\end{aligned} \tag{4.2}$$

Note that for local and semi–local operators, only on–site terms contribute, where

$R = R'$. Thus we can define pseudo operators $\mathcal{O}^{\text{ps}}$,

$$\mathcal{O}^{\text{ps}} = 1 + \sum_{R,ij} |p_{R,j}\rangle \left(\langle \phi^{\text{ae}}_{R,j} | \mathcal{O} | \phi^{\text{ae}}_{R,i} \rangle - \langle \phi^{\text{ps}}_{R,j} | \mathcal{O} | \phi^{\text{ps}}_{R,i} \rangle \right) \langle p_{R,i}|.$$

This pseudo operator, when acting on pseudo wavefunctions, will give the same matrix elements as the corresponding all–electron operator acting on all–electron wavefunctions.

To extract the hyperfine splitting from an *ab initio* calculation, we need the all–electron spin density at the nuclear positions,

$$\tilde{n}(R) = n_{\uparrow}(R) - n_{\downarrow}(R).$$

The simplicity of the delta function operator made this one of the first applications of the above reconstruction ideas. If the angular momentum expansion were complete, the first and third terms of equation (4.2) would cancel exactly, and only the term involving the all–electron atomic states needs to be calculated. However, only the s state has a non–vanishing density at the nucleus of an atom, so (4.2) reduces to

$$n^{\text{rec}}(R) = n^{\text{ps}}(R) \frac{n^{\text{ae}}_{\text{at}}(0)}{n^{\text{ps}}_{\text{at}}(0)}, \tag{4.3}$$

where the subscript "at" refers to atomic densities. Equation (4.3) amounts to a rescaling of the pseudo spin density. Blöchl calculated hyperfine constants for a variety semiconductor systems using equation (4.3) and obtained satisfactory results.

The problem with the $C_{59}N$ molecule is that the unpaired electron is in an orbital which has almost p-like symmetry near the nuclei. The spin density at the nucleus is therefore very small and its value becomes very sensitive to how well the charge density is resolved near the nucleus. We find that a more robust approach is to retain the projector form from equation (4.2) and use the formula

$$n^{\mathrm{rec}}(R) = n^{\mathrm{ae}}_{\mathrm{at}}(R)|\langle\Psi^{\mathrm{ps}}|p_{R,s}\rangle|^2. \tag{4.4}$$

Apart from the issue of resolution near the ionic core, the truncation of the angular momentum expansion in (4.2) implies that the operator form (4.4) is expected to underestimate the spin density, while the simple scaling method (4.3) overestimates it by assuming that each s channel has the same scaling factor. To investigate and compare these estimators, we chose a small molecule for which using large plane wave cutoffs were computationally feasible.

4.4 Hydrogenated pyridine

The structure of hydrogenated pyridine is shown in Figure (4-1). It is similar to one of the six–membered rings of $C_{59}N$ that contains the nitrogen atom and the same problems with estimating the small spin density at the nuclei are expected to arise. The molecule has an unpaired electron and its orbital is also expected to be p-like near the nuclei. Figure (4-2) shows a contour of our calculated charge density, the p-like character near the nuclei is visible in the lower panel. We computed the spin density at the nuclei and varied the plane wave cutoff between 20 and 80 Hartrees.

The spin densities at the positions of the nuclei are shown in Figure (4-3) for the "PAW projector" method and the simple "scaling" method. Both methods are reasonably converged at 40 Hartrees, but the projector method is already almost converged at 20 Hartrees. The converged values do no agree with each other for the two methods, highlighting the incompleteness of the angular momentum expansion in (4.2). As mentioned above, the "PAW projector" method tends to underestimate the spin density, since it is a truncated sum of positive terms. On the other hand,

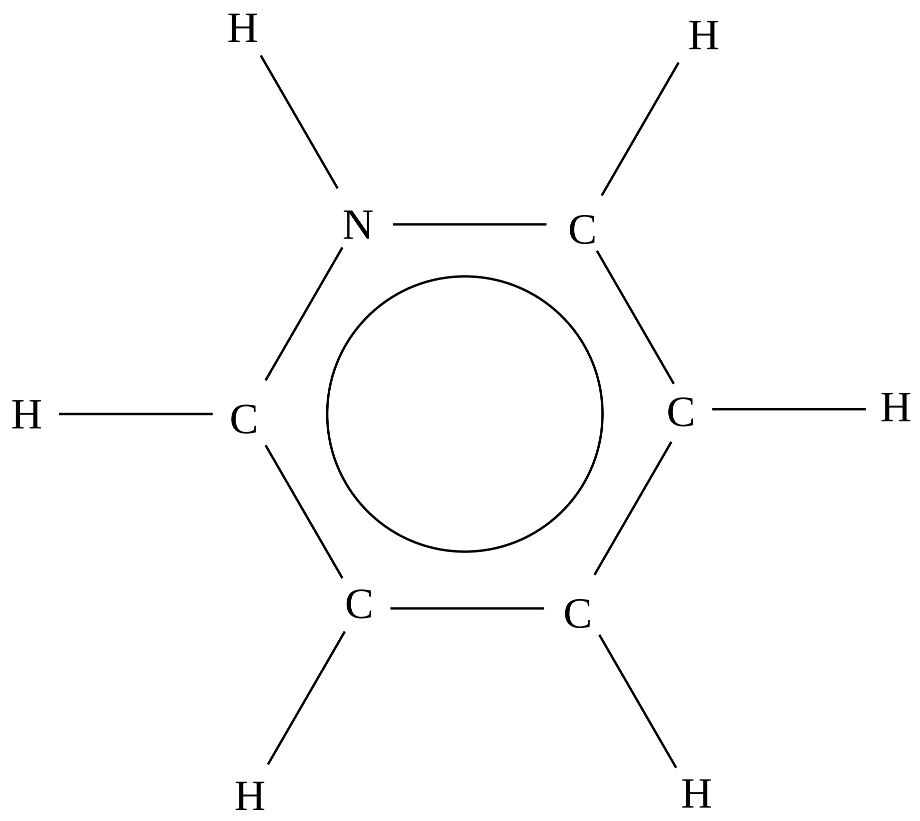

Figure 4-1: Structure of hydrogenated pyridine

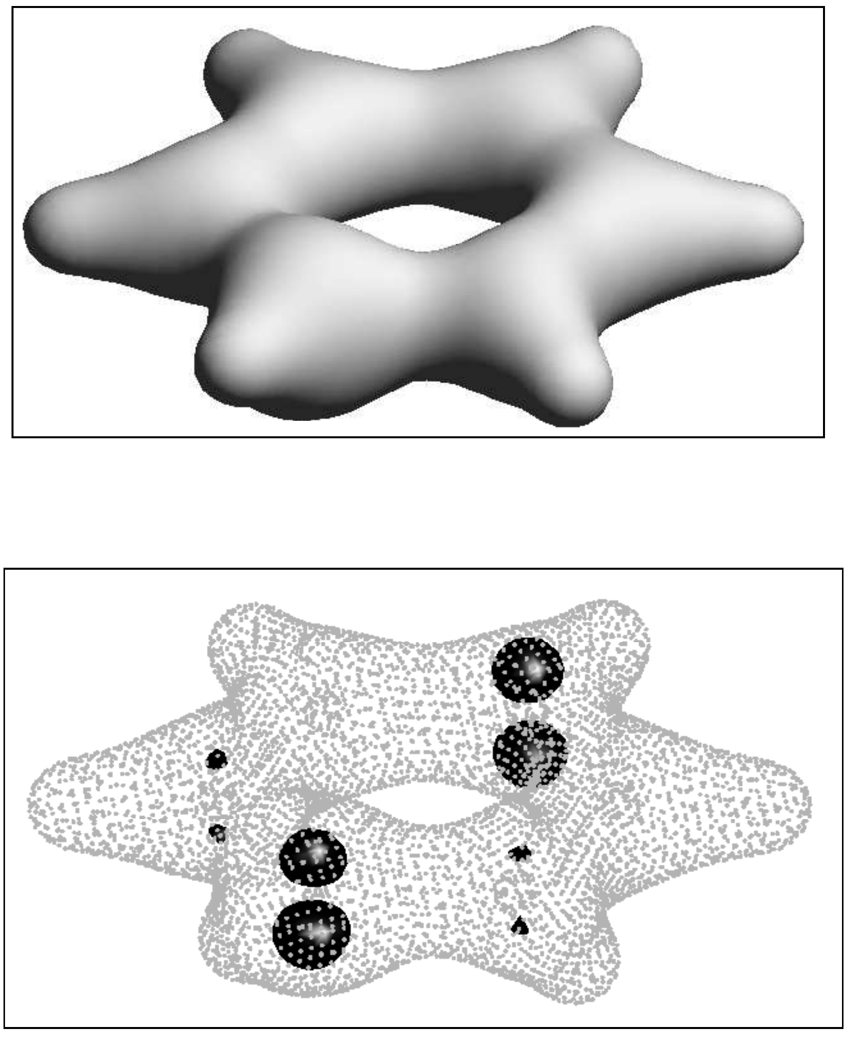

Figure 4-2: Contour of the charge density in hydrogenated pyridine (upper panel), and spin density (lower panel, with charge density as a dot cloud)

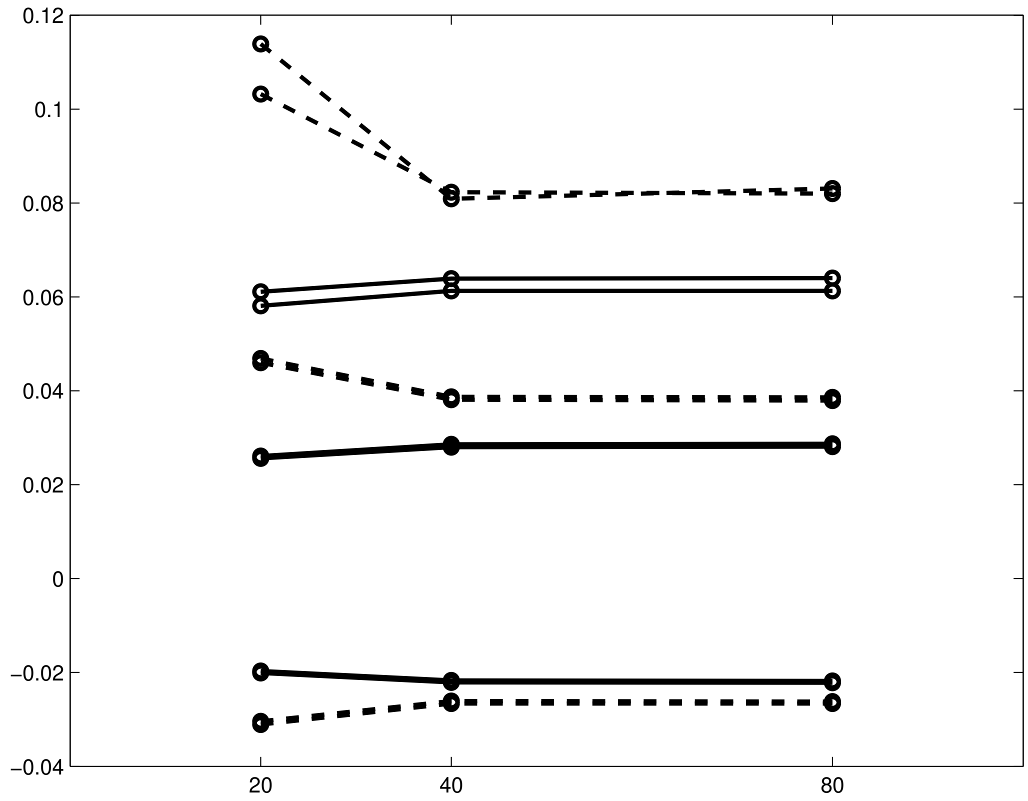

Figure 4-3: Spin densities at the six nuclear sites as a function of plane wave cutoff. The solid lines are computed with the "PAW projector" method, the dashed lines with the "scaling" method.

the "scaling" method overestimates the all–electron density, when it assumes the same scaling factor for all s channels. Thus we obtain upper and lower bounds for the all–electron spin density. It is important to emphasize that this incompleteness notwithstanding, with a 20 Hartree cutoff the "PAW projector" method gives a much better estimate of the true all–electron value. If one were to add another projector to the expansion, the spin density thus obtained should be very accurate.

4.5 Results

Now we turn to the $C_{59}N$ molecule. Figure (4-4) shows the contours of the charge density in gray, and the spin density in black, Figure (4-5) shows just the spin density. The significant delocalization of the unpaired electron is clearly visible, as well as the oscillations of the spin density as the distance increases away from the nitrogen atom. Table (4.1) shows our computed hyperfine coupling constants and the results using the Gaussian all–electron program which were published along with the experiment in [56]. It is clear that the full projection formula (4.4) is much more accurate, the error is at most 25% compared to experiment. It is interesting that the values obtained by simply rescaling the pseudo density are relatively close in agreement with the values from the Gaussian calculation. Both are overestimating the hyperfine coupling by over a factor of two.

Note that the experiment only measures the absolute value of the hyperfine coupling and that the identification of the experimental values with particular sites is actually inferred from the calculated values. This is one of the major use of such calculations.

The degree of delocalization can be determined from the relative values of the hyperfine parameters as a function of the distance from the maximum. Table (4.2) shows some relative values, which, due to cancellation of errors, are in general much

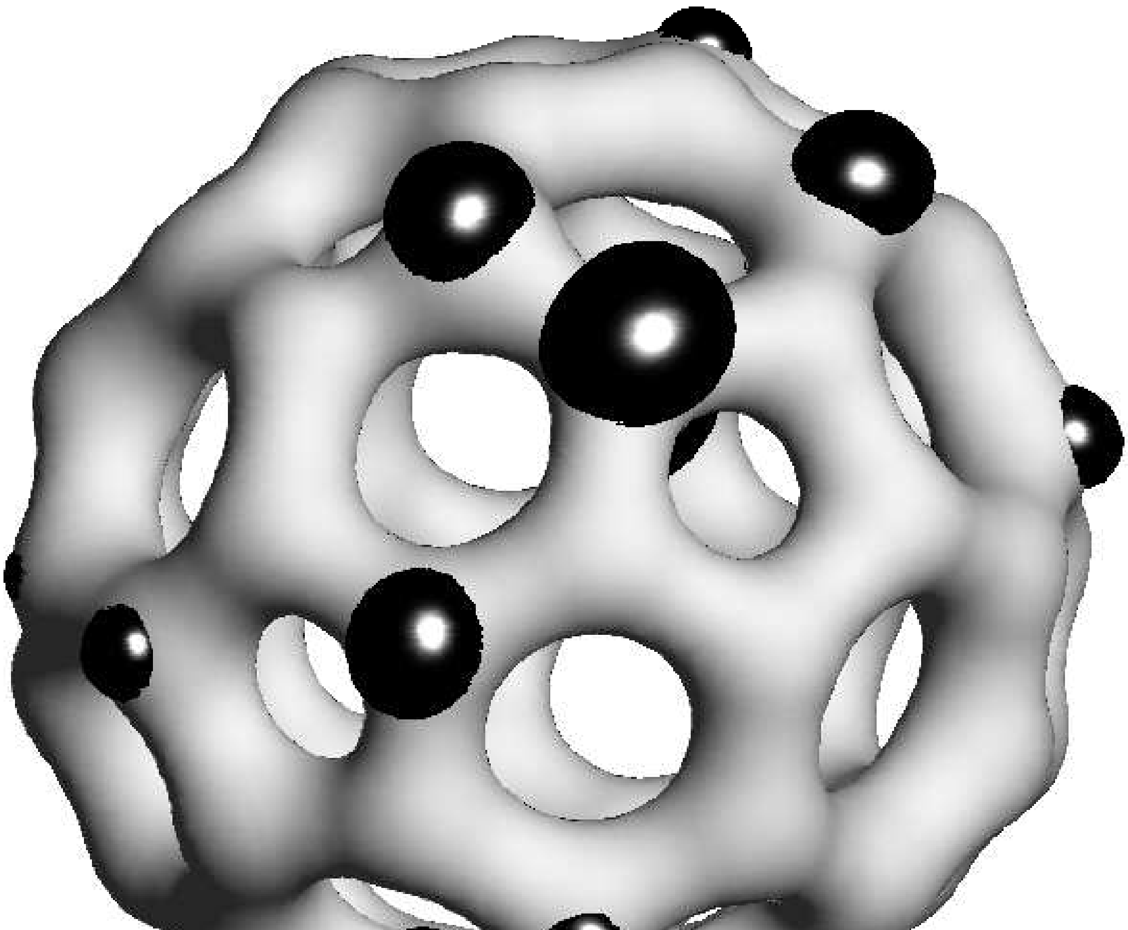

Figure 4-4: Contour of the charge density (gray) and the spin density (black). of the $C_{59}N$ molecule.

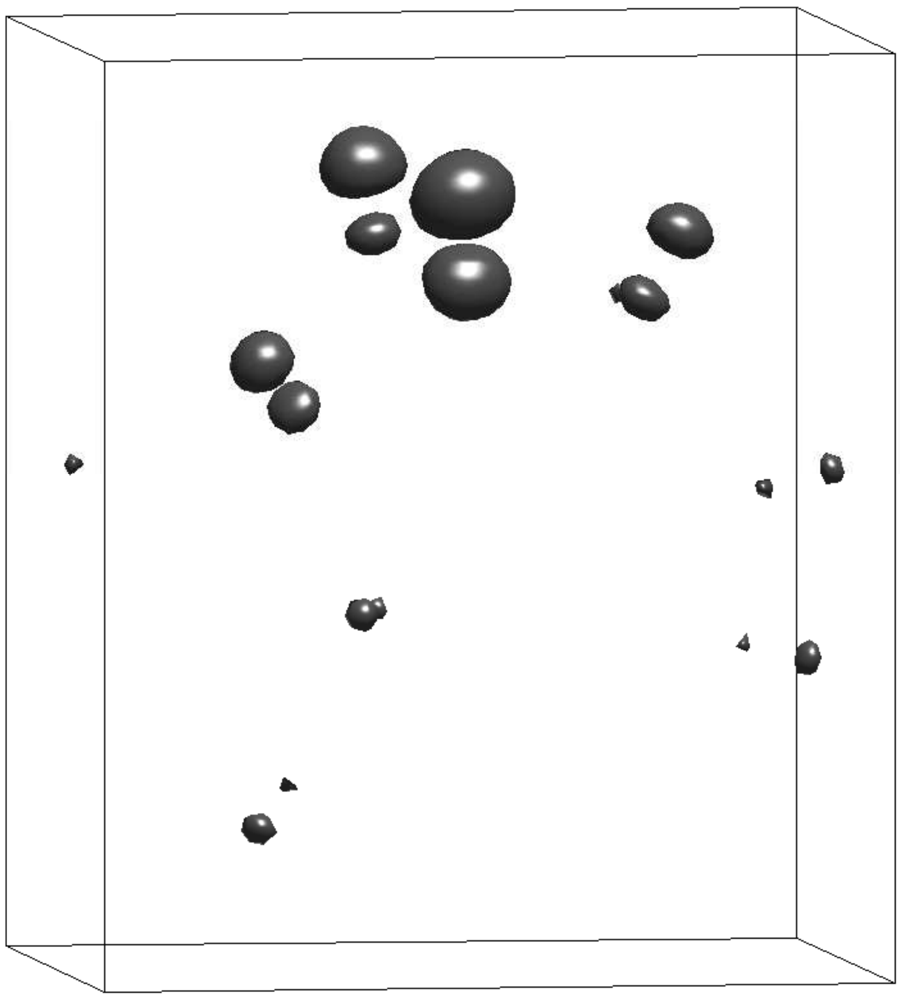

Figure 4-5: Contour of the spin density in the $C_{59}N$ molecule.

Atom	Expt.	PAW projector method	$n^{\mathrm{ps}}(R)$ scaling method	Gaussian B3LYP
N	0.36	0.33	0.62	0.87
C2	1.18	1.38	2.57	2.29
C3	0.48	-0.40	-0.74	-0.80
C4	0.52	0.55	1.02	0.88
C10	0.25	0.30	0.56	0.44

Table 4.1: Hyperfine coupling constants in mT. The present work is in the second and third column. The "PAW projector method" uses equation (4.4), the "scaling method" uses equation (4.3). Note that the experiment only measures the absolute value. The Gaussian and experimental values are from [56].

Atoms	Expt.	PAW projector method	Gaussian
C2/C4	2.3	2.5	2.6
C4/C10	2.1	1.8	2.0

Table 4.2: Relative hyperfine coupling constants, indicating the degree of delocalization.

better agreement with the experiment for all the computational methods.

4.6 Conclusion

In conclusion, we have computed isotropic hyperfine coupling constants for the azafullerene $C_{59}N$ and obtained a much better agreement with recently published experimental data than other calculations. We showed that the hyperfine parameters are very sensitive observables in this system because the unpaired electron is more or less in a p–like state with relatively small spin density at the nuclear sites. Thus a full projection–based formula was needed to obtain reasonable values for the hyperfine coupling, as opposed to a simple rescaling of the pseudo density, which has been the case earlier for semiconductors.

Chapter 5

Spin–Orbit interaction and the g tensor

Having seen in the last chapter that hyperfine couplings can be predicted reasonably well by the plane wave pseudopotential approach, we turn our attention to the g tensor. As mentioned before, in the ESR experiment a homogeneous external magnetic field breaks the Kramers degeneracy of the unpaired or conduction electrons and the system absorbs microwaves at a frequency corresponding to the difference in the energy levels. This energy difference is linear in the magnetic field for small fields and the proportionality constant is the g factor. Often the energy difference depends on the direction of the field, so g acquires a tensorial character. For a free electron, the value of the g factor is 2.0023. In molecules and solids, there is a shift away from the free electron value that arises due to the spin–orbit interaction.

There are two main difficulties associated with predicting g tensors in solids. First, the spin–orbit interaction which is responsible for the g shift is only appreciable near the cores of the atoms, where the potential is strong, therefore a good description of the electronic structure is necessary in the core region. Secondly, the introduction of a homogeneous magnetic field through the usual minimal coupling is problematic for an infinite solid, because the expectation value of the vector potential is infinite. This second problem is of course absent in the case of molecules and finite clusters, where g tensor calculations by all–electron methods are standard. To the best of our knowledge, to date, no *ab initio* calculation of the g tensor has been done in solids.

In this chapter, we collect the necessary ingredients from the literature to overcome these problems. We show how the spin–orbit interaction can be handled within the PAW formalism and test it by computing spin–orbit splittings in bulk silicon and germanium. The issue of the homogeneous magnetic in the case of extended systems was best addressed in a paper by de Graaf and Overhauser in 1969 [60]. We recall their results and show how it can be used in conjunction with the PAW method.

5.1 Background

In the quantum chemistry community, where g tensor calculations have become standard for molecules and clusters, it is customary to start with the single particle Hamiltonian,

$$h = \frac{p^2}{2} + V(\mathbf{r}),$$

where $V(\mathbf{r})$ is the self–consistent potential which includes the effect of nuclei and the other electrons. The spin–orbit interaction and the magnetic field are introduced perturbatively and the g tensor is extracted from the terms that are linear in the magnetic field. The standard result is given by

$$\delta g_n^{zz} = \frac{1}{2c^2}\langle\psi_n|\sum_{\mathbf{R}}(\tilde{r}_x^2 + \tilde{r}_y^2)\xi_{\mathbf{R}}(\tilde{r})|\psi_n\rangle + \frac{1}{c^2}\sum_{\mathbf{R}}\sum_{n'\neq n}\frac{\langle\psi_n|l_z|\psi_{n'}\rangle\langle\psi_{n'}|\xi_{\mathbf{R}}(\tilde{r})l_z|\psi_n\rangle}{E_{n'} - E_n}, \quad (5.1)$$

for the component of the g tensor in band n. Other components are obtained by permutation of indices. Here $\mathbf{R}$ runs through the nuclei, $\tilde{\mathbf{r}} = \mathbf{r} - \mathbf{R}$ is the position relative to a given nucleus and $\mathbf{l} = \tilde{\mathbf{r}} \times \mathbf{p}$ is the angular momentum. The self–consistent potential is assumed to be radially symmetric near the nuclei, with

$$\nabla V = \xi(r)\mathbf{r}.$$

An excellent derivation of (5.1) is given by A. J. Stone in [61].

In extended systems, however, the angular momentum $\mathbf{l}$ is undefined the and the expectation value of $\mathbf{A}$ over extended states which is needed in arriving at equation (5.1) is also ill defined. Moreover, the use of perturbation theory is questionable, since the Bloch states change qualitatively upon the introduction of an infinitesimal magnetic field. These problems were best addressed in the scheme of de Graaf and Overhauser[60]. They work with Bloch states that are eigenstates of a Hamiltonian that *includes* the spin–orbit interaction, given by

$$H_0 = \frac{p^2}{2} + V(\mathbf{r}) + \frac{1}{4c^2}\boldsymbol{\sigma} \cdot (\nabla V \times \mathbf{p}), \tag{5.2}$$

where $\boldsymbol{\sigma}$ are the Pauli spin matrices. The operator σ_z does not commute with H_0, thus it is necessary to work with 2-spinors explicitly. Given Bloch state solutions to (5.2), wave packets are constructed that are mostly upspin and mostly downspin. The difference in the energy of these wave packets in the presence of a small magnetic field give the g tensor. A special gauge is used which leaves the wave packets qualitatively unchanged upon the introduction of the homogeneous field, thus the energies of the wave packets can be computed by perturbation theory to first order in the field. The resulting expression for the g shift in a given band, designated by the subscript 0, is

$$\begin{aligned}
\delta g(\mathbf{k}) \;=\; & \frac{2}{i}\int d\mathbf{r}\,\left[\frac{\partial u^{\dagger}_{0\mathbf{k}\uparrow}}{\partial k_x}\boldsymbol{\pi}_y u_{0\mathbf{k}\uparrow} - \frac{\partial u^{\dagger}_{0\mathbf{k}\uparrow}}{\partial k_y}\boldsymbol{\pi}_x u_{0\mathbf{k}\uparrow}\right] \\
& +2\left[\left(X_{0\uparrow 0\uparrow} + \frac{\partial \alpha}{\partial k_x}\right)\boldsymbol{\pi}_{y,0\uparrow 0\uparrow} - \left(Y_{0\uparrow 0\uparrow} + \frac{\partial \alpha}{\partial k_y}\right)\boldsymbol{\pi}_{x,0\uparrow 0\uparrow}\right] \\
& +2(X_{0\uparrow 0\uparrow}k_y - Y_{0\uparrow 0\uparrow}k_x) + 2\frac{\mu_e}{\mu_B}(\sigma_{z,0\uparrow 0\uparrow} - 1),
\end{aligned} \tag{5.3}$$

for a field oriented along the z direction. The matrix elements of operators are taken between Bloch states, and X and Y are components of the periodic part of the coordinate operator,

$$\mathbf{X}_{ns,n's'} = i\int d\mathbf{r}\; u^{\dagger}_{ns\mathbf{k}}(\mathbf{r})\nabla_{\mathbf{k}}[u_{n's'\mathbf{k}}(\mathbf{r})],$$

where $u_{ns\mathbf{k}}$ is the periodic part of the two–component Bloch spinor in band n having spin s and wave vector $\mathbf{k}$ and is an eigenstate of (5.2). The generalized momentum is given by

$$\boldsymbol{\pi} = \mathbf{p} + \frac{1}{4c^2}\boldsymbol{\sigma}\times\nabla V$$

and α, the relative phase picked up by the downspin wave packet as the spin is flipped is given by

$$\begin{aligned}
Me^{i\alpha} \;=\; & \mu_e\int d\mathbf{r}\; u^{\dagger}_{0\uparrow\mathbf{k}}\sigma_x u_{0\uparrow\mathbf{k}} - \frac{i}{2c}\int d\mathbf{r}\;\frac{\partial u^{\dagger}_{0\mathbf{k}\downarrow}}{\partial k_y}(k_z + v_z + \pi_z)u_{0\uparrow\mathbf{k}} \\
& +\frac{i}{2c}\int d\mathbf{r}\;\frac{\partial u^{\dagger}_{0\mathbf{k}\downarrow}}{\partial k_z}(k_y + v_y + \pi_y)u_{0\uparrow\mathbf{k}},
\end{aligned}$$

where $\mathbf{v}$ is the group velocity of the wave packet in the energy band $E_0(\mathbf{k})$,

$$\mathbf{v} = \nabla_k E_0(\mathbf{k}).$$

5.2 Spin–orbit interaction and the PAW method

Having reviewed the necessary ingredients of a g tensor calculation, we turn to the implementation of the formulas within the Projector Augmented Wave approach. First, we need to compute wavefunctions that include the spin–orbit interaction. We do this by first order perturbation theory and evaluate the matrix elements of $H_{\rm so}$ between the unperturbed Bloch states, $\tilde{u}_{ns\mathbf{k}}$, from which the spin–orbit including states $u_{ns\mathbf{k}}$ can be calculated either by a standard sum over empty states or by minimizing an appropriate auxiliary functional. Substituting $\mathcal{O} = H_{\rm so}$ into equation (4.2) is problematic, because, as we mentioned above, the potential that enters into the spin–orbit Hamiltonian is self–consistent and thus depends on the pseudo wavefunctions. It is also not clear where the potential should come from in the case of the atomic matrix elements. If one used the fully self–consistent PAW method, then this is not an issue, since the single particle potential for the system is well defined everywhere. But in our *post processing* method, the atomic matrix elements are evaluated separately from the calculation in the extended system.

Suppose for a moment, that the all–electron self–consistent potential $V_{\rm sc}^{\rm ae}$ was available everywhere in space. Then the spin–orbit matrix element is given by (4.2) as

$$\begin{aligned}\langle \tilde{u}^{\rm ae}_{ns\mathbf{k}}|H_{\rm so}|\tilde{u}^{\rm ae}_{n's'\mathbf{k}}\rangle &= \langle \tilde{u}^{\rm ps}_{ns\mathbf{k}}|\frac{1}{4c^2}\boldsymbol{\sigma}\cdot(\nabla V^{\rm ae}_{\rm sc}\times\mathbf{p})|\tilde{u}^{\rm ps}_{n's'\mathbf{k}}\rangle \\ &+\sum_{R,ij}\langle \tilde{u}^{\rm ps}_{ns\mathbf{k}}|p_{R,j}\rangle\langle\phi^{\rm ae}_{R,j}|\frac{1}{4c^2}\boldsymbol{\sigma}\cdot(\nabla V^{\rm ae}_{\rm sc}\times\mathbf{p})|\phi^{\rm ae}_{R,i}\rangle\langle p_{R,i}|\tilde{u}^{\rm ps}_{n's'\mathbf{k}}\rangle \\ &-\sum_{R,ij}\langle \tilde{u}^{\rm ps}_{ns\mathbf{k}}|p_{R,j}\rangle\langle\phi^{\rm ps}_{R,j}|\frac{1}{4c^2}\boldsymbol{\sigma}\cdot(\nabla V^{\rm ae}_{\rm sc}\times\mathbf{p})|\phi^{\rm ps}_{R,i}\rangle\langle p_{R,i}|\tilde{u}^{\rm ps}_{n's'\mathbf{k}}\rangle.\end{aligned} \tag{5.4}$$

Even in the case of extended systems, the main contribution to the matrix elements come from the core region, where ∇V is large. Then we can approximate the matrix elements as integrals just over the core regions of the nuclei, so the first and third terms of equation (5.4) cancel. We write the self–consistent potential near the atomic core as

$$V_{\rm sc}^{\rm ae} = V_{\rm sc,atom}^{\rm ae} + \Delta V_{\rm sc},$$

where the "atom" subscript refers the self–consistent atomic result. If ΔV is small, we can approximate the matrix elements of $H_{\rm so}$ by the second term of equation (5.4) using just $V_{\rm sc,atom}^{\rm ae}$, which is readily available in the atomic calculation, as

$$\langle \tilde{u}_{ns\mathbf{k}}^{\rm ae} | H_{\rm so} | \tilde{u}_{n's'\mathbf{k}}^{\rm ae} \rangle \approx \sum_{R,ij} \langle \tilde{u}_{ns\mathbf{k}}^{\rm ps} | p_{R,j} \rangle \langle \phi_{R,j}^{\rm ae} | \frac{1}{4c^2} \boldsymbol{\sigma} \cdot (\nabla V_{\rm sc,atom}^{\rm ae} \times \mathbf{p}) | \phi_{R,i}^{\rm ae} \rangle \langle p_{R,i} | \tilde{u}_{n's'\mathbf{k}}^{\rm ps} \rangle. \quad (5.5)$$

Now $V_{\rm sc,atom}^{\rm ae}$ is exactly spherically symmetric, so the operator on the right hand side of equation (5.5) is just

$$|p_{R,j}\rangle \langle \phi_j | \frac{1}{4c^2} \xi(r) \mathbf{l} \cdot \boldsymbol{\sigma} | \phi_i \rangle \langle p_{R,i}|. \quad (5.6)$$

5.3 Numerical Test

There are a lot of approximations involved in arriving at (5.5). To test the validity of these, we computed the spin–orbit splitting in silicon and germanium at the Γ point using the LDA approximation. Such calculations have been done before in the context of the plane wave pseudopotential method[62, 63], but they necessarily involved the incorporation of a relativistic vector pseudopotential into the calculation. The present method is much simpler and here the focus is on testing the above approximations.

For germanium, we obtained a splitting of 0.31 eV, the experimental value is 0.3 eV. For silicon, our result is 0.053 eV, the true value is 0.044 eV. The error is larger for silicon, which is a lighter atom so the spin–orbit effects are overall much smaller. A recent calculation of the splitting in silicon using the all–electron FLAPW method and the GGA approximation to density functional theory gives 0.047 eV[64]. The level of agreement suggests that (5.5) is adequate for taking spin–orbit effects into account for heavier atoms. For lighter atoms the approximation of neglecting ΔV should still be good, but it would be necessary to use the full reconstruction equation (5.4), replacing $V^{\rm ae}$ by $V^{\rm ps}$,

$$
\begin{aligned}
\langle \tilde{u}^{\rm ae}_{ns\mathbf{k}}|H_{\rm so}|\tilde{u}^{\rm ae}_{n's'\mathbf{k}}\rangle \approx\; & \langle \tilde{u}^{\rm ps}_{ns\mathbf{k}}|\frac{1}{4c^2}\boldsymbol{\sigma}\cdot(\nabla V^{\rm ps}_{\rm sc}\times\mathbf{p})|\tilde{u}^{\rm ps}_{n's'\mathbf{k}}\rangle \\
& + \sum_{R,ij}\langle \tilde{u}^{\rm ps}_{ns\mathbf{k}}|p_{R,j}\rangle\langle \phi^{\rm ae}_{R,j}|\frac{1}{4c^2}\xi^{\rm ae}(r)\mathbf{l}\cdot\boldsymbol{\sigma}|\phi^{\rm ae}_{R,i}\rangle\langle p_{R,i}|\tilde{u}^{\rm ps}_{n's'\mathbf{k}}\rangle \\
& - \sum_{R,ij}\langle \tilde{u}^{\rm ps}_{ns\mathbf{k}}|p_{R,j}\rangle\langle \phi^{\rm ps}_{R,j}|\frac{1}{4c^2}\xi^{\rm ps}(r)\mathbf{l}\cdot\boldsymbol{\sigma}|\phi^{\rm ps}_{R,i}\rangle\langle p_{R,i}|\tilde{u}^{\rm ps}_{n's'\mathbf{k}}\rangle,
\end{aligned}
\tag{5.7}
$$

where $\xi^{\rm ae}$ and $\xi^{\rm ps}$ are defined by

$$
\begin{aligned}
\nabla V^{\rm ae}_{\rm sc,atom} &= \xi^{\rm ae}(r)\mathbf{r}, \\
\nabla V^{\rm ps}_{\rm sc,atom} &= \xi^{\rm ps}(r)\mathbf{r}.
\end{aligned}
$$

5.4 Putting it all together

In summary, the computational procedure to obtain the g tensor is as follows. Equation (5.5) or (5.7) is used to calculate the matrix elements of the spin–orbit Hamiltonian, then $u_{ns\mathbf{k}}$ is obtained by computing he first order shift in the wavefunctions $\tilde{u}_{ns\mathbf{k}}$. Then the following operators need to be evaluated over the Brillouin zone using equation (4.2),

$$
\begin{aligned}
\mathbf{X}_{0\uparrow,0\uparrow}(\mathbf{k}) &= \langle u_{0\uparrow\mathbf{k}}|i\nabla_{\mathbf{k}}|u_{0\uparrow\mathbf{k}}\rangle \\
\boldsymbol{\sigma}_{0\uparrow,0\uparrow}(\mathbf{k}) &= \langle u_{0\uparrow\mathbf{k}}|\boldsymbol{\sigma}|u_{0\uparrow\mathbf{k}}\rangle \\
\boldsymbol{\pi}_{0\uparrow,0\uparrow}(\mathbf{k}) &= \langle u_{0\uparrow\mathbf{k}}|(\mathbf{p}+\frac{1}{4c^2}\boldsymbol{\sigma}\times\nabla V)|u_{0\uparrow\mathbf{k}}\rangle
\end{aligned}
$$

Of these, $\boldsymbol{\pi}$ also contains a term involving ∇V, for which an approximation similar to (5.5) can be used.

Given these matrix elements, the g tensor can be computed using (5.3).

Bibliography

[1] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864.

[2] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).

[3] S. Goedecker and C. J. Umrigar, Phys. Rev. Lett. **81**, 866, (1998).

[4] J.P. Perdew and W. Yue, Phys. Rev. B **33**, 8800 (1986).

[5] Y.M. Juan and E. Kaxiras, Phys. Rev. B **48**, 14944 (1993).

[6] G. Ortiz, Phys. Rev. B **45**, 11328 (1992).

[7] A. D. Becke, Phys. Rev. A **38**, 3098 (1988).

[8] M. C. Payne *et al.*, Rev. Mod. Phys. **64** No. 4 (1992).

[9] E. R. Davidson, *Reduced Density Matrices in Quantum Chemistry*, Academic Press, New York (1976)

[10] F. Mauri *et al.*, Phys. Rev. B **47**, 9973 (1993).

[11] X.-P. Li *et al.*, Phys. Rev. B **47**, 10891 (1993).

[12] M. S. Daw, Phys. Rev. B **47**, 10895 (1993).

[13] P. Ordejon *et al.*, Phys. Rev. B **51**, 1456 (1995).

[14] E. Hernandez and M. J. Gillan, Phys. Rev. B **51**, 10157 (1995).

[15] W. Kohn, Phys. Rev. Lett. **76**, 3168 (1996).

[16] B. Farid *et al.*, Phys. Rev. B **48** 11602, (1993)

[17] G. D. Mahan, *Many–Particle Physics*, Plenum Press, New York, NY (1990).

[18] D. M. Ceperley and B. J. Alder, J. de Physique **C7 41** 295 (1980); D. M. Ceperley and B J. Alder, Phys. Rev. Lett. **45** 566 (1980).

[19] J. Cioslowski and K. Pernal, J. Chem. Phys. **111** 3396 (1999)

[20] M. Gell–Mann and K. Brueckner, Phys. Rev. **106** 364 (1957).

[21] Gábor Csányi and T. A. Arias, Phys. Rev. B

[22] M. Levy, in *Density Matrices and Densty Functionals*, edited by R. Erdahl and V. H. Smith Jr. (D. Reidel, 1987), p. 479

[23] J.P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77** 3865 (1996)

[24] V. V. Bulatov *et al.*, Phil. Mag. A **72** 453, (1995).

[25] H. Alexander in Dislocations in Solids, Elsevier Science Publishers (1986).

[26] H. Alexander and H. Teichler in Materials Science and Technology, **4** (ed W. Schröter), 249 (1991).

[27] C. Kisielowski–Kemmerich, Phys. Status Solidi (b) **161**, 11 (1990).

[28] J. Weber, Solid State Phenomena, **37-38**, 13 (1994).

[29] E. Weber and H. Alexander, J. de Physique **40**, 213 (1979).

[30] Wohler et al., J. Phys Chem Solids **31**, 1381 (1970).

[31] E. Weber et al., Inst. Phys. Conf. Ser. **31**, 266 (1977).

[32] Suezawa et al., Inst. Phys. Conf. Ser. **59**, Ch. 8 (1981).

[33] V. Kveder et al., Phys. Stat. Sol. (a) **72**, 701 (1982).

[34] E. G. Sieverts, Phys. Stat. Sol. (b) **120**, 11 (1983).

[35] C. Kisielowski–Kemmerich, G. Weber and H. Alexander, Thirteenth International Conference on Defects in Semiconductors, 1985, p. 387 and references therein; C. Kisielowski–Kemmerich, J. Elec. Mat. **14a** 387 (1985)

[36] P. Omling, E. R. Weber, L. Montelius, H. Alexander and J. Michel, Phys. Rev. B **32**, 6571 (1985).

[37] M. Heggie and R. Jones, Journal de Physique, **44 C4**, 43 (1983).

[38] M. Heggie and R. Jones, Philosophical Magazine B, **48**, 365 (1983).

[39] M. Heggie, in Dislocations (eds. P. Veyssière, L. Kubin, J. Castaing) Editions du CNSR p. 305 (1984)

[40] R. W. Nunes *et al.*, Phys. Rev. Lett. **77** 1516 (1996).

[41] J. Bennetto, R. W. Nunes and D. Vanderbilt, Phys. Rev. Lett. **79** 245 (1997).

[42] J. C. H. Spence, H. R. Kolar, Y. Huang and H. Alexander in Mater. Res. Soc. Symp. Proc. Vol. **408** 261 (1996).

[43] F. H. Stillinger and T. A. Weber, Phys. Rev. B **31** 5262, (1985).

[44] M. C. Payne *et al.*, Rev. Mod. Phys. **64** No. 4, 1045 (1992).

[45] J. Perdew and A. Zunger, Phys. Rev. B **23** 5048 (1981).

[46] D. M. Ceperly and B. J. Alder, Phys. Rev. Lett. **45** 566 (1980).

[47] L. Kleinmann and D. M. Bylander, Phys. Rev. Lett. **48**, 1425 (1982).

[48] J. R. K. Bigger *et al.*, Phys. Rev. Lett. **69** 2224 (1992).

[49] T. A. Arias, M. C. Payne, and J. D. Joannopoulos, Phys. Rev. Lett. **69**, 1077 (1992).

[50] T. Engeness and T. A. Arias, Phys. Rev. Lett. **79** 3006 (1997).

[51] L. C. Kimerling et al., Inst. Phys. Conf. Ser. **59**, Ch. 8 (1981)

[52] S. T. Pantelides, Phys. Rev. Lett. **57** 2979 (1986).

[53] M. S. Duesbery *et al.*, Phys. Rev. B **43** 5143 (1991).

[54] I. Lee and K. J. Chang, Phys. Rev. B **50** 18083 (1994).

[55] Technology Review, June 2001

[56] F. Fulop et al., Chem. Phys. Lett. **334** 233 (2001)

[57] W. Andreoni et al., Chem. Phys. Lett. **190** 159 (1992)

[58] J. C. Hummelen et al. in Topics in Current Chemistry **199** 93 (1999)

[59] P. E. Blöchl, Phys. Rev. B **50** 17953 (1994)

[60] A. M. de Graaf and A. W. Overhauser, Phys. Rev. **180** 701 (1969)

[61] A. J. Stone, (1962)

[62] M. S. Hybertsen and S. G. Louie, Phys. Rev. B **34** 2920 (1986)

[63] L. A. Hemstreet et al., Phys. Rev. B **47** 4238 (1993)

[64] Ramos et al., Phys. Rev. B **63** 165210 (2001)