

Towards Very Large Scale DFT Electronic Structure Calculations

Gerd Berghold

Max-Planck-Institut für Festkörperforschung

Stuttgart, 2001

Towards Very Large Scale DFT Electronic Structure Calculations

Von der Fakultät Physik der Universität Stuttgart zur Erlangung
der Würde eines Doktors der Naturwissenschaften

(Dr. rer. nat.)

genehmigte Abhandlung

vorgelegt von

Gerd Berghold

aus Wendlingen am Neckar

Hauptberichter:	Prof. Dr. M. Parrinello
Mitberichter:	Prof. Dr. A. Muramatsu
Promotionsausschussvorsitzender:	Prof. Dr. M. Dressel
Tag der mündlichen Prüfung:	11. Juli 2001

Max-Planck-Institut für Festkörperforschung

Stuttgart, 2001

Zusammenfassung

Während der letzten Jahrzehnte sind Ab-initio-Elektronenstruktur-Rechnungen zu einem wichtigen Werkzeug zur Untersuchung der strukturellen, dynamischen und elektronischen Eigenschaften von komplexen Systemen auf atomarem Niveau geworden. Vor allem der Einsatz moderner Computertechnologie ermöglicht heutzutage die quantenmechanische Behandlung von Systemen, die mehrere hundert Atome umfassen. Chemische Vorgänge können nun am Computer simuliert werden, ohne ein Reagenzglas zu benötigen. Gewürdigt wurde diese Entwicklung nicht zuletzt durch die Verleihung des Nobelpreises 1998 an John A. Pople und Walter Kohn für ihre theoretischen Arbeiten im Bereich der Quantenchemie bzw. der Dichtefunktionaltheorie.

Aufbauend auf der Arbeit von Kohn und Pople wurde eine kaum zu überschauende Vielzahl von Algorithmen entwickelt, um die elektronische Struktur von wechselwirkenden Systemen wenigstens näherungsweise zu beschreiben. Eine grobe Einteilung dieser Methoden wird im einleitenden Abschnitt gegeben.

Auf der einen Seite stehen die quantenchemischen Methoden, welche meist auf der Hartree-Fock Näherung basieren [1]. Wichtige Beispiele sind die CI-Methode („Configuration Interaction“) und störungstheoretische Ansätze. Nachteil der quantenchemischen Methoden ist, daß sie sehr empfindlich mit der Größe des untersuchten Systems skalieren. Günstiger verhalten sich, in Bezug auf den Rechenaufwand, die auf der Dichtefunktionaltheorie aufbauenden Methoden. Ihre theoretische Fundierung erhielt die Dichtefunktionaltheorie durch die Sätze von Hohenberg und Kohn [2], die bewiesen, daß die gesamte Information über ein System von Elektronen in einem äußeren Potential aus der Kenntnis der Elektronendichte abgeleitet werden kann.

Eine große Anzahl von Anwendungen der Dichtefunktionaltheorie, insbesondere im Bereich der Car-Parrinello-Molekulardynamik [3], beruht auf dem Ebene-Wellen-Ansatz für die Kohn-Sham-Orbitale und die Elektronendichte. Problematisch ist die

enorme Zahl von ebenen Wellen, die für eine akkurate Beschreibung benötigt werden. Ein weiterer Nachteil liegt in der nicht-lokalen Natur der ebenen Wellen. Dies führt dazu, daß leere Raumbereiche mit der gleichen Genauigkeit beschrieben werden wie mit Atomen besetzte Gebiete. Auf der anderen Seite vereinen ebene Wellen eine ganze Reihe von Vorteilen. So haben sie eine einfache funktionale Form und bilden eine orthonormale Basis, deren Vollständigkeit von nur einem Parameter abhängt. Ebene Wellen sind unabhängig von den Atompositionen, was die Berechnung der Hellmann-Feynman-Kräfte besonders einfach macht. Das Hartree-Potential ist im reziproken Raum lokal und daher mit ebenen Wellen leicht zu behandeln. Nicht zuletzt können viele algebraische Manipulationen durch den Gebrauch der Technik der schnellen Fourier-Transformation erheblich vereinfacht werden.

Gewissermaßen komplementär dazu ist die Situation bei den quantenchemischen Methoden, die die Kohn-Sham-Orbitale mit Basisfunktionen, die an den Atomen lokalisiert und meist in Gauß-Funktionen entwickelt sind, beschreiben. Dieser Ansatz kann die Lokalität der Gauß-Funktionen für die meisten Wechselwirkungen optimal ausnutzen, hat aber seine Schwäche in der Berechnung der Coulomb-Energie, die in den lokalen Basisfunktionen ausgedrückt eine nicht-lokale Kopplung darstellt.

Die Komplementarität der Vor- und Nachteile von Gauß-Funktionen und ebenen Wellen war die Motivation dafür, einen Dichtefunktional-Ansatz zu entwickeln, der sowohl Gauß-Funktionen als auch ebene Wellen verwendet, und dadurch möglichst die Vorteile des jeweiligen Funktionstyps nutzt, ohne daß seine Nachteile zum Tragen kommen. Als Resultat wurde von Lippert *et al.* die GAPW Methode („Gaussian Augmented Plane Wave“) entwickelt [4-6]. Der Ansatz, der ihr zugrundeliegt, beinhaltet die Verwendung von Gauß-Funktionen als Basisfunktionen für die Kohn-Sham-Orbitale und von ebenen Wellen und Gauß-Funktionen als Auxiliärbasis für die Darstellung der Elektronendichte. Auf der einen Seite kann mit diesem Ansatz das gesamte Instrumentarium der Quantenchemie zur rekursiven analytischen Integration von Matrixelementen zwischen Gauß-Funktionen genutzt, und auf der anderen Seite die Erfahrung beim Umgang mit ebenen Wellen und der Fourier-Transformation für die Berechnung des Hartree- und Austausch-Korrelations-Poten-

tials in der Ebene-Wellen-Darstellung eingesetzt werden. Der GAPW Ansatz führt dazu, daß alle Matrixelemente der Kohn-Sham-Matrix mit einem Rechenaufwand, der wie $O(N \log(N))$ mit der Systemgröße skaliert, berechnet werden können. Lediglich die Neuberechnung der Wellenfunktion aus der Kohn-Sham-Matrix im SCF-Zyklus („Self Consistent Field“), die durch Diagonalisierung der Kohn-Sham-Matrix realisiert ist, skaliert wie $O(N^3)$.

Die GAPW Methode wurde konzipiert, um die effiziente Simulation sehr großer Systeme zu ermöglichen. Obwohl der Ansatz in einer ersten Anwendung bereits seine Leistungsfähigkeit bewiesen hat [4], befindet sich das Projekt noch in der Entwicklungsphase. Erhebliche algorithmische Anstrengungen sind erforderlich, um mit über Jahre hinweg optimierten Programmpaketen, wie z.B. CPMD [7] oder Gaussian 94 [8], zu konkurrieren. Hauptziel dieser Arbeit ist es, solche Algorithmen zu entwickeln. Als Resultat entstanden vier, lose miteinander verbundene und in unterschiedlichen Bereichen der Dichtefunktionaltheorie angesiedelte, methodische Arbeiten, die jedoch alle dem gleichen Ziel dienen: Die Größenordnung der Systeme, die der GAPW Methode zugänglich sind, weiter zu steigern und dennoch die gleichen Maßstäbe an Zuverlässigkeit und Genauigkeit zu erfüllen wie etablierte Ab-initio-Elektronenstruktur-Programme.

Nach diesen einleitenden Erläuterungen werden im folgenden alle wichtigen Ergebnisse dieser Arbeit zusammengefaßt.

Im zweiten Kapitel führen wir eine Methode ein, welche die analytische Berechnung der Austausch-Korrelations-Wechselwirkung im Rahmen der Dichtefunktionaltheorie ermöglicht. Während viele Anteile des GAPW Energiefunktional mit Hilfe von rekursiver analytischer Integration berechnet werden können, wurde der Austausch-Korrelations-Anteil bisher ausschließlich unter Verwendung numerischer Verfahren integriert. Grund für diese Vorgehensweise ist die komplizierte algebraische Form vor allem der gradientenkorrigierten Austausch-Korrelations-Funktionale. Eine Alternative zu der numerischen Integration, basierend auf Ideen, die von Zheng und Almlöf entwickelt wurden [9,10], gründet sich auf einer analytischen Berech-

nung der Integrale durch die Anwendung matrixalgebraischer Verfahren. Die Vorgehensweise ist wie folgt. Zuerst wird eine Matrixdarstellung der Elektronendichte in einer Hilfsbasis berechnet. Anschließend werden die Integrale mit Hilfe von Basissatz-Vollständigkeitsrelationen und der spektralen Auflösung der Einheit als Matrixgleichungen geschrieben. Hierdurch treten Matrixfunktionen an die Stelle der Austausch-Korrelations-Funktionale.

Neuartig an unserem Ansatz ist die Verwendung von komplexen Kurvenintegralen zur Definition der auftretenden Matrixfunktionen. Dadurch gelingt es, neben den bereits von Zheng und Almlöf gefundenen Energieausdrücken, entsprechende analytische Formeln auch für das Austausch-Korrelations-Potential abzuleiten. Als Resultat erhält man ein System von in sich konsistenten Gleichungen, welche im Grenzfall einer vollständigen Basis exakt sind. Sowohl die Lokale-Dichte-Näherung als auch gradientenkorigierte Funktionale können verwendet werden.

Nicht nur methodische Gründe sprechen für solch einen Ansatz. Darüberhinaus wird durch die analytische Berechnung der Austausch-Korrelations-Wechselwirkung jegliches numerische Rauschen, welches bei den numerischen Methoden von den verwendeten Integrationsgittern herrührt, vermieden. Das ist deshalb von Bedeutung, da das numerische Rauschen die Konvergenz von Geometrie-Optimierungen erschwert.

In der durchgeführten Implementierung erweist sich die Frage nach der Vollständigkeit der verwendeten Hilfsbasis als kritischer Punkt. Während es im Fall der Lokalen-Dichte-Näherung möglich ist, die zur Darstellung der Kohn-Sham-Orbitale verwendeten Gauß-Basissätze auch als Hilfsbasis zu verwenden, muß im Fall der gradientenkorigierten Funktionale die Auxiliärbasis deutlich vergrößert werden, um die Ergebnisse der numerischen Integration befriedigend zu reproduzieren. Obwohl die erzielte Genauigkeit praktischen Zwecken genügt, wird dadurch doch der praktische Nutzen der Methode gemindert.

Im dritten Kapitel erweitern wir die GAPW Methode auf nicht-periodische Systeme. Gegenüber quantenchemischen Methoden mit ausschließlich lokalen Basisfunktionen

zeichnet sich die GAPW Methode durch die effiziente Berechnung der Coulomb-Wechselwirkung in der Ebene-Wellen-Darstellung im Impulsraum aus. Allerdings gelten im Rahmen dieser Vorgehensweise immer periodische Randbedingungen, so daß die Isolierung der periodischen Replica zur Simulation nicht-periodischer Systeme bisweilen sehr große Simulationszellen erfordert. Bei geladenen isolierten Systemen ist dieses Verfahren überhaupt nicht anwendbar. Die Berechnung der Coulomb-Wechselwirkung im Impulsraum generiert automatisch eine neutralisierende homogene Hintergrundladung.

Eine Lösungsmöglichkeit, entwickelt von P. E. Blöchl in Referenz [11], gründet sich auf der Tatsache, daß die Coulomb-Wechselwirkung zwischen räumlich entfernten Ladungsdichten vollständig durch die elektrostatischen Multipolmomente beschrieben werden kann. Durch Konstruktion von Punktladungen, welche die Multipolmomente der Elektronendichte reproduzieren, gelingt es so, den ausschließlich aus den periodischen Randbedingungen resultierenden Teil der Coulomb-Wechselwirkung zu substrahieren und dadurch die effiziente Berechnung der elektrostatischen Wechselwirkung im Impulsraum beizubehalten.

Wichtiger Bestandteil der Methode ist die Konstruktion der Punktladungen. Die Multipolmomente sind durch den Funktionswert und alle Ableitungen der Elektronendichte am Ursprung des reziproken Raumes vollständig bestimmt. Wegen der Diskretisierung des reziproken Raumes sind Informationen über die Ableitungen allerdings nur eingeschränkt zugänglich. Deshalb werden die Punktladungen näherungsweise entsprechend einer Anpassung von Gauß-Funktionen an die Elektronendichte nahe des Ursprungs des reziproken Raumes konstruiert.

Um obigen Ansatz für die GAPW Methode verwenden zu können, bedarf es einiger Modifikationen. Diese sind hauptsächlich technischer Natur und werden in Abschnitt 3.2.4 beschrieben.

Nachdem die GAPW Methode auf nicht-periodische Systeme erweitert wurde, sind nun auch geladene isolierte Systeme zugänglich. Um dies zu demonstrieren, werden Testrechnungen an einem H_3O^+ Kation und einem NH_2^- Anion vorgestellt. In

beiden Fällen ergibt sich quantitative Übereinstimmung mit den Rechnungen des quantenchemischen Programms Gaussian 94 [8]. Allerdings werden dazu relativ große Simulationszellen benötigt, was die Effizienz der Simulationen beeinträchtigt. Eine Erklärung hierfür liefert der Vergleich zwischen den analytisch berechneten Multipolmomenten und den Multipolmomenten der konstruierten Punktladungen. Da höhere Multipolmomente von dem Punktladungsmodell nur unzureichend reproduziert werden, muß die Simulationszelle groß genug sein, um die entsprechende Wechselwirkungen vernachlässigen zu können.

Als Anwendung wenden wir uns der Untersuchung von Substituenteneffekten in Ameisensäurederivaten zu. Hier lassen sich die berechneten Energiedifferenzen zwischen der jeweiligen Säure und ihrer konjugierten Base vollständig mit den experimentell bestimmten Säurestärken korrelieren.

Die Untersuchung komplexer Systeme mit Hilfe von Computersimulationen erfordert nicht nur die Entwicklung geeigneter Simulationsalgorithmen. Zusätzlich bedarf es auch Methoden zur Analyse der erhaltenen Resultate. In diesem Zusammenhang haben Wannier-Funktionen [12] in jüngster Zeit erhebliche Bedeutung erlangt. Sie ermöglichen es, die Sprache der Quantenmechanik in chemisch intuitive Bilder zu übersetzen. Als Beispiele seien die Identifizierung freier Elektronenpaare oder die Zuordnung von Valenzen genannt. Zusätzlich finden sie Anwendung in der Theorie der elektronischen Polarisierung und bei der Entwicklung von Algorithmen, welche linear mit der Systemgröße skalieren.

Trotz des breiten Anwendungsspektrums fehlt es an effizienten und allgemein anwendbaren Algorithmen, um Wannier-Funktionen zu berechnen. Inhalt des vierten Kapitels ist es, solche Algorithmen zu entwickeln.

Wannier-Funktionen können durch eine unitäre Transformation aus den Bloch-Orbitalen gewonnen werden. Die eigentliche Schwierigkeit besteht darin, die Transformation eindeutig zu definieren und anschließend effizient zu berechnen. Eine mögliche Definition gründet sich auf Konzepte, die von Mazari und Vanderbilt in Referenz [13] entwickelt wurden. Hier findet man die gesuchte Transformation mit Hilfe eines

Lokalisierungsfunktional, welches durch die aus der Transformation berechneten Wannier-Funktionen maximiert wird.

Im Rahmen dieser Arbeit gelingt es, das von Mazari und Vanderbilt vorgeschlagene Lokalisierungsfunktional auf Simulationszellen beliebiger Symmetrie zu verallgemeinern und einen Algorithmus abzuleiten, mit dessen Hilfe sich die gesuchte Transformation, und daraus resultierend die Wannier-Funktionen, iterativ berechnen lassen. Vorteil unseres Algorithmus ist, daß er sich mit effizienten numerischen Optimierungsmethoden, wie z.B. dem Verfahren der konjugierten Gradienten, kombinieren läßt, und sich dank seiner Allgemeinheit auch zur Berechnung lokalisierter Molekülorbitale in der Quantenchemie eignet. Andererseits läßt sich die Standardmethode der Quantenchemie zur Berechnung lokalisierter Molekülorbitale, die Methode der Orbitalrotationen, dahingehend verallgemeinern, daß sich damit auch Wannier-Funktionen berechnen lassen.

Testrechnungen zeigen, daß Wannier-Funktionen nun in Simulationszellen beliebiger Symmetrie zugänglich sind. Darüberhinaus wird ihre Berechnung durch die Anwendung numerischer Optimierungsmethoden erheblich beschleunigt. Auch die Methode der Orbitalrotationen erweist sich als überraschend effizient.

Als Anwendung werden die Wannier-Funktionen einer (10,0) Zig-Zag Nanotube berechnet. Hieraus erhält man ein intuitives Bild, wie die elektronischen Eigenschaften der Nanotube durch die Krümmung der Oberfläche, im Vergleich zu einer Graphitschicht oder zu einem C_{60} Fulleren, drastisch beeinflußt werden.

Im letzten Kapitel wenden wir uns der Neuberechnung der Wellenfunktion im SCF-Zyklus zu, die durch Diagonalisierung der Kohn-Sham-Matrix realisiert ist. Wie oben bereits diskutiert, skaliert die Diagonalisierung wie $O(N^3)$, während die Matrixelemente der Kohn-Sham-Matrix mit einem Rechenaufwand, der wie $O(N \log(N))$ mit der Systemgröße skaliert, berechnet werden können. Obwohl das kubische Verhalten erst bei großen Systemen dominiert, ist es für die GAPW Methode von entscheidender Bedeutung, hier alternative Verfahren anzuwenden. Der Grund ist, daß die GAPW Methode gerade zur Simulation sehr großer Systeme ausgerichtet

ist.

Zahlreiche Methoden, die sogenannten $O(N)$ Methoden, wurden in diesem Zusammenhang in den letzten Jahren entwickelt [14]. Obwohl diese Verfahren bereits in vielen Programmpaketen verwendet werden und ihre prinzipielle Anwendbarkeit unumstritten ist, sind doch erhebliche Probleme bekannt und ungelöst. Besondere Schwierigkeiten treten bei der Verwendung großer nicht-orthogonaler Basissätze auf. Der Grund liegt darin, daß hier lineare Abhängigkeiten in der Überlapp-Matrix auftreten und diese zu numerischen Instabilitäten und Konvergenzproblemen in den $O(N)$ Algorithmen führen. Da im Rahmen der GAPW Methode nicht-orthogonale Gauß-Funktionen verwendet werden und große Basissätze unvermeidlich sind, um Rechnungen hoher Qualität zu gewährleisten, ist die direkte Anwendung der $O(N)$ Methoden in unserem Fall problematisch.

Um $O(N)$ Methoden dennoch mit unserer GAPW Implementierung kombinieren zu können, verwenden wir ein Verfahren, welches auf Ideen von Lee und Head-Gordon basiert [15]. Hier wird aus einem großen Satz von Basisfunktionen, genannt die Sekundärbasis oder erweiterte Basis, eine Minimalbasis konstruiert, welche für das zu untersuchende System optimal ist. Die Vorgehensweise ist wie folgt. Die Minimalbasis ergibt sich aus einer linearen Transformation, angewendet auf die Sekundärbasis. Die optimale Minimalbasis ist nun dadurch ausgezeichnet, daß sie ein geeignetes Energiefunktional minimiert. Nach geschickter Parametrisierung der Transformation ist es möglich, das vorliegende Optimierungsproblem iterativ zu lösen. Lee und Head-Gordon verwenden hier Jacobi-Rotationen in Verbindung mit einem quadratisch konvergenten Newton-Raphson Verfahren. Problematisch an dieser Vorgehensweise ist, daß sie die Berechnung zweiter Ableitungen nach den zu optimierenden Parametern erforderlich macht. Für große Systeme, an denen wir vor allem interessiert sind, ist die Anzahl der zweiten Ableitungen entsprechend groß und ihre Handhabung sehr aufwendig.

Hier unterscheidet sich unsere Vorgehensweise von der Methode von Lee und Head-Gordon. Durch eine alternative Parametrisierung der linearen Transformation gelingt

es, die iterative Berechnung der Minimalbasis mit dem Verfahren der konjugierten Gradienten zu kombinieren und dadurch einen effizienten Algorithmus zu erhalten, welcher ausschließlich die Berechnung der ersten Ableitungen erfordert.

Der Grund, weshalb die so konstruierte Minimalbasis im Zusammenhang mit $O(N)$ Methoden interessant ist, liegt darin, daß sich einerseits qualitativ hochwertige Rechnungen durch eine große Sekundärbasis verwirklichen lassen und andererseits numerische Instabilitäten dadurch vermieden werden, daß man die $O(N)$ Methoden ausschließlich auf die Minimalbasis anwendet.

Durch Implementierung dreier unterschiedlicher $O(N)$ Methoden, der Tschebyscheff-Polynom-Entwicklung [16], der Dichtematrix-Minimierung [17] und der kanonischen Dichtematrix-Purifikation [18], gelingt es, einen Gesamtalgorithmus zu entwickeln, welcher vollständig linear mit der Systemgröße skaliert und darüberhinaus Rechnungen mit beliebig großen nicht-orthogonalen Basissätzen erlaubt.

Um den Algorithmus zu testen wurden umfangreiche Rechnungen durchgeführt. Insbesondere beinhalten diese die Untersuchung der strukturellen Eigenschaften von Chlorophyll A und eine Molekulardynamik-Simulation von Wasser. Die lineare Skalierung der verschiedenen $O(N)$ Methoden mit der Systemgröße wird schließlich am Beispiel langkettiger, unverzweigter Alkane demonstriert. Hier erweist sich die kanonische Dichtematrix-Purifikation als am effizientesten.

Literatur

- [1] A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry*, McGraw-Hill, New York, (1982).
- [2] P. Hohenberg and W. Kohn, Phys. Rev. B, **136**, 864, (1964).
- [3] R. Car and M. Parrinello, Phys. Rev. Lett., **55**, 2471, (1985).
- [4] G. Lippert, *Die GAPW-Dichtefunktional-Methode für Ab-Initio- Molekulardynamik-Simulationen*, Dissertation, Stuttgart, (1998).

- [5] G. Lippert, J. Hutter and M. Parrinello, *Theor. Chim. Acta*, **103**, 124, (1999).
- [6] M. Krack and M. Parrinello, *Phys. Chem. Chem. Phys.*, **2**, 2105, (2000).
- [7] J. Hutter *et al.*, computer code CPMD, MPI für Festkörperforschung and IBM Zürich Research Laboratory, (1995-1999).
- [8] Gaussian 94, Revision B.2, M. J. Frisch, G. W. Trucks, H. B. Schlegel, P. M. W. Gill, B. G. Johnson, M. A. Robb, J. R. Cheeseman, T. Keith, G. A. Petersson, J. A. Montgomery, K. Raghavachari, M. A. Al-Laham, V. G. Zakrzewski, J. V. Ortiz, J. B. Foresman, C. Y. Peng, P. Y. Ayala, W. Chen, M. W. Wong, J. L. Andres, E. S. Replogle, R. Gomperts, R. L. Martin, D. J. Fox, J. S. Binkley, D. J. Defrees, J. Baker, J. P. Stewart, M. Head-Gordon, C. Gonzalez, and J. A. Pople, Gaussian, Inc., Pittsburgh PA, (1995).
- [9] Y. C. Zheng and J. Almlöf, *Chem. Phys. Lett.*, **214**, 397, (1993).
- [10] Y. C. Zheng and J. Almlöf, *J. Mol. Struct.*, **388**, 277, (1996).
- [11] P. E. Blöchl, *J. Chem. Phys.*, **103**, 7422, (1995).
- [12] G. H. Wannier, *Phys. Rev.* **52**, 191, (1937).
- [13] N. Marzari and D. Vanderbilt, *Phys. Rev. B* **56**, 12847, (1997).
- [14] S. Goedecker, *Rev. Mod. Phys.*, **71** 1085 (1999).
- [15] M. S. Lee and M. Head-Gordon, *J. Chem. Phys.*, **107**, 9085, (1997).
- [16] S. Goedecker and M. Teter, *Phys. Rev. B*, **51**, 9455 (1995).
- [17] X. P. Li, R. W. Nunes and D. Vanderbilt, *Phys. Rev. B*, **47**, 10891 (1993).
- [18] A. H. R. Palser and D. E. Manolopoulos, *Phys. Rev. B*, **58**, 12704 (1998).

Contents

1	Introduction	4
1.1	Simulation Methods	4
1.2	Density Functional Theory	6
1.3	The Gaussian Augmented Plane Wave Method	8
1.3.1	The Electronic Density in the GAPW Formulation	9
1.3.2	The Electronic Energy Functional in the GAPW Formulation	11
2	Grid-free DFT Implementation of Exchange-Correlation Functionals	13
2.1	Introduction	13
2.2	Method	14
2.2.1	LDA	14
2.2.2	Gradient Corrections	17
2.3	Results and Discussion	18
2.4	Conclusions	21
3	Electrostatic Decoupling of Periodic Systems	24
3.1	Introduction	24
3.2	Method	25
3.2.1	Electrostatic Insulation of the Periodic Images	25
3.2.2	Construction of the Model Charge Density	27
3.2.3	Electron Potential and Atomic Forces	28
3.2.4	Implementation in the GAPW Method	29

3.3	Results and Discussion	29
3.4	Conclusions	35
4	General and Efficient Algorithms for Obtaining Maximally-Localized Wannier Functions	36
4.1	Introduction	36
4.2	Derivation of the Functional	38
4.3	A Generalized Localization Procedure	42
4.3.1	Orbital Rotations	46
4.3.2	Exponential Representation	47
4.4	Results and Discussion	51
4.5	Conclusions	62
4.6	Appendix	63
4.6.1	Determination of the Weights	63
4.6.2	Approximate Expressions for the Gradient and the Second Derivative	64
5	Polarized Atomic Orbitals for Linear Scaling Methods	66
5.1	Introduction	66
5.2	Polarized Atomic Orbital Method	67
5.3	Combination with Linear Scaling Methods	73
5.3.1	Sparsity of the Density Matrix	73
5.3.2	Implementation of a Combined PAO-O(N) Method	78
5.4	Results and Discussion	83

5.5 Conclusions 98

1 Introduction

1.1 Simulation Methods

Present-day computer technology has changed quite profoundly the way in which modern scientific research is conducted. This is due mainly to considerably progress in the scale and scope of computer simulations playing an increasingly important role in the study of complex systems.

In this respect a major role has been played by classical molecular dynamics (MD), which provides a set of numerical tools to solve Newton's equation of motion. Within a conventional MD framework [1], a functional form for interatomic forces is assumed on the basis of physical considerations, and potential parameters are fitted to reproduce given experimental or theoretical results. This approach has proven its usefulness for many cases but is not without drawbacks. One of the decisive ingredients for the simulation of a real system is the quality of the potentials used. The closer it is to reality, the more useful the simulation will be. Obviously, the predictive power of fitted potentials is limited to regions in configuration space included in the used data base. For instance, potentials fitted to describe liquid water can be inadequate to describe the properties of ice. It is therefore highly desirable to have alternative methods where potentials are generated in a consistent and accurate way.

In that context ab-initio electronic structure calculations provide a powerful solution, interatomic potentials are parameter-free and derived from first principles. Starting point of most ab-initio methods is the time-independent Schrödinger equation in the nonrelativistic Born-Oppenheimer approximation,

$$\mathcal{H}\Psi = E\Psi , \tag{1.1}$$

where E is the electronic energy, Ψ is the wave function, and \mathcal{H} is the electronic Hamilton operator,

$$\mathcal{H} = - \sum_{i=1}^N \frac{\Delta_i}{2} - \sum_{i=1}^N \sum_{A=1}^M \frac{Z_A}{r_{iA}} + \sum_{i=1}^N \sum_{j>i}^N \frac{1}{r_{ij}} . \tag{1.2}$$

Here, $r_{iA} = |\mathbf{r}_i - \mathbf{R}_A|$ and $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$ are the respective electron-nucleus and electron-electron distances, N and M are the number of electrons and nuclei, Δ is the Laplace operator, and Z_A denotes the charge of nucleus A . Atomic units are employed here and in the sequel.

An approximate way to solve Eq. (1.2) is provided by the Hartree-Fock (HF) method. Here, the wave function Ψ is written as a Slater determinant,

$$\Psi_{\text{HF}} = \frac{1}{\sqrt{N!}} \det[\psi_1 \psi_2 \dots \psi_N], \quad (1.3)$$

built from N occupied one-electron orbitals ψ_i . The variational principle states that the best wave function of this functional form, is one which gives the lowest possible energy, $E_{\text{HF}} = \langle \Psi_{\text{HF}} | \mathcal{H} | \Psi_{\text{HF}} \rangle$. By minimizing E_{HF} with respect to the ψ_i , one can derive an equation, called the HF equation, which determines the optimal one-electron orbitals.

The difference between the exact eigenvalue of the Schrödinger equation and the HF energy is termed the correlation energy. Thus, electron-correlation effects are determined in post-HF methods [2]. Of all methods to obtain the correlation energy, the method of configuration interaction (CI) is conceptually the simplest. The basic idea is to represent the wave function as a linear combination of Slater determinants, typically built from occupied and virtual one-electron orbitals of a HF calculation. A different systematic procedure for going beyond the HF approximation is perturbation theory (Møller-Plesset Perturbation Theory). Here, the HF Hamiltonian is treated as the zeroth-order Hamiltonian and the correlation energy is computed from a perturbation expansion. Further widespread procedures are the coupled-cluster approximation and Green's function techniques. However, the major drawback of all these methods is the unfavorable computational scaling with respect to the system size.

Due to the limitations in the transferability of empirical MD potentials and the use of time consuming post-HF methods, tight-binding (TB) [3,4] and semiempirical (SE) [5,6] techniques have been developed to simulate extended systems with moderate computational costs. SE approaches are normally formulated within the same

conceptual framework as ab-initio methods, but they neglect many smaller integrals to speed up the calculation. In order to compensate for the errors caused by these approximations, empirical parameters are introduced into the remaining integrals and calibrated against reliable experimental or theoretical reference data. The TB methods works by writing the eigenstates of the Hamiltonian in an atomic-like basis set, and replacing the exact many-body Hamiltonian operator with a parameterized Hamilton matrix. Both methods can only be successful if the used model retains the essential physics to describe the properties of interest. Provided this is the case, the parameterization can account for the quantum mechanical effects in an average way.

The position of density functional theory (DFT) [7] is somewhere between post-HF and SE/TB methods. With the improvement of both the methodology and the algorithms, DFT provides nowadays a way to incorporate electron correlation effects while allowing for a very efficient treatment of large systems. Since DFT is substantial in the following, the method is reviewed in more detail.

1.2 Density Functional Theory

DFT is based on two celebrated papers of Hohenberg and Kohn [8] and Kohn and Sham [9]. In the first it has been shown that the ground state energy of a non-uniform electron gas is a unique functional of the ground state density. This powerful theorem has been turned into a practical calculation scheme by Kohn and Sham who decomposed the electronic density $n(\mathbf{r})$ into mutually orthogonal one-electron orbitals, the so-called Kohn-Sham orbitals,

$$n(\mathbf{r}) = \sum_i f_i |\psi_i(\mathbf{r})|^2, \quad (1.4)$$

where f_i is the occupation number of state i . The electronic energy density functional can then be written as,

$$E[n] = \sum_i f_i \langle \psi_i | -\frac{\Delta}{2} | \psi_i \rangle + \int d\mathbf{r} V_{ext}(\mathbf{r}) n(\mathbf{r})$$

$$+ \frac{1}{2} \iint d\mathbf{r} d\mathbf{r}' \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} + E_{\text{XC}}[n] , \quad (1.5)$$

where the various terms represent, in order, the quantum kinetic energy, the interaction between electrons and the external potential $V_{\text{ext}}(\mathbf{r})$ generated by the nuclei, the Coulomb repulsion, and the rest of the electron density functional which is referred to as the exchange-correlation functional $E_{\text{XC}}[n]$.

Using the variational principle, with the orthonormality constraints,

$$\int d\mathbf{r} \psi_i^*(\mathbf{r}) \psi_j(\mathbf{r}) = \delta_{ij} , \quad (1.6)$$

it is possible to derive a Schrödinger-like equation for the Kohn-Sham orbitals [7],

$$\left(-\frac{\Delta}{2} + V_{\text{ext}}(\mathbf{r}) + V_{\text{H}}(\mathbf{r}) + V_{\text{XC}}(\mathbf{r}) \right) \psi_i = \epsilon_i \psi_i . \quad (1.7)$$

Here,

$$V_{\text{H}}(\mathbf{r}) = \frac{1}{2} \int d\mathbf{r}' \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \quad \text{and} \quad V_{\text{XC}}(\mathbf{r}) = \frac{\delta E_{\text{XC}}[n]}{\delta n(\mathbf{r})} , \quad (1.8)$$

are the Hartree and the exchange-correlation potential, and ϵ_i denotes the Kohn-Sham eigenvalue of electronic state i .

Eq. (1.7) is somewhat deceptive, in that it looks like a single particle Schrödinger equation. However, two features bring out the full many-body character of the problem. One is that Eq. (1.7) has to be solved self-consistently since both $V_{\text{H}}(\mathbf{r})$ and the $V_{\text{XC}}(\mathbf{r})$ depend on the density $n(\mathbf{r})$ which is a function of the orbitals $\psi_i(\mathbf{r})$. The other is the incomplete knowledge of $E_{\text{XC}}[n]$. For many years the most widely used scheme has been the so-called local density approximation (LDA),

$$E_{\text{XC}}[n] = \int d\mathbf{r} n(\mathbf{r}) \varepsilon_{\text{XC}}(n(\mathbf{r})) , \quad (1.9)$$

where $\varepsilon_{\text{XC}}(n(\mathbf{r}))$ is the exchange-correlation energy per electron at a point \mathbf{r} in a homogeneous electron gas, known with great accuracy from quantum Monte-Carlo calculations [10]. This approximation is obviously valid in the limit of slowly varying densities, but has proven its accuracy for a wide range of systems. More recently, however, new approximations to $E_{\text{XC}}[n]$ have been developed. These are the so-called gradient corrections (GC), which supplement the LDA term with one that

depends explicitly on the gradients of the density [11–13],

$$E_{\text{XC}}[n] = \int d\mathbf{r} n(\mathbf{r}) \varepsilon_{\text{XC}}(n, (|\nabla n|)) . \quad (1.10)$$

Although experience in GC is not as consolidated as that for LDA, the emerging picture is that in general GC is an improvement over LDA, especially for bonding energies [14] and the description of hydrogen bonds [15–17].

Having specified the theoretical framework, we turn now to the practical issue of numerically solving Eq. (1.7). In this context the most popular choice has been that based on the plane wave (PW) pseudopotential approach which was pioneered successfully by Cohen and collaborators [18]. There is, however a large number of reasons for wanting a different type of basis set. Some of the pitfalls of PW are that they are extremely demanding in terms of memory, they are wasteful when it comes to describe isolated molecular systems with large empty regions, and it is difficult to implement linear scaling methods using a PW basis. In order to remedy these deficiencies while at the same time preserving most of its advantages, Lippert et al. have recently proposed a novel scheme based on the most conventional type of quantum chemical basis functions, namely Gaussian functions. This scheme is called Gaussian augmented plane wave (GAPW) method [19–21] and the following section is devoted to its description.

1.3 The Gaussian Augmented Plane Wave Method

The GAPW method uses a Gaussian basis set to expand the Kohn-Sham orbitals, whereas an augmented PW basis set is introduced as an auxiliary basis set to represent the electronic charge density. The Gaussian basis functions make it possible to benefit from the efficient analytical integration schemes and screening algorithms of quantum chemistry. At the same time the use of a PW basis for the electron density permits efficient calculation of the Hartree energy using fast Fourier transform, thus circumventing one of the major bottlenecks of standard Gaussian based calculations. The basic idea is to divide the total density into a smooth extended part which is

represented in PW and parts localized close to the nuclei which are expanded in Gaussians. Using this representation of the charge, the localized and the extended part can be treated separately, leading to a very efficient computational scheme.

1.3.1 The Electronic Density in the GAPW Formulation

The basic idea of the GAPW method is to write the electronic density n as a sum of three contributions,

$$n = \tilde{n} - \tilde{n}^1 + n^1, \quad (1.11)$$

where \tilde{n} is smooth and distributed over all space, and,

$$n^1 = \sum_A n_A^1 \quad \text{and} \quad \tilde{n}^1 = \sum_A \tilde{n}_A^1, \quad (1.12)$$

are sums of the atom-centered contributions n_A^1 and \tilde{n}_A^1 which are hard and soft, respectively. Moreover, it is assumed that the difference between n_A^1 and \tilde{n}_A^1 is zero outside a spherical atomic region U_A . The atomic regions U_A of different atoms do not overlap. Inside the atomic region U_A the soft density \tilde{n} is equal to its atom-centered contribution \tilde{n}^1 ,

$$\tilde{n}(\mathbf{r}) = \tilde{n}^1(\mathbf{r}) \quad \text{for} \quad \mathbf{r} \in U_A, \quad (1.13)$$

and outside the atomic region, in the interstitial region I , \tilde{n} is equal to the total density n ,

$$\tilde{n}(\mathbf{r}) = n(\mathbf{r}) \quad \text{for} \quad \mathbf{r} \in I. \quad (1.14)$$

Therefore the following relations must be true based on the assumptions made,

$$n(\mathbf{r}) - \tilde{n}(\mathbf{r}) = 0 \quad \text{for} \quad \mathbf{r} \in I, \quad (1.15)$$

$$n_A^1(\mathbf{r}) - \tilde{n}_A^1(\mathbf{r}) = 0 \quad \text{for} \quad \mathbf{r} \in I, \quad (1.16)$$

$$\tilde{n}(\mathbf{r}) - \tilde{n}_A^1(\mathbf{r}) = 0 \quad \text{for} \quad \mathbf{r} \in U_A, \quad (1.17)$$

$$n(\mathbf{r}) - n_A^1(\mathbf{r}) = 0 \quad \text{for} \quad \mathbf{r} \in U_A. \quad (1.18)$$

Having divided the space into two parts, the atomic regions and the interstitial region, different types of basis functions can be used which are adapted to the

properties of the electronic density in the relevant region. In the atomic regions, especially close to the nuclei, the electronic density shows a strong variation. Therefore localized Gaussian functions are well suited to describe the density in these regions, whereas in the interstitial region a relatively soft density resides which can be easily expanded in a small number of PW.

Having established the basic notions of the GAPW method, we proceed with a detailed description how to construct the different parts of the density according to Eq. (1.11).

The total electronic density,

$$n(\mathbf{r}) = \sum_{\mu\nu} P_{\mu\nu} \varphi_{\mu}(\mathbf{r}) \varphi_{\nu}(\mathbf{r}) , \quad (1.19)$$

is expanded in a set of contracted Gaussian functions,

$$\varphi_{\mu}(\mathbf{r}) = \sum_a C_{a\mu} g_a(\mathbf{r}) , \quad (1.20)$$

which are used as an atomic basis set to represent the Kohn-Sham orbitals. $P_{\mu\nu}$ is a density matrix element, g_a is a primitive Gaussian function and $C_{a\mu}$ is the corresponding contraction coefficient. The strong variations of the electronic density close to the nuclei can be easily described by such an expansion. The soft part of the electronic density is represented by an auxiliary basis set of PW,

$$\tilde{n}(\mathbf{r}) = \frac{1}{\Omega} \sum_{|\mathbf{G}| < G_C} \tilde{n}(\mathbf{G}) e^{i\mathbf{G}\mathbf{r}} , \quad (1.21)$$

where Ω is the volume of the periodic cell. The plane wave expansion is truncated by the specification of a cutoff value for the kinetic energy,

$$E_C = \frac{1}{2} G_C^2 , \quad (1.22)$$

of the plane waves. The required cutoff value is relatively small, because only the soft part of the electronic density has to be described by plane waves. The one-center densities n^1 and \tilde{n}^1 are constructed using the primitive orbital basis functions g_a of the current atom A . For this purpose, the part contained in the atomic region U_A

has to be projected out of each orbital basis function φ_μ . Defining a new projector basis set $\{p_a\}$, the projection is performed by,

$$\langle p_b | \varphi_\mu \rangle = \sum_a C'_{a\mu} \langle p_b | g_a \rangle . \quad (1.23)$$

In analogy to the total electronic density the one-center density is written as,

$$n_A^1 = \sum_{ab \in A} \sum_{\mu\nu} C'_{a\mu} P_{\mu\nu} C'_{b\nu} g_a g_b . \quad (1.24)$$

Since the hard and soft basis functions coincide outside their atomic regions, the one-center expansion of the hard basis functions can also be applied to the soft one-center density,

$$\tilde{n}_A^1 = \sum_{ab \in A} \sum_{\mu\nu} \tilde{C}'_{a\mu} P_{\mu\nu} \tilde{C}'_{b\nu} g_a g_b . \quad (1.25)$$

Now all density contributions are summed up and the total electronic density n in the GAPW formulation reads,

$$\begin{aligned} n &= \tilde{n} - \tilde{n}^1 + n^1 \\ &= \frac{1}{\Omega} \sum_{|\mathbf{G}| < G_C} \tilde{n}(\mathbf{G}) e^{i\mathbf{G}\mathbf{r}} - \sum_{ab \in A} \sum_{\mu\nu} \left\{ \tilde{C}'_{a\mu} P_{\mu\nu} \tilde{C}'_{b\nu} + C'_{a\mu} P_{\mu\nu} C'_{b\nu} \right\} g_a g_b , \end{aligned} \quad (1.26)$$

which is equivalent to the formulation as a linear combination of atomic orbitals in Eq. (1.19) . Finally, it must be emphasized that the atomic regions U_A do not explicitly appear in the GAPW functional. They serve only as a conceptual tool during the derivation of the GAPW functional.

1.3.2 The Electronic Energy Functional in the GAPW Formulation

Starting point to derive the GAPW energy functional is Eq. (1.5), where the electronic energy of a molecular or crystalline system was defined as,

$$E[n] = E_T[n] + E_{\text{ext}}[n] + E_H[n] + E_{\text{XC}}[n] . \quad (1.27)$$

Blöchl [22] has shown that the Hartree and the exchange-correlation energy can be separated into independent global and local contributions using the set of relations introduced in Eq. (1.15). The described density partition can be applied

straightforwardly to an exchange-correlation energy functional,

$$E_{\text{XC}}[n] = E_{\text{XC}}[\tilde{n}] - \sum_A E_{\text{XC}}[\tilde{n}_A^1] + \sum_A E_{\text{XC}}[n_A^1], \quad (1.28)$$

which is semi-local. By contrast, the non-local character of the Hartree energy functional causes a more complicated calculation requiring the introduction of appropriate localized screening densities expanded in a set of hard Gaussian functions $g_A^{lm}(\mathbf{r})$,

$$n^0 = \sum_A n_A^0 = \sum_{lm} Q_A^{lm} g_A^{lm}, \quad (1.29)$$

and soft Gaussian functions \tilde{g}_A^{lm} ,

$$\tilde{n}^0 = \sum_A \tilde{n}_A^0 = \sum_{lm} Q_A^{lm} \tilde{g}_A^{lm}. \quad (1.30)$$

The Q_A^{lm} are defined as,

$$Q_A^{lm} = N q^{lm} [n_A^1 - \tilde{n}_A^1 + n_A^Z], \quad (1.31)$$

where $q_A^{lm}[n]$ is the multipole moment operator. n_A^Z is the charge density of the ionic core at atom A and N is a normalization constant. These densities exactly cancel the electrostatic multipole moments of the one-center densities and thus allow a separation of the contributions to the Hartree energy. Finally, Eq. (1.27) can be rewritten due to the density partition of Eq. (1.11) as,

$$\begin{aligned} E[n] = & E_{\text{T}}[n] + E_{\text{ext}}[n] + E_{\text{H}}[\tilde{n} + \tilde{n}^0] + \\ & \sum_A \left\{ E_{\text{H}}[n_A^1 + n_A^Z] - E_{\text{H}}[\tilde{n}_A^1 + n_A^0] \right\} - \\ & E_{\text{H}}[\tilde{n}^0] + E_{\text{H}}[n^0] + \int V_{\text{H}}[n^0 - \tilde{n}^0] \tilde{n} \, d\mathbf{r} + \\ & E_{\text{XC}}[\tilde{n}] + \sum_A \left\{ E_{\text{XC}}[n_A^1] - E_{\text{XC}}[\tilde{n}_A^1] \right\}. \end{aligned} \quad (1.32)$$

One comment is in order. In the first implementation of the GAPW method, the interaction with the ionic cores $E_{\text{ext}}[n]$ was described by a pseudopotential operator in order to integrate out the core electrons [19,20]. Recently, the GAPW method was extended to all-electron calculations [21].

2 Grid-free DFT Implementation of Exchange-Correlation Functionals

2.1 Introduction

The main result of section 1.2 can be stated as follows, we would be able to predict the electronic properties of many-particle systems very accurately if the exchange-correlation (XC) functional were in its exact form. However, the goal of obtaining this exact functional has not been accomplished, and approximations continue to be made. Beyond the simplest local density approximation (LDA, Eq. (1.9)), a variety of additional gradient corrections (GC, Eq. (1.10)) have been proposed [11–13]. However, one drawback of these methods is that they involve integration over complicated algebraic expressions precluding an analytic evaluation. Therefore, a rather different route has traditionally been chosen, involving numerical quadrature over a grid, where the grid points must be chosen with great care and where numerical noise due to limited accuracy is unfavorable.

An alternative to these traditional grid-based methods was proposed by Zheng and Almlöf [23,24], who use matrix techniques to calculate the required integrals analytically. The suggested approach is based on a matrix representation of the density in an auxiliary basis set. Within this representation, matrix elements of XC functionals of the density are obtained by simply calculating matrix functions. The additional use of basis set completeness relations and the spectral resolution of the identity operator allows analytic expressions for the XC energy and potential to be derived. The method is not limited to local XC functionals. Gradient corrected functionals can be treated in a similar way, using a matrix representation of the absolute value of the gradient.

It is important to stress that such an analytical approach is not only desirable from an esthetic point of view, it also avoids the noise resulting from numerical quadrature. Clearly, the use of a finite basis set introduces a new source of inaccuracy

but the error due to basis set incompleteness is smooth, independent of the choice of coordinate system and can be eliminated in a controlled way by increasing the basis set.

The aim of this chapter is to derive all the expressions required to implement this grid-free method, to present a number of test calculations, and to address the apparent problems.

2.2 Method

The starting point for the grid-free method is a matrix representation of the density computed from Eq. (1.4). If we expand the Kohn-Sham orbitals in a basis set (Latin letters) and use an auxiliary basis to calculate the matrix representation of the density (Greek letters) we obtain,

$$M_{\alpha\beta}[n] = \langle \alpha | n | \beta \rangle = \sum_{i,j} \mathcal{P}_{ij}(ij\alpha\beta) . \quad (2.1)$$

Here, $(ij\alpha\beta)$ is a generalized one-electron overlap integral and \mathcal{P}_{ij} is an element of the reduced, first order density matrix. In the framework of the GAPW method, the element reads, $\mathcal{P}_{ij} = C_{is}P_{st}C_{jt}$, according to Eq. (1.19) and Eq. (1.20). Other representations of the density (e.g. in auxiliary basis set methods) require a straightforward change of Eq. (2.1).

To derive the required expressions for the XC energy and potential it is convenient to first consider the local part.

2.2.1 LDA

Inserting the expansion of the Kohn-Sham orbitals into Eq. (1.9) and applying the spectral resolution of the identity operator with respect to the auxiliary basis, the local part of the XC energy can be rewritten,

$$E_{\text{XC}}[n] = \sum_{i,j} \mathcal{P}_{ij} \langle i | \varepsilon_{\text{XC}}(n) | j \rangle$$

$$\begin{aligned}
&= \sum_{i,j} \sum_{\alpha,\beta} \sum_{\mu,\nu} \mathcal{P}_{ij} \langle i|\alpha\rangle S_{\alpha\beta}^{-1} \langle \beta|\varepsilon_{\text{XC}}(n)|\mu\rangle S_{\mu\nu}^{-1} \langle \nu|j\rangle \\
&= \text{Tr}[\mathcal{P}\hat{S}S^{-1}M[\varepsilon_{\text{XC}}(n)]S^{-1}\hat{S}^T] .
\end{aligned} \tag{2.2}$$

Here, Tr denotes the trace of a matrix, S^{-1} is the inverse of the overlap matrix of the auxiliary basis, and, $\hat{S}_{i\alpha} = \langle i|\alpha\rangle$, is the overlap matrix between the basis set used to expand the electron orbitals and the auxiliary basis. Due to the appearance of the matrix, $M_{\alpha\beta}[\varepsilon_{\text{XC}}(n)] = \langle \alpha|\varepsilon_{\text{XC}}(n)|\beta\rangle$, Eq. (2.2) is not yet suited for practical use. The key to obtain a practical scheme, lies in the fact that we can obtain matrix elements of any function of the density by transforming $M[n]$ to an orthonormal basis ($= \tilde{M}[n]$), diagonalizing it, evaluating the functions of the eigenvalues, and transforming back to the original basis. Thus, the final form of the energy expression is given by,

$$E_{\text{XC}}[n] = \text{Tr}[\mathcal{P}\hat{S}X\varepsilon_{\text{XC}}(\tilde{M}[n])X\hat{S}^T] , \tag{2.3}$$

where, $X = S^{-\frac{1}{2}}$, is the transformation matrix from $M[n]$, in terms of the auxiliary basis, to the orthonormal representation $\tilde{M}[n]$. Since we are working with real basis functions, the transformation matrix is symmetric, which means that Eq. (2.3) is simplified using, $X = X^T$.

Based on Eq. (2.3) it is easy to compute the XC energy by carrying out the following steps:

1. Calculate the matrix representation of the density using Eq. (2.1).
2. Transform to an orthonormal basis,

$$\tilde{M}[n] = XM[n]X . \tag{2.4}$$

3. Calculate the matrix function by diagonalizing $\tilde{M}[n]$ and evaluate the function of the eigenvalues,

$$\varepsilon_{\text{XC}}(\tilde{M}[n]) = Y\varepsilon_{\text{XC}}(\Lambda)Y^T . \tag{2.5}$$

Here, Y^T is the eigenvector matrix of $\tilde{M}[n]$ with eigenvalues λ_k and, $\varepsilon_{\text{XC}}(\Lambda)_{kl} = \varepsilon_{\text{XC}}(\lambda_k)\delta_{kl}$.

4. Calculate the energy using Eq. (2.3).

In order to perform a DFT calculation we also need an expression for the XC potential, which is defined as the functional derivative of the energy, $V_{\text{XC}} = \frac{\delta E_{\text{XC}}[n]}{\delta n}$. The corresponding equation in matrix form reads,

$$\langle i|V_{\text{XC}}|j\rangle = \frac{\partial E_{\text{XC}}[n]}{\partial \mathcal{P}_{ij}} = (\hat{S}X\varepsilon_{\text{XC}}(\tilde{M}[n])X\hat{S}^T)_{ij} + \text{Tr}[\mathcal{P}\hat{S}X\frac{\partial \varepsilon_{\text{XC}}(\tilde{M}[n])}{\partial \mathcal{P}_{ij}}X\hat{S}^T] . \quad (2.6)$$

The first part of Eq. (2.6) is straightforward to calculate, but the second part must be handled with care. The reason is that a derivation of the function which involves a diagonalization of $\tilde{M}[n]$ is not possible because the needed unitary transformation Y itself depends on \mathcal{P} . However we can express the matrix function in an alternative way using a complex contour integral [25],

$$\begin{aligned} \frac{\partial \varepsilon_{\text{XC}}(\tilde{M}[n])}{\partial \mathcal{P}_{ij}} &= \frac{1}{2\pi i} \frac{\partial}{\partial \mathcal{P}_{ij}} \oint \varepsilon_{\text{XC}}(z)(z\mathbf{1} - \tilde{M}[n])^{-1} dz \\ &= \frac{1}{2\pi i} \oint \varepsilon_{\text{XC}}(z)(z\mathbf{1} - \tilde{M}[n])^{-1} \frac{\partial \tilde{M}[n]}{\partial \mathcal{P}_{ij}} (z\mathbf{1} - \tilde{M}[n])^{-1} dz \\ &= Y \frac{1}{2\pi i} \oint \varepsilon_{\text{XC}}(z)(z\mathbf{1} - \Lambda)^{-1} Y^T \frac{\partial \tilde{M}[n]}{\partial \mathcal{P}_{ij}} Y (z\mathbf{1} - \Lambda)^{-1} dz Y^T . \end{aligned} \quad (2.7)$$

Here, $\mathbf{1}$ is the identity matrix and, $\Lambda_{kl} = \lambda_k \delta_{kl}$. Carrying out the integration over z one obtains,

$$\frac{1}{2\pi i} \oint \frac{\varepsilon_{\text{XC}}(z)}{(z - \lambda_k)(z - \lambda_l)} dz = \begin{cases} \mu_{\text{XC}}(\lambda_k), & \lambda_k = \lambda_l , \\ \frac{\varepsilon_{\text{XC}}(\lambda_k) - \varepsilon_{\text{XC}}(\lambda_l)}{\lambda_k - \lambda_l}, & \lambda_k \neq \lambda_l , \end{cases} \quad (2.8)$$

where μ_{XC} denotes the derivative of $\varepsilon_{\text{XC}}(n(\mathbf{r}))$ with respect to $n(\mathbf{r})$. Performing some simple algebraic transformations, the second part of Eq. (2.6) becomes,

$$\text{Tr}[\mathcal{P}\hat{S}X\frac{\partial \varepsilon_{\text{XC}}(\tilde{M}[n])}{\partial \mathcal{P}_{ij}}X\hat{S}^T] = \text{Tr}[Y\{A, B\}Y^T\frac{\partial \tilde{M}[n]}{\partial \mathcal{P}_{ij}}] , \quad (2.9)$$

where the B_{kl} components are defined through Eq. (2.8), $\{A, B\}$ denotes a component-wise matrix multiplication and,

$$A = (\hat{S}XY)^T \mathcal{P} (\hat{S}XY) , \quad \frac{\partial \tilde{M}[n]}{\partial \mathcal{P}_{ij}} = X((\alpha\beta ij))X . \quad (2.10)$$

For the special case when the auxiliary basis coincides with the original basis the formulas can be simplified by using, $\hat{S}X = X^{-1} = S^{\frac{1}{2}}$.

2.2.2 Gradient Corrections

The approach outlined earlier must be slightly modified for gradient corrected functionals. Firstly all the non-local XC functionals that we have considered allow us to rewrite Eq. (1.10) using a function f depending solely on $n(\mathbf{r})$ and another function g depending only on a variable $s(\mathbf{r})$,

$$E_{\text{XC}}[n] = \int d\mathbf{r} n(\mathbf{r}) f(n(\mathbf{r})) g(s(\mathbf{r})) \quad \text{with} \quad s(\mathbf{r}) = \frac{|\nabla n(\mathbf{r})|}{n(\mathbf{r})^{\frac{4}{3}}}. \quad (2.11)$$

In close analogy with Eq. (2.2) and Eq. (2.3), an energy expression in matrix form can be derived,

$$E_{\text{XC}}[n] = \text{Tr}[\mathcal{P} \hat{S} X f(\tilde{M}[n]) g(\tilde{M}[s]) X \hat{S}^T]. \quad (2.12)$$

However, what is needed is a scheme to calculate $\tilde{M}[s]$. Such a scheme has already been proposed by Zheng and Almlöf [24]. Here we obtain a matrix representation of the x-component of the gradient through,

$$\tilde{M}[(\nabla n)_x] = \tilde{M}[(\nabla)_x] \tilde{M}[n] - \tilde{M}[n] \tilde{M}[(\nabla)_x], \quad (2.13)$$

where the tilde denotes that we are already using an orthonormal basis set. With analogous equations for all components we derive first an expression for the absolute value of the gradient,

$$\tilde{M}[|\nabla n|] = \sqrt{\tilde{M}[(\nabla n)_x]^2 + \tilde{M}[(\nabla n)_y]^2 + \tilde{M}[(\nabla n)_z]^2}, \quad (2.14)$$

and finally for the matrix representation of $s(\mathbf{r})$,

$$\tilde{M}[s] = \tilde{M}[|\nabla n|] (\tilde{M}[n])^{-\frac{4}{3}}. \quad (2.15)$$

It is worth noting that working with Eq. (2.15) causes some problems. Firstly, $|\nabla n(\mathbf{r})|$ and $n(\mathbf{r})$ interpreted as operators should commute because they are multiplicative. However, when we use a finite basis set this property is no longer true. To overcome this deficiency, we replace Eq. (2.15) by the corresponding anticommutator relation. Secondly a product of two positive definite matrices is not necessarily positive definite. This means that when we diagonalize the matrix representation

of $s(\mathbf{r})$ to calculate a matrix function, we cannot expect to obtain merely positive eigenvalues. Unfortunately, as discussed in the next section, the use of large basis sets appears inevitable to handle this problem.

Nevertheless, applying the matrix techniques introduced in the last section, also the non-local XC energy can be computed analytically by the use of the equations above.

To calculate now the potential we get an expression resembling Eq. (2.6) but with a third part including the derivative $\frac{\partial g(\tilde{M}[s])}{\partial \mathcal{P}_{ij}}$,

$$\begin{aligned} \langle i|V_{\text{XC}}|j\rangle &= (\hat{S}Xf(\tilde{M}[n])g(\tilde{M}[s])X\hat{S}^T)_{ij} + \text{Tr}[\mathcal{P}\hat{S}X\frac{\partial f(\tilde{M}[n])}{\partial \mathcal{P}_{ij}}g(\tilde{M}[s])X\hat{S}^T] \\ &+ \text{Tr}[\mathcal{P}\hat{S}Xf(\tilde{M}[n])\frac{\partial g(\tilde{M}[s])}{\partial \mathcal{P}_{ij}}X\hat{S}^T]. \end{aligned} \quad (2.16)$$

In agreement with Eqs. (2.7) to (2.9) we compute the third part through repeated use of the chain rule according to the Eqs. (2.15), (2.14) and (2.13) until we obtain nothing but terms containing the derivative $\frac{\partial \tilde{M}[n]}{\partial \mathcal{P}_{ij}}$. Adding up all terms allows the straightforward calculation of the potential in a recursive way.

It is important to notice that all derived equations are exact in the limit of a complete basis set. In a practical implementation, however, we have to use a finite basis set and the effect of this restriction is discussed in the next section.

2.3 Results and Discussion

The matrix approach has been implemented in the GAPW method (see section 1.3). Following Eq. (1.28), the XC energy is divided into a smooth part $E_{\text{XC}}[\tilde{n}]$, which is evaluated on a regular grid using plane waves, and two sums of one-center contributions $E_{\text{XC}}[\tilde{n}_A^1]$ and $E_{\text{XC}}[n_A^1]$ which are evaluated according to the grid-free approach. The corresponding one-center densities are expanded, in agreement with Eq. (2.1), in an uncontracted Gaussian type orbital (GTO) basis set. All calculations were done using a s^5p^2 basis for hydrogen and a $(sp)^5d^2$ basis for the remainder. To obtain satisfactory results, particularly when handling the gradient corrected functionals,

		LDA		B+LYP		B+P	
Mol.	Meth.	E_{XC}	E_{TOT}	E_{XC}	E_{TOT}	E_{XC}	E_{TOT}
H ₂	ana.	-0.6492	-1.1357	-0.6872	-1.1658	-0.6969	-1.1752
	num.	-0.6492	-1.1357	-0.6908	-1.1683	-0.6993	-1.1768
H ₂ O	ana.	-4.1211	-17.1792	-4.2202	-17.2591	-4.2778	-17.3147
	num.	-4.1211	-17.1791	-4.2283	-17.2659	-4.2862	-17.3215
HCN	ana.	-4.3733	-16.1653	-4.4440	-16.2172	-4.5304	-16.3003
	num.	-4.3733	-16.1653	-4.4513	-16.2238	-4.5384	-16.3073
N ₂	ana.	-4.7912	-19.8835	-4.8769	-19.9492	-4.9620	-20.0304
	num.	-4.7912	-19.8835	-4.8850	-19.9564	-4.9708	-20.0382
NH ₃	ana.	-3.5204	-11.6997	-3.5929	-11.7544	-3.6588	-11.8159
	num.	-3.5203	-11.6997	-3.6011	-11.7609	-3.6633	-11.8218
CO ₂	ana.	-8.4701	-37.7316	-8.6273	-37.8521	-8.7583	-37.9798
	num.	-8.4700	-37.7316	-8.6394	-37.8625	-8.7731	-37.9924
CH ₄	ana.	-3.0837	-8.0326	-3.1481	-8.0772	-3.2121	-8.1406
	num.	-3.0836	-8.0326	-3.1569	-8.0338	-3.2192	-8.1461
C ₂ H ₂	ana.	-3.9874	-12.4557	-4.0417	-12.4930	-4.1309	-12.5791
	num.	-3.9874	-12.4557	-4.0485	-12.4990	-4.1382	-12.5854

Table 2.1: Comparison of exchange-correlation and total energies of various molecules ($[E_H]$) obtained from the grid-free approach (ana.) respectively obtained from numerical quadrature over a grid (num.). The XC functionals considered are a pure LDA functional, an exchange functional by Becke (B), a correlation functional by Lee, Yang and Parr (LYP) and a correlation functional by Perdew (P).

Mol.	E_{XC}	E_{TOT}	Mol.	E_{XC}	E_{TOT}
H ₂	-0.6492	-1.1357	H ₂ O	-4.1218	-17.1796
HCN	-4.3741	-16.1659	N ₂	-4.7921	-19.8842
NH ₃	-3.5208	-11.7000	CO ₂	-8.4714	-37.7327
CH ₄	-3.0839	-8.0328	C ₂ H ₂	-3.9881	-12.4563

Table 2.2: Exchange-correlation and total energies of various molecules ($[E_H]$) obtained from the grid-free approach using the XC functional LDA. In contrast to table 2.1, the auxiliary basis coincides with the basis used to expand the one-center density.

enlarged auxiliary basis sets to compute the matrix representations are needed. For that, we employ an auxiliary basis of type $(sp)^7(pd)^4$ for hydrogen and of type $(spd)^7(df)^4$ otherwise. The grid-based numerical calculations were performed using a product grid consisting of 40 radial points and an angular part with 50 points chosen according to the Lebedev method [26,27].

A comparison of the results obtained from our grid-free approach and the results obtained from a grid-based numerical integration is shown in table 2.1. All calculations were done at fixed geometry and all energies are calculated from self-consistent densities. The XC functionals considered are a pure LDA functional in a Pade approximation [28], an exchange functional by Becke (B) [11], a correlation functional by Lee, Yang and Parr (LYP) [12] and a correlation functional by Perdew (P) [13]. It is evident that the results involving gradient corrected functionals are less satisfactory than for the LDA case. Although the accuracy achieved is sufficient for practical purposes, we cannot completely reproduce, even with that large auxiliary basis, the results of the numerical integration. This is even more unfavorable if one considers table 2.2. Here, we have revised the previous LDA calculations, but with an auxiliary basis coinciding with the original basis. Obviously, for the LDA case, a large auxiliary basis is not mandatory to reproduce the results of the numerical integration.

How can such a discrepancy be explained? Firstly, it appears that the gradient of the density calculated via Eq. (2.13) is not well represented by small basis sets. The next point is that the calculation of the matrix representation of $s(\mathbf{r})$ (Eq. (2.15)) involves a matrix representation of $n(\mathbf{r})^{-\frac{4}{3}}$. This expression becomes unbounded for small densities and we have no appropriate representation in this region. A further problem is that when we diagonalize the matrix representation of the variable $s(\mathbf{r})$ (Eq. (2.15)) in order to calculate a matrix function, using small basis sets, we obtain negative eigenvalues which are not physically meaningful. This is because a product of two positive definite matrices is not necessarily positive definite. However in our experience, with the enlarged auxiliary basis introduced above, the appearance of negative eigenvalues is avoided and all operators are reasonably represented.

The last statement is confirmed by a further investigation. Table 2.3 shows the structural parameters of various small molecules computed from optimized geometries. Here, an excellent agreement between the grid-free and the numerical approach, for local as well as for gradient corrected functionals, is evident. This suggests that the size of the used auxiliary basis is large enough to give reliable relative energies, even though the absolute energies differ from numerical integration.

2.4 Conclusions

We have derived all the expressions required to implement a grid-free method to compute local and gradient corrected XC functionals. The approach is based on a matrix representation of the density in an auxiliary basis set. The use of basis set completeness relations and the spectral resolution of the identity operator allows analytic expressions for the XC energy and potential to be derived. The key point of our approach is the definition of a matrix function via a complex line integral, allowing to generalize the analytic formulas of Zheng and Almlöf [23,24] for the XC energy to the potential.

All derived equations are exact in the limit of a complete basis set. In a practical implementation, however, a finite basis set has to be used and this restriction effects

		LDA		B+LYP		B+P	
Mol.	Par.	ana.	num.	ana.	num.	ana.	num.
H ₂	r(HH)	0.764	0.764	0.745	0.744	0.749	0.748
H ₂ O	r(OH)	0.971	0.972	0.967	0.968	0.966	0.965
	∠(HOH)	104.8	104.8	105.2	104.9	104.9	104.9
HCN	r(CH)	1.079	1.079	1.066	1.066	1.069	1.068
	r(CN)	1.149	1.149	1.144	1.144	1.145	1.145
N ₂	r(NN)	1.095	1.095	1.092	1.092	1.092	1.092
NH ₃	r(NH)	1.022	1.022	1.015	1.016	1.015	1.015
	∠(HNH)	107.1	107.1	107.3	107.3	107.2	107.3
CO ₂	r(CO)	1.163	1.163	1.163	1.164	1.161	1.161
CH ₄	r(CH)	1.096	1.096	1.088	1.088	1.089	1.088
C ₂ H ₂	r(CC)	1.200	1.200	1.192	1.193	1.194	1.194
	r(CH)	1.100	1.100	1.091	1.090	1.092	1.091

Table 2.3: Comparison of structural parameters [$\mathbf{rA},^\circ$] of various molecules obtained from the grid-free approach (ana.) respectively obtained from numerical quadrature over a grid (num.). The XC functionals considered are a pure LDA functional, an exchange functional by Becke (B), a correlation functional by Lee, Yang and Parr (LYP) and a correlation functional by Perdew (P).

the usefulness of our analytic approach. In the case of LDA, we can use the original GTO basis set to compute the matrix representations and the overall performance of the operator approach compares favorably with numerical integration. In the case of gradient corrected functionals, even with a large auxiliary basis set, we cannot completely reproduce the results from numerical integration. Although, the accuracy achieved is sufficient for practical purposes and we obtain reliable relative energies, the usefulness of such an implementation is questionable.

3 Electrostatic Decoupling of Periodic Systems

3.1 Introduction

The major difficulty of applying plane wave based electronic structure methods, such as the GAPW method (see section 1.3), to isolated molecules is that the periodic images of the charge distribution are automatically generated due to the discrete sampling of reciprocal space. The interaction energy between the periodic images, however, is undesired. The interaction can be divided into two contributions, the wave function overlap and the electrostatic interaction. Whereas wave function overlap can be made negligible in a straightforward way by separating the images by $5 - 7 \text{ \AA}$, the long-range nature prohibits using such an approach to determine the electrostatic interaction. Due to the importance of handling isolated systems, many solutions to that problem have been found and applied successfully [29–31].

Recently a different approach was proposed by Bloechl [32] that is useful not only to decouple the periodic images. It also provides a point charge model that can be used to couple the isolated charge density of the quantum mechanical calculation to an environment, e.g. in hybrid quantum mechanics/molecular mechanics approaches.

The guiding principle of the method is that the interaction energy between separated densities is determined entirely through their electrostatic multipoles. The electrostatic multipole moments can be expressed in reciprocal space by the value and all derivatives of the charge density at the origin. Unfortunately, the information on the derivatives is lost if reciprocal space is discretized, as is usually done in plane wave calculations. However, one can recover the information from an interpolation near the origin.

For computational reasons it is advantageous to first expand the density into atom-centered Gaussians and to construct the point charge model from the superimposed Gaussians. As mentioned above, it is sufficient to choose the Gaussians such that their superposition reproduces the density near the origin of reciprocal space and

thus its multipole moments.

The chapter is organized as follows. First we explain how the charge densities can be decoupled given a suitable model density in terms of superimposed Gaussians and how to obtain such a model density. In addition the expressions for the correction of the potential acting on the electrons and the forces acting on the nuclei are given. Next, the modifications needed to implement the present scheme in the GAPW method are described. Finally, applications to various systems are presented to demonstrate the reliability of the method.

3.2 Method

3.2.1 Electrostatic Insulation of the Periodic Images

To evaluate the electrostatic self-energy of an isolated charge density, we start out with a periodic array $n(\mathbf{r})$ of separated charge densities $n_V(\mathbf{r})$, each member of which is entirely localized within a unit cell V . The periodic array of unit cells should be space filling, and each member of $n(\mathbf{r})$ should be entirely localized within one of these cells. The self-energy of an isolated density is then given by,

$$E = \frac{1}{2} \int_V \int_V d\mathbf{r} d\mathbf{r}' \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}, \quad (3.1)$$

where the double integration is performed over a single unit cell V .

Let us now introduce a model charge density $\hat{n}_V(\mathbf{r})$, which is localized within the same volume V and which reproduces the multipole moment Q^{lm} of $n_V(\mathbf{r})$. The model density is written as a sum of Gaussians centered at atomic site \mathbf{R}_i ,

$$\hat{n}_V(\mathbf{r}) = \sum_i q_i g_i(\mathbf{r} - \mathbf{R}_i), \quad (3.2)$$

which are normalized such that they possess a charge of one,

$$g_i(\mathbf{r}) = \frac{1}{(\sqrt{\pi}r_{c,i})^3} \exp\left(-\left(\frac{r}{r_{c,i}}\right)^2\right). \quad (3.3)$$

Every atomic site may be the center of various Gaussians with different decay length $r_{c,i}$, chosen such that the tails of the Gaussians do not spill over the boundaries of the unit cell. Further details of the model density will be specified in the next section.

In addition to the model density, we construct a point charge model,

$$\bar{n}(\mathbf{r}) = \sum_i q_i \delta_i(\mathbf{r} - \mathbf{R}_i) , \quad (3.4)$$

such that each point charge equals the total charge of the Gaussians at that site. By construction, the multipole moments of the model charge density agree with those of the point charge model.

Using the point charge model, the self-energy of the isolated density can be written as the sum of four terms,

$$E = E_1 + E_2 + E_3 + E_4 , \quad (3.5)$$

where the individual terms are defined as follows,

$$\begin{aligned} E_1 &= \frac{1}{2} \sum_{V'} \int \int_{VV'} d\mathbf{r} d\mathbf{r}' \frac{(n(\mathbf{r}) + n_C)(n(\mathbf{r}') + n_C)}{|\mathbf{r} - \mathbf{r}'|} , \\ E_2 &= -\frac{1}{2} \sum_{V'} \int \int_{VV'} d\mathbf{r} d\mathbf{r}' \frac{(\bar{n}(\mathbf{r}) + n_C)(\bar{n}(\mathbf{r}') + n_C)}{|\mathbf{r} - \mathbf{r}'|} , \\ E_3 &= \frac{1}{2} \int \int_V d\mathbf{r} d\mathbf{r}' \frac{\bar{n}(\mathbf{r})\bar{n}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} , \\ E_4 &= -\frac{1}{2} \int \int_V d\mathbf{r} d\mathbf{r}' \frac{(n(\mathbf{r}) - \bar{n}(\mathbf{r})) n_C}{|\mathbf{r} - \mathbf{r}'|} . \end{aligned} \quad (3.6)$$

The off-diagonal elements, $V \neq V'$, in E_1 and E_2 cancel because the electrostatic interaction of separated charge distribution depends only on its multipole moments, which by construction are identical between $n(\mathbf{r})$ and $\bar{n}(\mathbf{r})$. The compensating constant charge background, $n_C = \frac{1}{V} \int_V d\mathbf{r} n(\mathbf{r})$, has been added, because otherwise E_1 and E_2 , which are the self-energies of periodic densities, are individually infinite when the system is charged.

The first term E_1 is the self-energy of the periodic array of isolated charge densities

and is evaluated in reciprocal space to be,

$$E_1 = \frac{2\pi}{V} \sum_{\mathbf{G} \neq 0} \frac{n(\mathbf{G}) n^*(\mathbf{G})}{G^2} . \quad (3.7)$$

The next term E_2 subtracts the self-energy of the periodic point charge model and is evaluated using standard Ewald summation techniques [1]. The self-energy of the isolated point charge model E_3 is a simple double sum over the charges from a single unit cell,

$$E_3 = \frac{1}{2} \sum_{(\mathbf{R}_i \neq \mathbf{R}_j) \in V} \frac{q_i q_j}{|\mathbf{R}_i - \mathbf{R}_j|} . \quad (3.8)$$

The evaluation of E_4 is more subtle. However, exploiting the properties of the point charge model one obtains [32],

$$E_4 = -\frac{\pi}{2V} \sum_{i,j;\mathbf{R}_i,\mathbf{R}_j \in V} q_i (r_{c,i}^2 + r_{c,j}^2) q_j . \quad (3.9)$$

Using the formulas above, one can finally write the correction to the electrostatic energy, $\Delta E = E_2 + E_3 + E_4$, relative to that of the periodic system as a quadratic form,

$$\Delta E = \sum_{i,j} q_i M_{ij} q_j . \quad (3.10)$$

3.2.2 Construction of the Model Charge Density

The model density is obtained from a fit to the original charge density, which is biased by a weight function such that the original density is retrieved near the origin of reciprocal space and thus the multipole moments are reproduced. A reasonable choice for the weight function is,

$$w(\mathbf{G}) = \begin{cases} 4\pi \frac{(G^2 - G_c^2)}{G^2 G_c^2}, & |G| < G_c , \\ 0, & |G| \geq G_c , \end{cases} \quad (3.11)$$

enhancing the importance of the low- G components, while ignoring any high- G components of the density. Because the weight function diverges at the origin, the $\mathbf{G} = 0$ term is treated separately as a constraint using the method of Lagrangian

multipliers. Consequently, the parameters of the model density q_i are obtained by minimizing,

$$F(q_i, \lambda) = \frac{1}{2} \sum_{\mathbf{G} \neq 0} w(G) |n(\mathbf{G}) - \sum_i q_i g_i(\mathbf{G})|^2 - \lambda (n(\mathbf{G} = 0) - \sum_i q_i g_i(\mathbf{G} = 0)) , \quad (3.12)$$

with respect to the coefficients q_i and the Lagrangian multiplier λ . A straightforward computation leads to,

$$q_i = \sum_j A_{ij}^{-1} (b_j - \frac{\sum_{k,l} A_{kl}^{-1} b_l}{\sum_{k,l} A_{kl}^{-1}}) , \quad (3.13)$$

with,

$$\begin{aligned} A_{ij} &= \sum_{\mathbf{G} \neq 0} w(G) g_i^*(\mathbf{G}) g_j(\mathbf{G}) , \\ b_i &= \sum_{\mathbf{G} \neq 0} w(G) \text{Re}[g_i(\mathbf{G}) n^*(\mathbf{G})] . \end{aligned} \quad (3.14)$$

3.2.3 Electron Potential and Atomic Forces

The correction to the electron potential is determined as the derivative of ΔE with respect to the charge density,

$$\Delta v(\mathbf{G}) = \frac{\partial \Delta E}{\partial n^*(\mathbf{G})} = \sum_i \frac{\partial \Delta E}{\partial q_i} \frac{\partial q_i}{\partial n^*(\mathbf{G})} . \quad (3.15)$$

Using Eq. (3.10) the potential acting on the charges of the Gaussians reads,

$$\frac{\partial \Delta E}{\partial q_i} = u_i = 2 \sum_j M_{ij} q_j . \quad (3.16)$$

The correction to the electron potential is then obtained by combining Eq. (3.16) with $\frac{\partial q_i}{\partial n^*(\mathbf{G})}$ computed from Eqs. (3.13) and (3.14),

$$\Delta v(\mathbf{G}) = \begin{cases} \frac{\sum_{i,j} A_{ij}^{-1} u_j}{\sum_{i,j} A_{ij}^{-1}}, & \mathbf{G} = 0 . \\ \sum_{i,j} w(G) g_i(\mathbf{G}) A_{ij}^{-1} (u_j - \frac{\sum_{k,l} A_{kl}^{-1} u_l}{\sum_{k,l} A_{kl}^{-1}}), & \mathbf{G} \neq 0 . \end{cases} \quad (3.17)$$

The correction to the atomic forces is determined as the derivative of ΔE with respect to the atomic position \mathbf{R}_k ,

$$\mathbf{F}_k = -\frac{d\Delta E}{d\mathbf{R}_k} = -\sum_{i,j} q_i \frac{dM_{ij}}{d\mathbf{R}_k} q_j - \sum_i \frac{dq_i}{d\mathbf{R}_k} u_i . \quad (3.18)$$

The first term is due to the electrostatic field of the periodic images and its evaluation is straightforward. The second term is analogous to the Pulay forces of electronic structure methods [33]. It vanishes if the Gaussians form a complete basis. Otherwise they describe the sensitivity of the results on the positions of the centers of the Gaussians.

The dependency of the charges on the atomic position is evaluated in two steps. The first neglects the constraint of charge conservation and the second applies the constraint,

$$\begin{aligned} \frac{dq_i^{(0)}}{d\mathbf{R}_k} &= \sum_j A_{ij}^{-1} \left(\frac{db_j}{d\mathbf{R}_k} - \sum_l \frac{dA_{jl}}{d\mathbf{R}_k} q_l \right), \\ \frac{dq_i}{d\mathbf{R}_k} &= \frac{dq_i^{(0)}}{d\mathbf{R}_k} - \sum_j A_{ij}^{-1} \frac{\sum_m \frac{dq_m^{(0)}}{d\mathbf{R}_k}}{\sum_{mn} A_{mn}^{-1}}. \end{aligned} \quad (3.19)$$

3.2.4 Implementation in the GAPW Method

The present approach has been implemented in the GAPW method. Here some slight modifications to the original decoupling method are required. Following Eq. (1.32), the long-range part of the GAPW electrostatic energy is provided by $E_{\text{H}}[\tilde{n} + \tilde{n}^0]$. Consequently, the correction to the electrostatic energy ΔE , relative to that of the periodic system, is merely applied to the term above.

As the spherical parts of the compensation density \tilde{n}^0 are themselves represented as spherical, atom-centered Gaussians, they are not included in the fit, but rather added directly to the model density fitted to the soft part of the electronic density \tilde{n} alone. In the present implementation we have ignored the usually small, nonspherical contributions of the compensation density.

3.3 Results and Discussion

In order to test the reliability of the present decoupling method we report the results of a series of all-electron calculations for two charged molecules, namely the

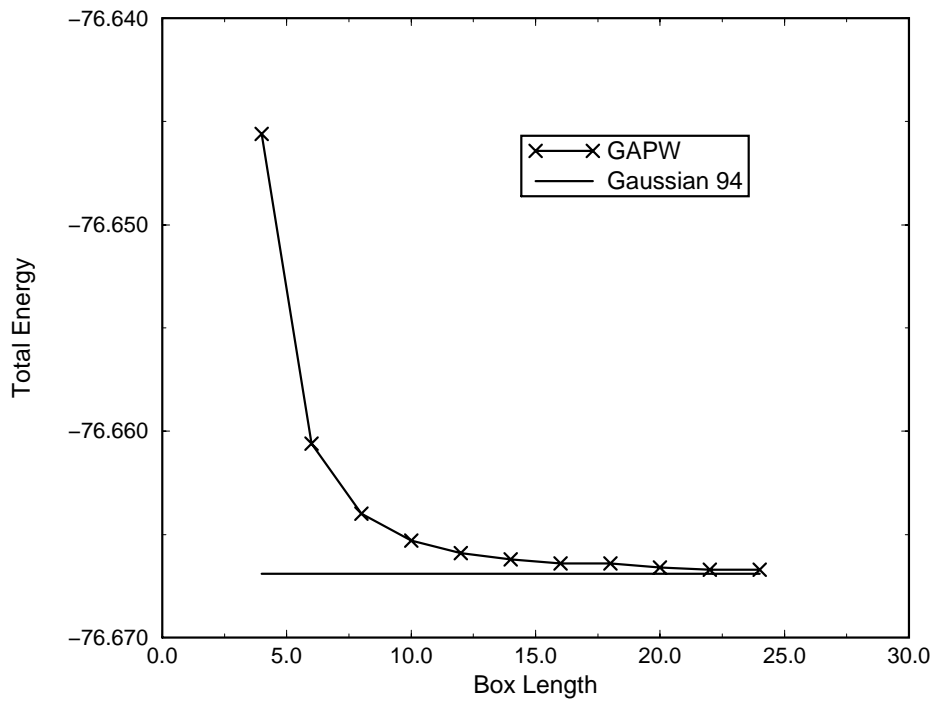


Figure 3.1: Convergence of the total energy [Hartrees] of the hydronium ion H_3O^+ with respect to the box length [rA]. All-electron calculations using the exchange-correlation functional BLYP are shown. As a comparison the total energy of an isolated hydronium ion is computed with Gaussian 94.

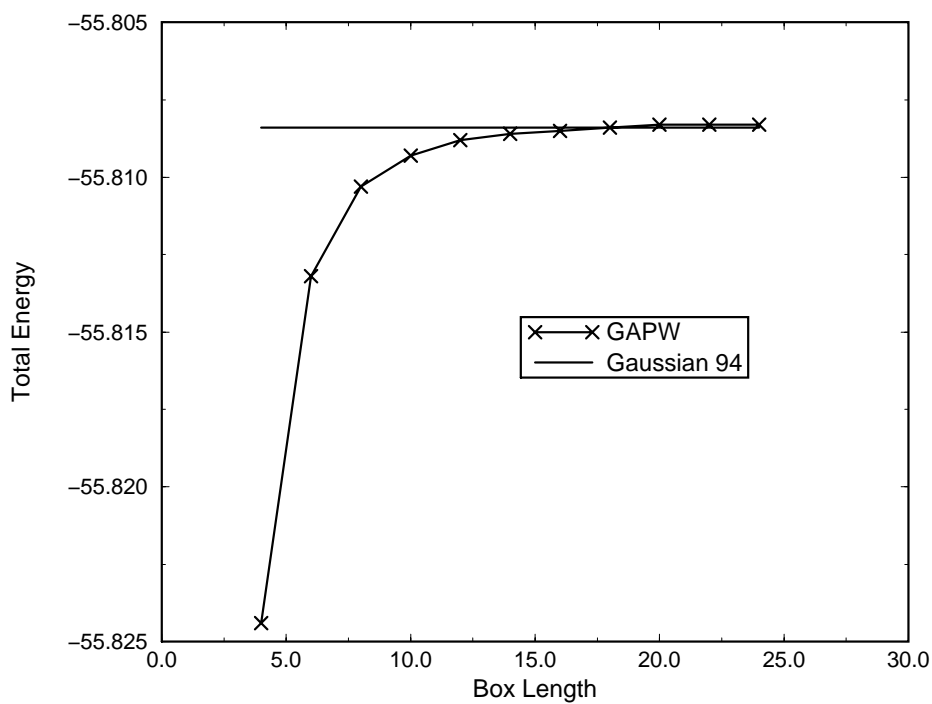


Figure 3.2: Convergence of the total energy [Hartrees] of the NH_2^- anion with respect to the box length [rA]. All-electron calculations using the exchange-correlation functional BLYP are shown. As a comparison the total energy of an isolated NH_2^- anion is computed with Gaussian 94.

Dipole Moments [Debye]					
X		Y		Z	
0.0	(0.0)	0.0	(0.0)	1.72	(1.75)
Quadrupole Moments [rA Debye]					
XX		XY		YY	
-2.49	(3.24)	0.0	(0.0)	-2.50	(3.23)
XZ		YZ		ZZ	
0.0	(0.0)	0.0	(0.0)	-6.13	(0.37)
Octapole moments [$\text{rA}^2 \text{ Debye}$]					
XXX		XXY		XYY	
1.71	(1.54)	-0.02	(-0.02)	-1.71	(-1.54)
YYY		XXZ		XYZ	
0.02	(0.02)	0.84	(0.74)	0.0	(0.0)
YYZ		XZZ/YZZ		ZZZ	
0.84	(0.74)	0.0	(0.0)	0.21	(0.09)

Table 3.1: Electrostatic multipole moments of the hydronium ion H_3O^+ computed from an all-electron calculation using the exchange-correlation functional BLYP. The moments computed from the point charge model are given in parentheses.

hydronium ion H_3O^+ and the NH_2^- anion. The results from the calculation of the corresponding isolated systems with the Gaussian 94 program [34] serve as a reference. All computations were done using the 6-31G* basis set, which is one of the well-known standard basis sets of Gaussian 94, and the gradient-corrected exchange-correlation functional of Becke [11] and Lee, Yang and Parr [12] (BLYP). Figure 3.1 and 3.2 show the convergence of the total energy of the H_3O^+ cation and the NH_2^- anion with respect to the box length of the cubic simulation cell. It is obvious that the converged total energies show excellent agreement with the results of the corresponding isolated systems, although large box lengths are needed to achieve convergence.

Dipole Moments [Debye]				
CO	HF	H ₂ O	H ₂ S	H ₂ O ₂
0.23 (0.22)	1.86 (1.86)	2.05 (2.04)	1.46 (1.36)	1.94 (2.01)
HCN	NH ₃	CH ₃ F	CH ₃ Cl	CHONH ₂
3.01 (2.99)	1.99 (1.98)	1.93 (1.96)	2.01 (2.01)	3.99 (3.97)

Table 3.2: Molecular dipole moments of various small molecules. The dipole moments computed from the point charge model are given in parentheses. All calculations were done using pseudopotentials and the exchange-correlation functional BLYP.

The requirement of employing such large simulation cells to converge the total energy is certainly a drawback. An explanation can be drawn from table 3.1. Here we have computed the electrostatic multipole moments of the hydronium ion to be compared with the moments calculated from the point charge model. Although the dipole moments are reproduced quantitatively and the octapole moments at least qualitatively, the point charge model fails to reproduce the quadrupole moments. Consequently, the simulation cell has to be large enough to make the interaction energy of the quadrupole moments negligible. These findings are confirmed by further calculations. We find the dipole moments always in good agreement with those computed from the point charge model (see table 3.2), but the higher multipole moments are often poorly reproduced. Thus, the use of large simulation cells appears inevitable to obtain accurate total energies.

As an application of the decoupling method we have investigated the substituent effects of various formic acid derivatives on the acid strength. For that we have computed in table 3.3 the total energy difference between the acid and its conjugate base using pseudopotentials and the exchange-correlation functional BLYP. A cubic cell of box length 15 \AA is employed. From Figure 3.1 and 3.2 we can expect to produce results which are in milihartree agreement in total energy from converged calculation. The chloro and the nitro group possess a dipole directed away from car-

Acid/Acid ⁻	ΔE [Hartrees]	pK_a
CCl ₃ COOH/CCl ₃ COO ⁻	-0.519	0.89
CH ₂ NO ₂ COOH/CH ₂ NO ₂ COO ⁻	-0.525	1.32
CHCl ₂ COOH/CHCl ₂ COO ⁻	-0.530	1.30
CH ₂ ClCOOH/CH ₂ ClCOO ⁻	-0.544	2.81
HCOOH/HCOO ⁻	-0.558	3.77
CH ₃ COOH/CH ₃ COO ⁻	-0.571	4.76

Table 3.3: Comparison of the total energy difference ΔE between an acid and its conjugate base and the negative logarithm of the acidity constant pK_a for various formic acid derivatives. Data have been computed using a cubic cell of box length 15 rA. Pseudopotentials and the exchange-correlation functional BLYP were employed.

bon and placing partial positive character on the carbon. By electrostatic attraction this tends to stabilize the conjugate base. On the other hand the dipoles introduced also destabilize the acid since the carboxyl group has a dipole also. Since the two are bound together with their positive dipole ends adjacent, they repel each other electrostatically. Both effects lead to an acid-strengthening in perfect agreement with a lowering of the measured pK_a values [35]. The range of these dipole effects can be seen in the acidity of trichloroacetic acid ($pK_a = 0.89$), about 10^4 times more acidic than formic acid ($pK_a = 3.77$) which corresponds to a stabilization energy of 39 milihartrees.

The opposite is true for the small electron-donating effect of the methyl group. Here, the conjugate base is destabilized while the acid is stabilized. Both effects lead to an acid-weakening having the consequence that formic acid is about 10 times more acidic than acetic acid. The corresponding stabilization energy amounts to 13 milihartrees.

3.4 Conclusions

A method to subtract the electrostatic interaction between periodic images of the densities of isolated molecules represented in a plane wave expansion has been presented. The subtraction scheme is carried out by the use of a point charge model reproducing the multipole moments of the system. The basic steps are to first expand the density into atom-centered Gaussians and to construct the point charge model from the superimposed Gaussians.

To implement the present approach into the GAPW method slight modifications are required. Here, one has to consider the splitting of the GAPW density in a smooth extended part and parts localized close to the nuclei. While the smooth extended part can be treated straightforwardly, the localized parts require special care.

The present approach has proven accurate in a comparison of the total energies of charged isolated systems with the results of the Gaussian 94 program. The requirement of employing large simulation cells is certainly a drawback. An explanation can be drawn from the failure of the point charge model to reproduce the higher multipole moments of the system. Consequently, the simulation cell has to be large enough to make corresponding interaction energy negligible.

As an application we have investigated the substituent effects of various formic acid derivatives on the acid strength. By comparing the total energy difference between the respective acid and its conjugate base we are able to reproduce the experimental acid strength qualitatively.

4 General and Efficient Algorithms for Obtaining Maximally-Localized Wannier Functions

4.1 Introduction

Computer simulations have enjoyed great success in recent years, with applications that range from materials science to chemistry and biochemistry. This is due mainly to considerably algorithmic progress, leading to fast and reliable programs. Consequently, the development of new simulation methods has played a central role in the previous sections. In the present section, however, we address to a further requirement modern computer technology has to fulfill. Namely, to provide analysis tools helping to understand the accomplished simulations. In that context Wannier functions [36] play an important role, thus, they are in the focus of what follows.

The representation of the electronic ground state in terms of localized Wannier orbitals provides a powerful tool in the study of periodic solids. Recent advances in the formulation of a theory of electronic polarization [37,38] and the development of linear-scaling methods [39] have rejuvenated the use of Wannier functions as an analysis tool. Namely, Wannier functions afford an insightful picture to the nature of chemical bonding and aid in the understanding of classical chemical concepts (*e.g.* non-bonding electron pairs or valency) in terms of quantum mechanics.

In spite of this wide field of applications, a lack of a practical, general, and efficient method to calculate Wannier functions is evident. This is in contrast to finite systems, where many different criteria for producing localized orbitals have been developed[40–43].

Wannier functions (WF) are defined in terms of a unitary transformation performed on the occupied Bloch orbitals (BO) [36]. One major problem in a practical calculation is their non-uniqueness. This is a result of the indeterminacy of the BO's, which are, in the case of a single band, only determined up to a phase factor, in the multi-band case, up to an arbitrary unitary transformation among all occupied orbitals at

every point in the Brillouin zone. As proposed recently by Marzari and Vanderbilt [44], one can resolve this non-uniqueness by requiring that the total spread of the localized function be minimal. This criterion is in close analogy with the Boys-Foster method [40] for finite systems, here one uses the spread defined through the conventional position operator. The new technique has been successfully applied to crystal systems and to small molecules within a general \mathbf{k} -point scheme[44,45]. An extension to disordered systems within the Γ -point approximation was recently performed[46]. This is of particular interest when one would like a localized orbital picture within the framework of Car-Parrinello molecular dynamics (CPMD).

Here we reexamine the problem focusing on the Γ -point approximation only. For this case, Silvestrelli [47] has derived a formula for the spread in three dimensions for a molecular dynamics cell of arbitrary shape. The formula is based on the formulation of Marzari and Vanderbilt. Recently Resta has proposed a formula for the spread in one dimension which reduces to that of Marzari and Vanderbilt in the limit of large cell size [48,49]. We generalize Resta's formulation to three dimensions and arbitrary molecular dynamics cells. In addition we derive a novel formula for the spread and examine the convergence properties of these three different formulations.

Upon minimization of the spread functional the appropriate unitary transformation to the localized orbitals can be calculated. With explicit knowledge of the spread functional we can derive the complete expressions required to implement the iterative minimization procedure beyond a steepest descent scheme. The steepest descents method, used in [44,46,47], performs reasonably for simple molecular systems, but it is known to have convergence problems when more complicated systems are analyzed. This can only be remedied by calculating the required gradient of the spread functional without any simplification, thus allowing the iterative calculation of maximally-localized orbitals with very efficient optimization schemes (*e.g.* conjugate gradient [50] and the direct inversion in the iterative subspace [51] method). Comparisons of the efficiency of the aforementioned optimization schemes to the well known Jacobi optimization procedure is made.

The present methodology can be straightforwardly generalized to finite (non-periodic) systems, thus providing a tool for Boys-Foster localization [40]. Furthermore we apply our optimization scheme also to the Pipek-Mezey localization [41] which is based on Mulliken population analysis.

4.2 Derivation of the Functional

We begin by reviewing the work of Resta [49]. In his treatment, the fundamental object for studying localization of an electronic state within Born-Von Karman boundary conditions is the dimensionless complex number,

$$z = \int_L dx \exp(i2\pi x/L) |\psi(x)|^2. \quad (4.1)$$

Here, L is the linear dimension, and $\psi(x)$ denotes the wavefunction. By considering the definition of the spread of the wavefunction to be $\Omega = \langle x^2 \rangle - \langle x \rangle^2$, where $\langle \dots \rangle$ denotes an expectation value, Resta has shown that to $O(1/L^2)$ the functional for the spread in one-dimension to be,

$$\Omega = \frac{1}{(2\pi)^2} \ln |z|^2. \quad (4.2)$$

One goal of this study is to generalize Eq. (4.1) to three-dimensions and obtain the appropriate generalization of Eq. (4.2). Thus, we choose to study the following dimensionless complex number within Born-Von Karman boundary conditions,

$$z_I = \int_V d\mathbf{r} \exp(i\mathbf{G}_I \cdot \mathbf{r}) |\psi(\mathbf{r})|^2. \quad (4.3)$$

Here, I labels a general reciprocal lattice vector, $\mathbf{G}_I = l\mathbf{b}_1 + m\mathbf{b}_2 + n\mathbf{b}_3$, where \mathbf{b}_α are the primitive reciprocal lattice vectors, the integers l , m , and n are the Miller indices, V is the volume of the supercell, and $\psi(\mathbf{r})$ denotes the wavefunction. We must find an appropriate function of the z_I 's that gives the three dimensional spread in the case of an arbitrary simulation cell. We proceed by noting that in a molecular dynamics simulation the cell parameters (primitive lattice vectors) to describe systems of general symmetry are given by \mathbf{a}_1 , \mathbf{a}_2 and \mathbf{a}_3 . It is convenient

to form a matrix of these cell parameters, $\vec{\mathbf{h}} = (\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3)$ where the volume V of the simulation cell is given by the determinant of $\vec{\mathbf{h}}$. It is also very useful to define scaled coordinates, $\mathbf{s} = \vec{\mathbf{h}}^{\leftrightarrow -1} \cdot \mathbf{r}$ that lie in the unit cube. In molecular dynamics simulations, this allows one to perform periodic boundary conditions for systems with general symmetry by first transforming to the unit cube, performing cubic periodic boundary conditions, and transforming back to the general cell with the action of $\vec{\mathbf{h}}$ [52]. One can also compute the reciprocal space vectors for systems of general symmetry with knowledge of the matrix of cell parameters. Thus, the I -th reciprocal lattice vector,

$$\mathbf{G}_I = 2\pi \left(\vec{\mathbf{h}}^{\leftrightarrow -1} \right)^T \cdot \hat{\mathbf{g}}_I . \quad (4.4)$$

Here, the superscript T denotes transposition, and $\hat{\mathbf{g}}_I = (l_I, m_I, n_I)$ is the I -th Miller index. We then substitute this expression into Eq. (4.3) and use the definition of \mathbf{r} to obtain,

$$z_I = \det \vec{\mathbf{h}} \int_0^1 ds \exp \left(i2\pi \hat{\mathbf{g}}_I^T \cdot \mathbf{s} \right) |\psi(\vec{\mathbf{h}} \cdot \mathbf{s})|^2 . \quad (4.5)$$

Note that the exponential in Eq. (4.5) is independent of any coordinate system. Following Resta [49] we can write the electron density in terms of a superposition of localized density and its periodic images, $|\psi(\vec{\mathbf{h}} \cdot \mathbf{s})|^2 = \sum_{\hat{\mathbf{m}}=-\infty}^{\infty} n_{\text{loc}}(\vec{\mathbf{h}} \cdot \mathbf{s} - \vec{\mathbf{h}} \cdot \mathbf{s}_0 - \vec{\mathbf{h}} \cdot \hat{\mathbf{m}})$. Here $\hat{\mathbf{m}}$ is a vector of integers and $\vec{\mathbf{h}} \cdot \mathbf{s}_0$ is the center of the distribution such that $\int_{-\infty}^{\infty} ds \vec{\mathbf{h}} \cdot \mathbf{s} n_{\text{loc}}(\vec{\mathbf{h}} \cdot \mathbf{s}) = 0$. Using the Poisson summation formula [53], we rewrite Eq. (4.5),

$$z_I = \exp \left(i2\pi \hat{\mathbf{g}}_I^T \cdot \mathbf{s}_0 \right) \hat{n}_{\text{loc}} \left(-2\pi \hat{\mathbf{g}}_I^T \cdot \vec{\mathbf{h}}^{\leftrightarrow -1} \right) , \quad (4.6)$$

where \hat{n}_{loc} denotes the Fourier transform of n_{loc} . Furthermore, since we are considering n_{loc} to be localized, its Fourier transform is smooth over reciprocal distances and we can be assured that it is well represented about $\hat{g}_I = 0$. We expand $\hat{n}_{\text{loc}}(-2\pi \hat{\mathbf{g}}_I^T \cdot \vec{\mathbf{h}}^{\leftrightarrow -1})$ to second order, obtaining,

$$\hat{n}_{\text{loc}}(-2\pi \hat{\mathbf{g}}_I^T \cdot \vec{\mathbf{h}}^{\leftrightarrow -1}) = 1 + \sum_{\alpha} \hat{g}_{\alpha, I} \frac{\partial \hat{n}_{\text{loc}}}{\partial \hat{g}_{\alpha, I}} \Big|_{\hat{g}_I=0} + \frac{1}{2} \sum_{\alpha, \beta} \hat{g}_{\alpha, I} \hat{g}_{\beta, I} \frac{\partial^2 \hat{n}_{\text{loc}}}{\partial \hat{g}_{\alpha, I} \partial \hat{g}_{\beta, I}} \Big|_{\hat{g}_I=0} + \dots \quad (4.7)$$

The second term in Eq.(4.7) is zero given our imposed condition $\langle \vec{\mathbf{h}} \cdot \mathbf{s} \rangle = 0$. Thus, we are left with,

$$\hat{n}_{\text{loc}}(-2\pi \hat{\mathbf{g}}_I^T \cdot \vec{\mathbf{h}}^{\leftrightarrow-1}) = 1 - \frac{(2\pi)^2}{2} V \sum_{\alpha,\beta} \hat{g}_{\alpha,I} \hat{g}_{\beta,I} \int_{-\infty}^{\infty} d\mathbf{s} s_{\alpha} s_{\beta} n_{\text{loc}}(\vec{\mathbf{h}} \cdot \mathbf{s}) . \quad (4.8)$$

Combining Eq. (4.8) and Eq. (4.6), we obtain,

$$1 - |z_I| = V \frac{(2\pi)^2}{2} \sum_{\alpha,\beta} \hat{g}_{\alpha,I} \hat{g}_{\beta,I} \int_{-\infty}^{\infty} d\mathbf{s} s_{\alpha} s_{\beta} n_{\text{loc}}(\vec{\mathbf{h}} \cdot \mathbf{s}) . \quad (4.9)$$

Keeping in mind that $\int_{-\infty}^{\infty} d\mathbf{s} \vec{\mathbf{h}} \cdot \mathbf{s} n_{\text{loc}}(\vec{\mathbf{h}} \cdot \mathbf{s}) = 0$, one can define the spread of the electronic distribution for the case of a general box through,

$$\langle r^2 \rangle - \langle r \rangle^2 = \langle \left(\vec{\mathbf{h}} \cdot \mathbf{s} \right)^2 \rangle = \sum_{\alpha,\beta} g_{\alpha\beta} V \int_{-\infty}^{\infty} d\mathbf{s} s_{\alpha} s_{\beta} n_{\text{loc}}(\vec{\mathbf{h}} \cdot \mathbf{s}) . \quad (4.10)$$

Here, $g_{\alpha\beta} = \sum_{\mu} \overset{\leftrightarrow T}{h}_{\alpha\mu} \overset{\leftrightarrow}{h}_{\mu\beta}$ can be thought of as a metric tensor to describe the corresponding distances in the unit cube. Eq. (4.10) shows us exactly how the length scales are built into the spread through the metric tensor. From direct comparison of Eq. (4.9) and Eq. (4.10) we see that for supercells of general symmetry we need to choose linear combinations of $\hat{g}_{\alpha,I} \hat{g}_{\beta,I}$ that reproduce the metric tensor, $g_{\alpha\beta}$. However, as stated earlier, $\hat{g}_{\alpha,I}$ are dimensionless numbers. Thus, an appropriate generalization takes the form of a sum rule,

$$g_{\alpha\beta} = \sum_I \omega_I \hat{g}_{\alpha,I} \hat{g}_{\beta,I} . \quad (4.11)$$

Here, ω_I are the ‘‘weights’’ with the appropriate dimensions to be determined in appendix 4.6.1. Thus, it should also be clear that $g_{\alpha\beta}$ will have at most six independent entries (for triclinic symmetry) and thus a maximum of six weights are needed. It is interesting to note that by multiplying Eq. (4.11) on the left and right hand sides by $\overset{\leftrightarrow-1}{\mathbf{h}}$ and using the definition of \mathbf{G}_I , one will recover the rule used by Silvestrelli [47] and by Marzari and Vanderbilt [44]. Finally, we generalize to more than one state, $|\psi\rangle \rightarrow |\psi_n\rangle$ and the desired expression for the spread, Ω in a supercell of general symmetry is,

$$\Omega = \frac{2}{(2\pi)^2} \sum_n^{\text{Nstates}} \sum_I \omega_I (1 - |z_{I,n}|) + O(2\pi \hat{\mathbf{g}}_I^T \cdot \vec{\mathbf{h}}^{\leftrightarrow-1})^2$$

$$z_{I,n} = \int_V d\mathbf{r} \exp(i\mathbf{G}_I \cdot \mathbf{r}) |\psi_n(\mathbf{r})|^2, \quad (4.12)$$

where Eq. (4.11) determine the \mathbf{G}_I .

At this point it is useful to make contact with other spread formulas that are present in the current literature. Following Resta's derivation one finds the formula [49], that in our notation reads,

$$\Omega = -\frac{1}{(2\pi)^2} \sum_n^{\text{Nstates}} \sum_I \omega_I \log |z_{I,n}|^2, \quad (4.13)$$

with $z_{I,n}$ defined as above. Eq. (4.13) is obtained by inserting Eq. (4.8) into Eq. (4.6), taking the log of the absolute value and expanding to consistent order.

Silvestrelli[47] on the other hand uses (again, in our notation),

$$\Omega = \frac{1}{(2\pi)^2} \sum_n^{\text{Nstates}} \sum_I \omega_I (1 - |z_{I,n}|^2), \quad (4.14)$$

with a similar definition for $z_{I,n}$. Obviously Eq. (4.14) is obtained from Eq. (4.13) by an expansion of the log.

At first glance, it seems confusing that there are different definitions for the spread. Admittedly, one has to keep in mind that all formulae are only valid up to the order given in Eq. (4.12). Thus, although different, they are consistent and there is no fundamental reason to choose one definition of the spread over another. Consequently, we compare in sec. 4.4 the spreads of various model systems using all different definitions.

One can also derive a general expression for the expectation value of the periodic position operator for computing the center of the localized function. Recall, that for a cubic simulation supercell the expectation value of the position operator is given as,

$$r_{\alpha,n} = -\frac{L}{2\pi} \text{Im} \log z_{\alpha,n}$$

$$z_{\alpha,n} = \int_V d\mathbf{r} \exp(i\mathbf{g}_\alpha \cdot \mathbf{r}) |\psi_n(\mathbf{r})|^2, \quad (4.15)$$

where $\hat{\mathbf{g}}_1 = (1, 0, 0)$, $\hat{\mathbf{g}}_2 = (0, 1, 0)$, and $\hat{\mathbf{g}}_3 = (0, 0, 1)$, and Im denotes the imaginary part. Again, the salient feature of Eq.(4.15) is that the expectation value of the exponential is invariant with respect to the choice of cell. Thus, a general equation for the expectation value of the position operator in supercells of arbitrary symmetry is,

$$r_{\alpha,n} = - \sum_{\beta} \frac{\overleftrightarrow{h}_{\alpha\beta}}{2\pi} \text{Im} \log z_{\alpha,n} . \quad (4.16)$$

Having established the definition of the spread formulae in the context of WF's we proceed with a detailed description of their computation.

4.3 A Generalized Localization Procedure

The mathematical problem which defines the localization procedure is to find the unitary transformation, U on the orbitals,

$$|\tilde{\psi}_n\rangle = \sum_i U_{in} |\psi_i\rangle, \quad (4.17)$$

that simultaneously minimizes the spread functional, Ω . To present a general formulation it is convenient to work with a generalized form,

$$\begin{aligned} \Omega &= \sum_n \sum_I f(|z_{I,n}|^2) \\ z_{I,n} &= \langle \psi_n | O^I | \psi_n \rangle , \end{aligned} \quad (4.18)$$

where f and O^I denote an appropriate function and operator.

If we neglect the weights and constants in favor of simplicity, we obtain the different spread functionals of the last section, defined through Eq. (4.12), Eq. (4.13) and Eq. (4.14) by setting,

$$\begin{aligned} O^I &= \exp(i\mathbf{G}_I \cdot \mathbf{r}) \\ f_1(|z_{I,n}|^2) &= \sqrt{|z_{I,n}|^2} = |z_{I,n}| , \\ f_2(|z_{I,n}|^2) &= \log(|z_{I,n}|^2) , \\ f_3(|z_{I,n}|^2) &= |z_{I,n}|^2 . \end{aligned} \quad (4.19)$$

The values of Index, I will range at most from one to six. It is important to notice that maximizing Eq. (4.18) is equivalent to minimizing the spread functional.

At this point it is useful to make the connection with methods used in traditional applications of quantum chemistry. Due to its general form, Eq. (4.18) is also suited for finite systems. The Boys-Foster and the Pipek-Mezey method [40,41], which are widely used to obtain localized molecular orbitals (MO) in quantum chemistry, can both be formulated by working with f_3 . Using MO's instead of BO's we define the operator, O^I in Eq. (4.19) for the Boys-Foster and Pipek-Mezey scheme respectively as,

$$O^I = r_I \quad \text{or} \quad O^I = \sum_{\mu \in I} \frac{1}{2} \{ |\tilde{\mu}\rangle \langle \mu| + |\mu\rangle \langle \tilde{\mu}| \} . \quad (4.20)$$

Here, r_I denotes the conventional position operator, I runs from one to the number of atoms, $\{|\mu\rangle, \mu \in I\}$ denotes the atomic basis set of atom I , and $|\tilde{\mu}\rangle = \sum_{\nu} S_{\nu\mu}^{-1} |\nu\rangle$, where S is the overlap matrix of the atomic basis set.

While the Boys-Foster method minimizes the total spread, $\Omega = \langle r^2 \rangle_n - \langle r \rangle_n^2$ (and hence maximizes $\langle r \rangle_n^2$ [40]), the Pipek-Mezey method is different and warrants some explanation. The projection operator in Eq. (4.20) used by Pipek and Mezey (PM) is closely related to the Mulliken population analysis [41]. In fact, maximizing the PM functional corresponds to a minimization of the number of atoms over which an orbital is spread. The popularity of the PM functional originates from two sources. First, it's very easy to implement and leads to a fast algorithm when combined with the linear combination of atomic orbitals method. Second, the PM functional, unlike the Boys-Foster method, preserves the σ - π separation of double bonds. This picture of double bonds is in chemistry usually preferred over the τ (banana) bond picture generated by the Boys-Foster localization procedure. As an illustration we have computed in figure 4.1 and 4.2 the maximally-localized orbitals of an ethene molecule generated by the Pipek-Mezey and the Boys-Foster localization procedure. Here, the different symmetry properties of the localization functionals lead to different types of localized orbitals.

The actual calculation of maximally localized WF's or maximally localized MO's

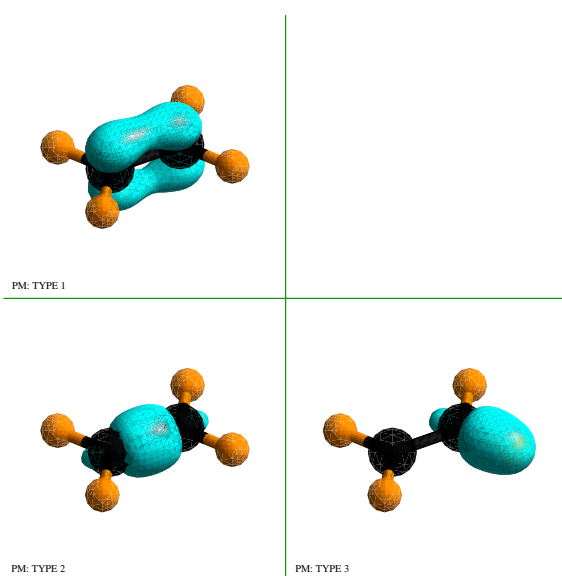


Figure 4.1: Maximally-localized orbitals of an ethene molecule generated by the Pipek-Mezey (PM) localization procedure. The PM functional preserves the σ - π separation of double bonds leading to three different types of localized orbitals.

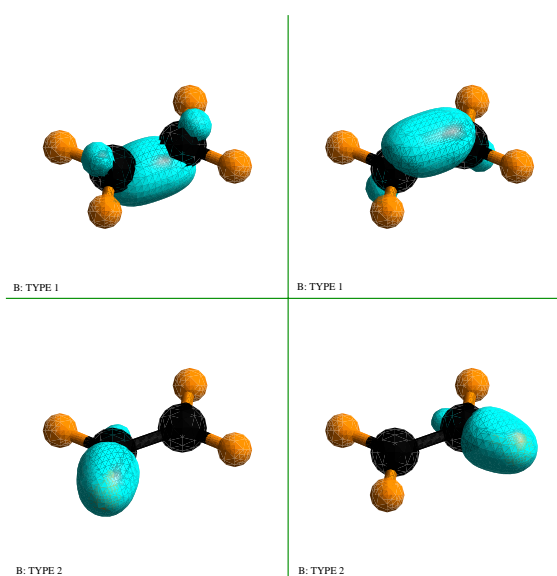


Figure 4.2: Maximally-localized orbitals of an ethene molecule generated by the Boys-Foster (BF) localization procedure. The BF functional results in a τ (banana) bond picture leading to two different types of localized orbitals.

within our localization procedure is relatively simple. First, we take the output of a conventional electronic structure calculation, (BO's in the periodic, MO's in the finite case) choose a spread functional and solve for the unitary transformation producing the orbitals that maximize Eq. (4.18). As stated earlier, we have two choices for spread functionals for finite systems and three choices of spread functionals for periodic systems. The details of the calculation are described in the following.

One must now focus on the computation of U . To ensure a maximally localized function, we would like to find an efficient solution to,

$$\frac{\partial \Omega}{\partial U_{ij}} = 0, \quad (4.21)$$

where U is considered to be real since we are working within the Γ -point approximation. There are two principal alternatives for parameterizing the unitary transformation, U , first as a direct product of elementary plane rotations, and second as the exponential of an antisymmetric matrix. The first parameterization scheme, discussed in the next subsection, amounts to the well known Jacobi optimization procedure for finding eigenvalues of general matrices. The second parameterization choice of U , used in our method, is based on the exponential alternative as investigated in sec. 4.3.2.

4.3.1 Orbital Rotations

The traditional method in quantum chemistry for computing localized MO's is the method of two-by-two orbital rotations first introduced by Edmiston and Ruedenberg [42]. The basic idea of the method is to tackle the problem of finding U by performing a sequence of consecutive two-by-two rotations among all pairs of orbitals. The elementary step consists of a plane rotation where two orbitals i and j are rotated through an angle, ϕ . To proceed we select an optimal angle to ensure that our spread functional, as defined in Eq. (4.18), is iteratively maximized. The transformed expectation values are denoted with $\tilde{z}_{I,i/j}$ and are obtained as,

$$\tilde{z}_{I,i} = \cos(\phi)z_{I,i} + \sin(\phi)z_{I,j} , \quad \tilde{z}_{I,j} = -\sin(\phi)z_{I,i} + \cos(\phi)z_{I,j} . \quad (4.22)$$

Thus, combining Eq. (4.22) with Eq. (4.18) it is straightforward to calculate the change in the functional value, $\Delta\Omega$ as a function of ϕ . The most natural way to obtain the optimal angle which maximizes the change in the functional value is to compute the derivative of $\Delta\Omega$ with respect to ϕ , set it to zero and solve for ϕ . This is precisely the way the method of orbital rotations is implemented and an explicit calculation yields,

$$\tan(4\phi) = -\frac{a}{b}, \quad (4.23)$$

$$a = \text{Re}[M_{ij}(\bar{M}_{ii} - \bar{M}_{jj})], \quad b = |M_{ij}|^2 - \frac{1}{4}|M_{ii} - M_{jj}|^2,$$

where $M_{ij} = \sum_I z_{I,ij}$ and Re denotes the real part. $\phi + n\pi/4$, $n \in \mathbb{Z}$, are the solutions of Eq. (4.23) corresponding to maxima and minima. For a maximum the condition, $\partial^2\Delta\Omega/\partial\phi^2 = 16 b \cos(4\phi) - 16 a \sin(4\phi) < 0$ has to be fulfilled.

Unfortunately there is one severe restriction. Eq. (4.23) is only valid in the case, $f_3(x) = x$, namely in connection with the Silvestrelli-Marzari-Vanderbilt, the Boys-Forster and the Pipek-Mezey functional (see sec. 4.3.2). In the other cases, $f_1(x) = \sqrt{x}$ (functional of the present work) and $f_2(x) = \log(x)$ (Resta), no analogous formula is derivable. The reason is that the explicit solution of $\partial\Delta\Omega/\partial\phi = 0$ with respect to ϕ seems not analytically tractable. Nevertheless, one can still implement the method of orbital rotations in the above cases by a numerical maximization of $\Delta\Omega$ as a function of ϕ using derivative information.

4.3.2 Exponential Representation

The ansatz, $|\tilde{\psi}_n\rangle = \sum_{\tilde{i}} U_{in} |\psi_{\tilde{i}}\rangle$, where U is an unitary matrix, leads to the transformed expectation value,

$$\tilde{z}_{I,n} = \sum_{ij} U_{in}^\dagger U_{jn} z_{I,ij}, \quad (4.24)$$

with $z_{I,ij} = \langle \psi_i | O^I | \psi_j \rangle$. As discussed above, we parameterize $U = \exp(A)$ as the exponential of an antisymmetric matrix and calculate the gradient with respect to

A. Using the chain rule the gradient splits into two pieces,

$$\frac{\partial \Omega}{\partial A_{ij}} = \sum_{st} \frac{\partial \Omega}{\partial U_{st}} \frac{\partial U_{st}}{\partial A_{ij}} = \sum_{I,n} \sum_{st} \frac{\partial f(|z_{I,n}|^2)}{\partial U_{st}} \frac{\partial U_{st}}{\partial A_{ij}}. \quad (4.25)$$

It is worth while to note that only the first piece depends on the type of spread functional and its evaluation is straightforward,

$$\begin{aligned} \frac{\partial f(|z_{I,n}|^2)}{\partial U_{st}} &= f'(|z_{I,n}|^2) \frac{\partial |z_{I,n}|^2}{\partial U_{st}} \\ &= 2 f'(|z_{I,n}|^2) \left(\sum_i U_{in} \delta_{tn} z_{I,is} \right) \left(\sum_{kl} U_{kn} U_{ln} \bar{z}_{I,kl} \right) + \text{c.c.}, \end{aligned} \quad (4.26)$$

where \bar{z} denotes the complex conjugate (**c.c.**) of z and f' is the derivative of f .

Combining Eq. (4.25) and Eq. (4.26) a general form for the gradient is obtained as,

$$\frac{\partial \Omega}{\partial A_{ij}} = \sum_{st} M_{st} \frac{\partial U_{st}}{\partial A_{ij}} = \left[M^T \frac{\partial U_{st}}{\partial A_{ij}} \right], \quad (4.27)$$

where M is defined via Eq. (4.26). However, the Pipek-Mezey functional has to be treated with special care. Since the summation index, I runs from one to the number of atoms of the system, one has to use the separability of the corresponding operator, O^I to calculate M in an efficient way.

The calculation of $\partial U_{st}/\partial A_{ij}$ is more subtle. We have to calculate the derivative of a matrix function, here the exponential function, $U = \exp(A)$ with respect to A . This can be done by writing the matrix function in an alternative way using a complex contour integral [25],

$$\begin{aligned} \frac{\partial U}{\partial A_{ij}} &= \frac{\partial \exp(A)}{\partial A_{ij}} \\ &= \frac{1}{2\pi i} \frac{\partial}{\partial A_{ij}} \oint \exp(z) (z\mathbf{1} - A)^{-1} dz \\ &= \frac{1}{2\pi i} \oint \exp(z) (z\mathbf{1} - A)^{-1} (1^{ij} - 1^{ji}) (z\mathbf{1} - A)^{-1} dz \\ &= R^\dagger \frac{1}{2\pi i} \oint \exp(z) (z\mathbf{1} - \Lambda)^{-1} R (1^{ij} - 1^{ji}) R^\dagger (z\mathbf{1} - \Lambda)^{-1} dz R. \end{aligned} \quad (4.28)$$

$\mathbf{1}$ denotes the identity matrix, $(1^{ij})_{kl} = \delta_{ki} \delta_{lj}$, R is the eigenvector matrix of A with eigenvalues λ_k and $\Lambda_{kl} = \lambda_k \delta_{kl}$. Carrying out the integration over z , one obtains,

$$\frac{1}{2\pi i} \oint \frac{\exp(z)}{(z - \lambda_k)(z - \lambda_l)} dz = \begin{cases} e^{\lambda_k}, & \lambda_k = \lambda_l, \\ \frac{e^{\lambda_k} - e^{\lambda_l}}{\lambda_k - \lambda_l}, & \lambda_k \neq \lambda_l, \end{cases} \quad (4.29)$$

Performing some simple algebraic transformations, Eq. (4.27) becomes,

$$\begin{aligned} \frac{\partial \Omega}{\partial A_{ij}} &= \text{Tr} \left[M^T R^\dagger \{C^{ij}, B\} R \right] \\ &= (R^\dagger \{RM^T R^\dagger, B\} R)_{ji} - (R^\dagger \{RM^T R^\dagger, B\} R)_{ij}, \end{aligned} \quad (4.30)$$

where the B_{kl} components are defined through Eq. (4.29), $\{C^{ij}, B\}$ denotes a component-wise matrix multiplication and $C^{ij} = R(1^{ij} - 1^{ji})R^\dagger$. The final transformation in Eq. (4.30) is verified by inserting the explicit definition of the matrix 1^{ij} .

Using the results above we have now a very useful scheme to optimize Ω by iterating the following steps:

- Start with an arbitrary matrix, *e.g.* $A = 0$.
- Diagonalize A to obtain the eigenvector matrix R and the diagonal matrix Λ with the eigenvalues as diagonal elements.
- Calculate the unitary matrix via $U = \exp(A) = R^\dagger e^\Lambda R$.
- Compute M and B defined in Eq. (4.27) respectively (4.29).
- Calculate the gradient according to Eq. (4.30).
- Update A and repeat process.

Within the above scheme the desired gradient is obtained analytically and we are able to combine our iterative localization procedure with gradient methods developed to accelerate convergence.

In addition we have calculated a simplified expression for the second derivative to be used as a preconditioner for a further speed up of the iterative localization (see Appendix 4.6.2). There we also compute the gradient at $A = 0$ which is equivalent to the gradient used by Sivestrelli et al. [46,47]. Sec. 4.4 is devoted to a detailed comparison of the methods presented in this section.

Symmetry	Functional	d_{C1}	Ω_{C1}	d_{C2}	Ω_{C2}	d_{L1}	Ω_{L1}	d_{L2}	Ω_{L2}
simple cubic	PW	0.53	0.72	0.53	0.72	0.30	0.75	0.30	0.75
	R	0.53	0.72	0.53	0.72	0.30	0.76	0.30	0.76
	SMV	0.53	0.71	0.53	0.71	0.30	0.75	0.30	0.75
orthorhombic	PW	0.53	0.72	0.53	0.72	0.30	0.75	0.30	0.75
	R	0.53	0.72	0.53	0.72	0.31	0.76	0.30	0.76
	SMV	0.54	0.71	0.53	0.71	0.30	0.75	0.30	0.75
fcc	PW	0.54	0.71	0.53	0.71	0.30	0.75	0.30	0.74
	R	0.53	0.72	0.53	0.72	0.30	0.76	0.30	0.75
	SMV	0.54	0.70	0.54	0.70	0.30	0.74	0.30	0.73
bcc	PW	0.53	0.73	0.53	0.72	0.31	0.77	0.30	0.76
	R	0.53	0.73	0.53	0.72	0.30	0.78	0.30	0.76
	SMV	0.53	0.73	0.53	0.72	0.30	0.77	0.29	0.76
hexagonal	PW	0.53	0.74	0.53	0.72	0.31	0.78	0.30	0.77
	R	0.53	0.73	0.53	0.72	0.30	0.78	0.30	0.76
	SMV	0.54	0.74	0.53	0.72	0.31	0.78	0.30	0.77
triclinic	PW	0.53	0.71	0.53	0.71	0.31	0.75	0.30	0.75
	R	0.53	0.72	0.53	0.72	0.30	0.76	0.30	0.76
	SMV	0.54	0.71	0.53	0.71	0.30	0.75	0.30	0.75

Table 4.1: Distances , with respect to the position of the oxygen ion, and spreads of the two covalent orbital WF's ($d_{C1/2}$ and $\Omega_{C1/2}$ in \AA) and the two lone-pair WF's ($d_{L1/2}$ and $\Omega_{L1/2}$ in \AA) of the isolated water molecule . Data have been computed for different supercell symmetries, using the functional of present work (PW,Eq. (4.12)), the Resta functional (R,Eq. (4.13)) and the Silvestrelli-Marzari-Vanderbilt functional (SMV,Eq. (4.14)). For details see text.

4.4 Results and Discussion

The iterative localization algorithm, presented in this work, has been implemented in the GAPW method (see section 1.3) and in the CPMD code [54] which is based on a plane wave expansion. Both programs are suited for periodic and for finite systems (see section 3).

One of the goals of our work is to make a comparison of the different spread functionals proposed in the literature, including the one derived in this work for supercells of general symmetry. In order to perform a comparison, we have chosen a simple system which has been previously studied using different spread functionals. Thus, we consider the case of one isolated water molecule periodically replicated in different supercell symmetries (simple cubic, orthorhombic, hexagonal, fcc, bcc, hexagonal and triclinic) where the same cell volume is maintained [55]. It is clear that the centroid of the localized functions (Eq. (4.16)) should remain in the same positions relative to the atoms for all supercells considered (ignoring the negligible effects of distortion from the periodic images for the cell parameters chosen). Following the work of Ref. [47] we use the general form of the functional (Eq. (4.18)), and specifically the spread functionals given by Eqs. (4.12), (4.13) and (4.14) for the iterative maximization.

Table 4.1 reports our results for the distances between the oxygen atom and the centroids of the four different WF's (calculated using Eq. (4.16)), and the spreads for the different supercell symmetries. No systematic differences in the distances are recognizable. Besides numerical noise, all considered spread functionals lead, independent from the choice of the supercell, to the same position of the centroids relative to the oxygen atom. This fact is in contrast to the calculated spread, here the distinct definitions result in slightly different values, even though one finds qualitative agreement.

After having compared our functional to previous work in literature, we examine possible ways to improve the iterative localization beyond a simple steepest descent scheme. To achieve a faster convergence the following optimization methods, in

Fun.	SD	CG	CG+PR	DIIS	OR
Formamide (9)					
BF	394	62	52	48	19
PM	187	39	36	30	24
Pentane (16)					
BF	57	26	25	24	34
PM	26	18	15	16	30
Naphtalene (24)					
BF	614	104	76	75	62
PM	172	47	39	31	58
Histidine (30)					
BF	2036	182	156	117	77
PM	365	90	70	66	66
Testosterone (58)					
BF	1444	236	202	122	152
PM	567	77	73	54	160

Table 4.2: Number of iteration steps to reach convergence for different optimization methods. Convergence is assumed if the relative value of the spread functional is decreasing by less than 10^{-8} for successive iteration steps. The considered methods are a steepest descent (SD) with line search procedure, the Polak-Ribière conjugate gradient method (CG), the Polak-Ribière conjugate gradient method with a preconditioner (CG+PR) the direct inversion in iterative subspace method (DIIS) with a preconditioner and the orbital rotation method (OR). One step of OR is about three times more expensive in CPU time compared to the other methods. BF denotes the Boys-Forster, PM the Pipek-Mezey functional. The number of double occupied orbitals of the respective system is added in parenthesis.

Fun.	SD	CG	CG+PR	DIIS	OR
Si_8 (16)					
PW	28	19	19	19	25
R	29	19	19	19	37
SMV	28	19	18	18	32
$(\text{H}_2\text{O})_8$ (32)					
PW	323	60	44	42	67
R	740	89	72	61	92
SMV	248	71	61	53	87
C_{32} (64)					
PW	216	52	38	32	109
R	301	73	66	52	141
SMV	197	57	39	35	184
$(\text{SiO}_2)_{16}$ (128)					
PW	5199	479	253	178	261
R	8634	2932	478	394	307
SMV	3348	415	210	172	375

Table 4.3: Number of iteration steps to reach convergence for different optimization methods. Convergence is assumed if the relative value of the localization functional is decreasing by less than 10^{-8} for successive iteration steps. The considered methods are a steepest descent (SD) with line search procedure, the Polak-Ribière conjugate gradient method (CG), the Polak-Ribière conjugate gradient method with a preconditioner (CG+PR), the direct inversion in iterative subspace method (DIIS) with a preconditioner and the method of orbital rotations (OR). One step of OR is about three times more expensive in CPU time compared to the other methods. PW denotes the functional of present work (Eq. (4.12)), R the Resta functional (Eq. (4.13)) and SMV the Silvestrelli-Marzari-Vanderbilt functional (Eq. (4.14)). The number of double occupied orbitals of the respective system is added in parenthesis.

connection with the scheme described in sec. 4.3 are implemented:

- A steepest descent (SD) with line search procedure performed in the following way. After bracketing the maximum, a parabolic fit through three points is carried out. The SD serves as a benchmark, since one can combine the method also with the approximate gradient (Eq. (4.33)).
- The Polak-Ribière conjugate gradient method (CG) [50].
- The Polak-Ribière conjugate gradient method in combination with a preconditioner (CG+PR). Close to the maximum we replace the pure gradient \mathbf{g} by $H^{-1}\mathbf{g}$ where H^{-1} denotes the inverse of the approximate diagonal Hessian matrix calculated via Eq. (4.33).
- The direct inversion in the iterative subspace method (DIIS) [51] in combination with a preconditioner. In addition the diagonal approximation of the Hessian matrix is improved by making use of the information gained by calculating exact first derivatives at a series of iteration points. In our implementation the Limited Memory - BFGS method [56] is used which directly updates the inverse of the Hessian matrix.
- Method of orbital rotations (OR) as discussed in sec. 4.3.1.

To study the effect of the different optimization methods on the convergence of our scheme we consider four periodic systems, using the functional proposed in this work (PW, Eq. (4.12)), the Resta functional (R, Eq. (4.13)) and the Silvestrelli-Marzari-Vanderbilt functional (SMV, Eq. (4.14)) and five finite systems using the Pipek-Mezey (PM) and the Boys-Foster (BF) functional. As finite systems we have chosen the formaldehyde derivate formamide CHONH_2 (9), the alkane pentane C_5H_{12} (16), the aromate naphthalene C_{10}H_8 (24), the amino acid histidine $\text{C}_6\text{H}_9\text{N}_3\text{O}_2$ (30) and the steroid testosterone $\text{C}_{19}\text{H}_{28}\text{O}_2$ (58). As periodic benchmark systems serve bulk silicon Si_8 (16), bulk water $(\text{H}_2\text{O})_8$ (32), diamond C_{32} (64) and β -Cristobalite $(\text{SiO}_2)_{16}$ (128). The number of doubly occupied orbitals is added in parenthesis.

The criterion used to classify the properties of a given method is as follows, we have always maintain the same accuracy in convergence and only evaluate the number of steps required to reach this.

The results of our calculation are summarized in table 4.2 for the finite and in table 4.3 for the periodic systems. Obviously there is a wide range in convergence behavior depending on the system size but also on the specific chemical properties. Small, fully saturated systems are easy to localize ($\text{Si}_8, \text{C}_5\text{H}_{12}$) and in those cases a steepest descent scheme is sufficient. However, this is no longer true if one switches to more complicated systems including electron lone pairs, double bonds or aromatic rings in connection with an increasing number of orbitals. In these cases, a considerable gain is obtained using high level optimization methods (CG, CG+PR, DIIS). For the biggest system we have studied, more than an order of magnitude increase in convergence speed is observed compared to the steepest descent with line search procedure. On the other hand the OR scheme, in spite of its simplicity, is remarkably efficient [57].

It is not surprising that the more sophisticated iterative schemes are more efficient than the steepest descent procedure. However, for cases where one would like the WF's not only for the final configuration, but for evenly sampled times along the trajectory, an efficient scheme is imperative.

One comment is in order. We find the best convergence using the preconditioned DIIS scheme. Unfortunately, the method converges to the 'closest' stationary point which in many cases is not the global maxima. Thus, it seems more advantageous to work in a practical implementation with the CG+PR method.

Having finished the methodical part, we look at some applications. The present methodology has already been employed in different areas of electronic structure calculations, e.g. in the context of linear scaling methods [58], in a new ab-initio approach for NMR chemical shifts in periodic systems [59] and as an analysis tool to study microsolvation and chemical reactivity of sodium and water clusters [60] or bond breaking by mechanical stress [61]. As a further example how WF's can serve

Molecule	Type	Spread	Type	Spread	Type	Spread
Graphite	GT 1	0.77	GT 2	1.14	—	—
Nanotube	NT 1A	0.80	—	—	NT 3	1.65
	NT 1B	0.75				
Fullerene	FT 1	0.82	FT 2A	1.29	—	—
			FT 2B	1.05		

Table 4.4: Spreads (in \AA) of the different types of maximally-localized Wannier Functions in graphite, in a $(10,0)$ zig-zag nanotube and in fullerene C_{60} . Data have been computed using the Silvestrelli-Marzari-Vanderbilt functional. For details see text.

to analyze chemical bonding, we turn to the field of carbon nanotubes.

Carbon nanotubes have excited a considerable interest in the condensed-matter and materials research communities in the last few years, and much experimental and theoretical work has been devoted to them as prototype of one-dimensional ordered systems with promising technological applications [62–64]. Classifications of the tubes as metals or semiconductors were given on the basis of how the underlying graphite band structure is folded [65,66]. Recently, nanotubes with very small radii were theoretically studied and it was found that the electronic properties of small tubes are significantly altered from those obtained by purely folding a graphite sheet [67]. This is a consequence of strong σ^* - π^* hybridization due to the effect of the large curvature. To confirm, at least qualitatively, the aforementioned curvature effect on nanotubes with small radii, we have depicted in figure 4.3 and 4.4 the WF's in graphite and in a $(10,0)$ zig-zag nanotube. The WF's are taken from pseudopotential LDA calculations using the GAPW method.

In graphite, two different types of localized orbitals were found, a σ -type orbital (GT 1) and a τ -type orbital (GT 2). This banana bond picture generated by the Silvestrelli-Marzari-Vanderbilt functional corresponds to the more familiar double

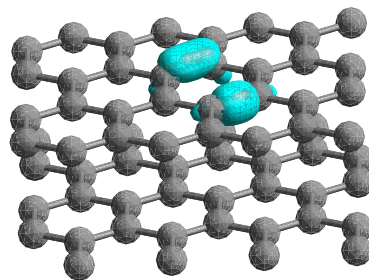


Figure 4.3: Maximally-localized Wannier Functions in graphite. Two different types of localized orbitals were found: A σ -type orbital (GT 1) and a τ -type (banana-type) orbital (GT 2). The spreads are given in table 4.4

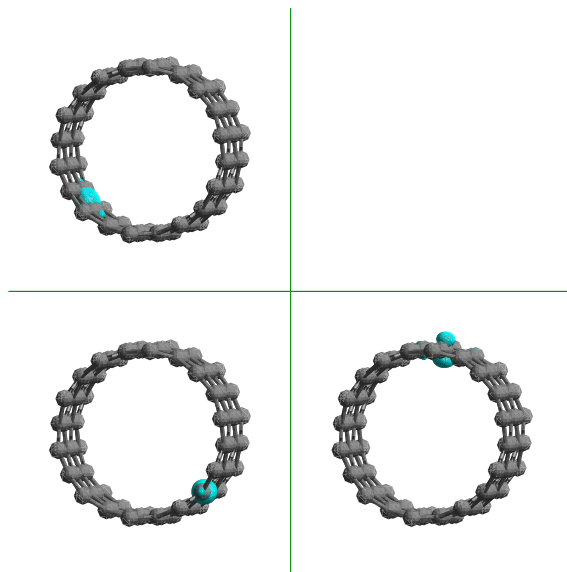


Figure 4.4: Maximally-localized Wannier Functions in a $(10, 0)$ zig-zag nanotube. Three different types of localized orbitals were found: A non-coaxial σ -type orbital (upper left panel, NT 1A), a coaxial σ -type orbital (lower left panel, NT 1B) and a third type significantly altered from the τ -type (banana-type) orbital found in figure 4.3 (lower right panel, NT 3). The spreads are given in table 4.4

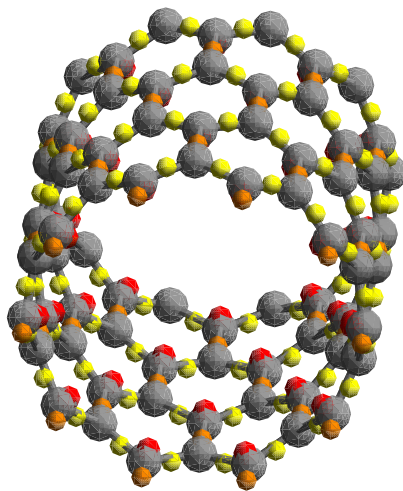


Figure 4.5: Centroids of the maximally-localized Wannier Functions depicted in figure 4.4. Grey denotes carbon, yellow denotes centroids of type NT 1A, orange denotes centroids of type ZT 1B and red denotes centroids of type NT 3. From the red centroids one can see that the localized orbitals of type NT 3 are preferably located inside the tube.

bond picture of graphite (compare figure 4.1 and 4.2). Namely, a degenerated σ -system with conjugated double bonds.

Let us now assume that the (10,0) zig-zag nanotube is generated simply by rolling the graphite sheet segment. From the symmetry of the tube, four different types of localized orbitals are expected. A non-coaxial σ -type orbital, a coaxial σ -type orbital, a τ -type orbital preferably located inside the tube and a τ -type orbital preferably located outside the tube. If we compare our expectation with figure 4.4, we can actually identify a non-coaxial (NT 1A) and a coaxial (NT 1B) σ -type orbital having different spreads (table 4.4). Obviously, the curvature corresponding to the non-coaxial direction leads to a delocalization of the NT 1A orbital. Nevertheless, since NT 1A, NT 1B and GT 1 have comparable spreads and shapes, the curvature only weakly affects the σ -type orbitals. Contrary to what we expected, figure 4.4 shows no τ -type orbitals, but rather a third type (NT 3) significantly altered from the GT 2 τ -type orbital found in figure 4.3. Two main features are apparent. Most of the NT 3 orbital is located inside the tube (figure 4.5) and the orbital is much more delocalized than a GT 2 τ -type orbital (table 4.4). Thus, by plotting the WF's we are able to visualize how the electronic properties of a small nanotube are modified from those generated simply by folding a graphite sheet. A careful analysis shows that these modifications are mainly due to a strong, curvature-induced σ^* - π^* hybridization [67].

In addition we have computed the WF's in fullerene C_{60} (figure 4.6). Three different types of localized orbitals were found. A σ -type orbital (FT 1), and two τ -type orbital preferably located inside (FT 2A) respectively outside (FT 2B) the fullerene. Although the curvature leads to a slight delocalization of the FT 1 orbital and to a splitting of the τ -type orbital into a less localized endohedral FT 2A orbital and a more localized exohedral FT 2B orbital (table 4.4), the WF's compare favorably with those found in graphite. The result suggests that the chemical behavior of fullerene C_{60} resembles graphite, but with the peculiarity that the spherical structure allows exohedral and endohedral chemistries to be distinguished.

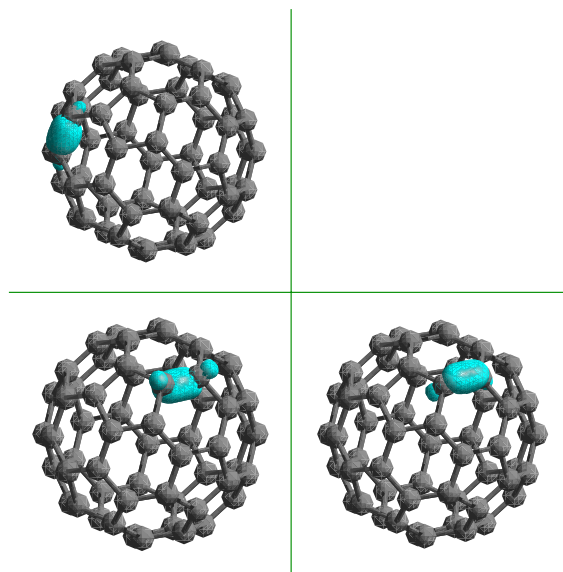


Figure 4.6: Maximally-localized Wannier Functions in fullerene C_{60} . Three different types of localized orbitals were found: A σ -type orbital (upper left panel, FT 1), a τ -type (banana-type) orbital preferably located *inside* the fullerene (FT 2A) and a τ -type orbital preferably located *outside* the fullerene (FT 2B). The spreads are given in table 4.4

4.5 Conclusions

We have generalized the work of Resta to three dimensions to derive a spread functional as a starting point to calculate maximally-localized Wannier functions in simulation cells of general symmetry suitable for periodic systems in the Γ -point approximation. Thus, one can easily perform a calculation in a supercell with arbitrary symmetry with knowledge of only the cell parameters.

The functional turns out to be equivalent to others discussed in the literature, it differs, besides its more general form, only in the higher-order terms from either the Silvestrelli-Marzari-Vanderbilt or Resta proposals.

In addition, we have derived an iterative scheme to obtain maximally-localized orbitals. The method is also suited for finite systems and combination with very efficient optimization procedures is possible.

To compare the efficiency of our approach with the standard method in quantum chemistry proposed by Edmiston and Ruedenberg [42], we generalize the method of orbital rotations to a wider class of spread functionals.

Test calculation show that with increasing system size and chemical complexity the use of high level optimization methods lead to a considerably faster convergence, for the biggest systems the performance is improved by more than a factor of ten. On the other hand, the simple orbital rotation scheme is surprisingly efficient. With respect to an practical implementation both methods have advantages, the orbital rotation method is easier to implement, nevertheless the efficiency is fairly good, our method is more general by being, at least for larger systems, clearly faster.

As an application we have computed the maximally-localized Wannier functions of the (10,0) zig-zag nanotube. By comparing the results with the Wannier functions of graphite and of the fullerene C_{60} , we are able to provide an intuitive picture how the electronic properties of nanotubes with very small radii are significantly altered from a graphite sheet by curvature effects.

4.6 Appendix

4.6.1 Determination of the Weights

In this appendix we determine the weights, ω_I as defined in the sum rule Eq. (4.11) for supercells of general symmetry. Recall that the metric, $\overleftrightarrow{\mathbf{g}}$ will contain at most six independent entries as defined by the case of least symmetry, triclinic. Thus, Eq. (4.11) is a linear set of six equations with six unknowns. We have freedom to choose the six Miller indices, $\hat{\mathbf{g}}^I$ of which we are to take the linear combinations of. For computational convenience of computing z_i we choose the first six indices that take you from one to the next point in the Brillouin zone. Namely, $\hat{\mathbf{g}}^1 = (1, 0, 0)$, $\hat{\mathbf{g}}^2 = (0, 1, 0)$, $\hat{\mathbf{g}}^3 = (0, 0, 1)$, $\hat{\mathbf{g}}^4 = (1, 1, 0)$, $\hat{\mathbf{g}}^5 = (1, 0, 1)$, $\hat{\mathbf{g}}^6 = (0, 1, 1)$. With this choice of $\hat{\mathbf{g}}^i$ the explicit system of equations based on Eq. (4.11) takes the following simple form,

$$\begin{pmatrix} 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \\ \omega_4 \\ \omega_5 \\ \omega_6 \end{pmatrix} = \begin{pmatrix} g_{11} \\ g_{12} \\ g_{13} \\ g_{22} \\ g_{23} \\ g_{33} \end{pmatrix} \quad (4.31)$$

Thus, the solution to Eq. (4.31) yields the following set of general weights,

$$\begin{aligned} \omega_1 &= g_{11} - g_{12} - g_{13} \\ \omega_2 &= g_{22} - g_{12} - g_{23} \\ \omega_3 &= g_{33} - g_{13} - g_{23} \\ \omega_4 &= g_{12} \\ \omega_5 &= g_{13} \\ \omega_6 &= g_{23} \end{aligned} \quad (4.32)$$

Eq. (4.32) indeed reduces to the specific cases computed in Ref.[47]. However, here, the case for triclinic symmetry is also included. Thus, with knowledge of the cell

parameters, in conjunction with Eq. (4.12) allows one to compute the maximally localized WF.

4.6.2 Approximate Expressions for the Gradient and the Second Derivative

We can also calculate the second derivative analytically by the same methods as described in sec. 4.3. However, for our purposes an approximate solution is sufficient. The simplest approximation is to neglect all off-diagonal elements, thus allowing for an easy matrix inversion as needed in optimization schemes. This approximation is justified near the maximum where the Hessian matrix is diagonal dominant.

Further we can calculate the diagonal elements in an approximate way too. Using a power series expansion of the exponential function and exchanging limit and summation, it is easy to show that $\lim_{A \rightarrow 0} \exp(\pm A) = 1$, $\lim_{A \rightarrow 0} \partial \exp(\pm A) / \partial A_{ij} = \pm(1^{ij} - 1^{ji})$ and $\lim_{A \rightarrow 0} \partial^2 \exp(\pm A) / \partial A_{ij}^2 = (1^{ij} - 1^{ji})^2 = -(1^{ii} + 1^{jj})$. With this we get,

$$\begin{aligned} \lim_{A \rightarrow 0} \frac{\partial^2 \Omega}{\partial A_{ij}^2} = & -4 \sum_I \left(|z_{I,i}|^2 f'(|z_{I,i}|^2) + |z_{I,j}|^2 f'(|z_{I,j}|^2) \right) \\ & + 16 \sum_I \left(\text{Re}[z_{I,ij} \bar{z}_{I,j}]^2 f''(|z_{I,j}|^2) + \text{Re}[z_{I,ij} \bar{z}_{I,i}]^2 f''(|z_{I,i}|^2) \right), \end{aligned} \quad (4.33)$$

ending up with an expression for the second derivative in the limiting case, $A \rightarrow 0$. In order to use Eq. (4.33), we first perform an optimization until close to the maximum, then we update the expectation values via Eq. (4.24) and restart the iteration with $A = 0$. At that point Eq. (4.33) is a reasonable estimate of the diagonal elements of the second derivative and we can use the inverse of the approximate diagonal matrix as a preconditioner.

We can also calculate the gradient in the same limit,

$$\lim_{A \rightarrow 0} \frac{\partial \Omega}{\partial A_{ij}} = 4 \sum_I \left(\text{Re}[z_{I,ij} \bar{z}_{I,j}] f'(|z_{I,j}|^2) - \text{Re}[z_{I,ij} \bar{z}_{I,i}] f'(|z_{I,i}|^2) \right). \quad (4.34)$$

Eq. (4.34) amounts to the gradient used by Sivestrelli et al. [46,47] for general functionals. In each iteration step one has to update the expectation values via Eq. (4.24)

and to restart with $A = 0$. This amounts to a redefinition of the optimization parameters, making the use of global schemes impossible. In addition, to update the expectation values means to diagonalize A in order to obtain the unitary matrix, $U = \exp(A)$. Since the diagonalization is also the most time consuming step in the calculation of the full gradient, the use of Eq. (4.30) is in any case advantageous.

5 Polarized Atomic Orbitals for Linear Scaling Methods

5.1 Introduction

The GAPW method has been introduced in detail in section 1.3. The essential feature was to achieve a computational cost for the calculation of the Kohn-Sham matrix that scales with the system size N as $O(N \log N)$. Therefore, the rate-determining step for GAPW calculations in the large molecule limit is the density update procedure, which is conventionally done as a diagonalization. Although the diagonalization step has a very small prefactor, the procedure dominates the overall time scaling for large systems due to its cubic dependence on system size. Therefore, interest in the search for more efficient update methods has been strongly increasing.

Recently, Lee and Head-Gordon introduced the polarized atomic orbital (PAO) method [70] to construct minimal basis sets optimized in the molecular environment. A flexible PAO basis set with a dimension typically equal to the size of a minimal basis, is formed from atom-centered linear combinations of a larger set of atomic orbitals. While atom-optimized minimal basis sets, e.g. of the STO-nG type, are often performing poorly in practical calculations, the PAO's derive their flexibility from the fact that they can adapt to the molecular environment by the admixture of higher angular momentum functions.

From a computational point of view, PAO's have the attractive feature of greatly reducing the number of independent variables to be determined during a density update. This offers the possibility for studying large systems currently intractable with larger than minimal basis sets, while retaining a considerable amount of flexibility in the basis set.

Besides, being considerably less memory and CPU intensive, the PAO method is also potentially useful in combination with linear scaling density update methods. These $O(N)$ methods, are based either on localized orbitals [71,72], or a sparse

density matrix without any explicit construction of orbitals [73–75]. Although many algorithms have been proposed, the use of linear scaling density update methods within HF or DFT is not yet widespread. One of the major drawbacks of linear scaling approaches are numerical instabilities resulting from ill-conditioned overlap matrices encountered when large nonorthogonal basis sets are used. PAO’s provides a possible solution to this problem. We find that the condition number of the PAO overlap matrix is independent from the condition number of the underlying extended basis set and consequently no numerical instabilities are encountered.

The outline of the remainder of this paper is as follows. In the first section we propose a modified version of the PAO method of Lee and Head-Gordon. After deriving all expressions required to implement this approach, we compare its accuracy and performance with extended basis set calculations. Next, the condition number of the PAO overlap matrix is examined to validate the usefulness of a combined PAO-O(N) method for large nonorthogonal basis sets. Finally, we use the Chebyshev polynomial expansion method [73], the conjugate gradient density matrix search [74] and the canonical purification of the density matrix [75] to implement the combined method and a series of benchmark calculations is presented.

5.2 Polarized Atomic Orbital Method

Polarized atomic orbitals (PAO’s) are a small set of basis functions variationally optimized during an iterative SCF calculation. The PAO basis is defined as a linear transformation of an underlying extended basis set. The transformation is strictly limited to atom-centered linear combinations of basis functions. In other words, only basis functions from the same atom are allowed to mix,

$$|\tilde{\phi}_i\rangle = \sum_j B_{ji}|\phi_j\rangle, \quad (5.1)$$

where i and j belong to the same atom. Here, $|\phi\rangle$ denotes an atomic orbital basis functions, B is the transformation from the extended to the PAO basis set and the tilde distinguishes throughout this paper quantities related to the PAO basis

set. As mentioned above, B is a strictly atom-centered block diagonal matrix. The complement of the PAO basis set within the extended basis set will be called the excluded functions.

Using Eq. (5.1), it is straightforward to compute the one-particle density matrix in the extended basis as a function of the PAO density matrix and the transformation matrix B ,

$$P = B\tilde{P}B^T . \quad (5.2)$$

We proceed by writing the transformation matrix as a product of an atom-centered orthogonalization term and an atom-centered unitary transformation,

$$B_{ij} = \sum_k N_{ik}U_{kj} = \sum_k T_{ik}^{-\frac{1}{2}}U_{kj} , \quad (5.3)$$

where T is the atom-centered overlap matrix of the atomic orbitals.

Up to now we have just established the fundamental definitions of the PAO method given by Lee and Head-Gordon [70]. Next, to motivate our modified approach, we briefly review their scheme to solve for the PAO's.

The unitary matrix U is parameterized as the product of a series of two-by-two rotation matrices, each of which is characterized by a single rotation angle that mixes a PAO with an excluded function centered at the same atom. An energy functional is chosen,

$$\Omega = \text{Tr} [(3PSP - 2PSPSP)H] , \quad (5.4)$$

identical to that used in the conjugate gradient density matrix search for non-orthogonal basis sets [76]. H is the Hamiltonian matrix and S is the overlap matrix. Inserting Eq. (5.2) into Eq. (5.4) the energy functional becomes a matrix function of the PAO density matrix and the transformation matrix, i.e. $\Omega = f(B, \tilde{P})$. Minimizing Ω has to be done with respect to the degrees of freedom that define \tilde{P} and B , namely the matrix elements \tilde{P}_{ij} and the rotation angles of the two-by-two rotation matrices. During the minimization process a Newton-Raphson quadratically convergent scheme is used. The basics of this scheme are:

- The PAO density matrix and the transformation matrix are updated simultaneously by minimizing an unconstrained energy functional.
- All equations are defined and manipulated using a non-orthogonal basis. Therefore, the authors employ tensor methods for dealing with minimization problems involving non-orthogonal basis sets.
- The use of a Newton-Raphson quadratically convergent scheme requires the calculation of the second derivative matrix with respect to the degrees of freedom.

In the following we propose an alternative to this scheme that builds on existing methods for solving the Hartree–Fock or Kohn–Sham equations.

- We split the optimization process into two parts. Optimization of the transformation matrix is, at each step, accompanied by an optimization of the density matrix (DM) in the current PAO basis. The calculation of the density matrix is performed in the PAO basis and can be done either by conventional methods (e.g. diagonalization) or linear scaling methods mentioned in section 4.1. Thus, this scheme allows for great flexibility and requires only minor changes in existing implementations.
- We avoid a minimization process involving non-orthogonal basis sets by the following procedure. During the MO/DM update we transform both the MO's/DM and the Hamiltonian matrix into an orthonormal basis. Consequently, no tensor manipulation is employed throughout the paper.
- The transformation matrix is updated by minimizing a constraint energy functional. The use of a conjugate gradient scheme avoids the cumbersome manipulation of second derivative information as needed in a Newton-Raphson scheme.

Having established our modifications we proceed with a detailed description of the computation based on Eqs. (5.1-5.3). Starting point is a parameterization of the

block diagonals of the unitary matrix, $U_I = \exp(A_I)$, as the exponential of an antisymmetric matrix. There is no loss in generality in parameterizing A_I as done in Eq. (5.5),

$$A_I = \begin{pmatrix} 0 & X_I \\ -X_I^T & 0 \end{pmatrix}, \quad (5.5)$$

where X_I is an $n_I \times (m_I - n_I)$ matrix. Here n_I denotes the number of PAO basis functions and $(m_I - n_I)$ is the number of excluded basis functions centered at atom I . The special form of A_I ensures that the energy functional defined in Eq. (5.6) is invariant with respect to unitary rotations within the PAO and the excluded basis set.

The density matrix in the PAO basis is calculated by minimizing the energy functional Ω given in matrix representation

$$\Omega = \text{Tr}[\tilde{P}B^T H B] + \text{Tr}[\tilde{\Lambda}(\tilde{P} - \tilde{P}B^T S B \tilde{P})], \quad (5.6)$$

where $\tilde{\Lambda}$ is the Lagrangian multiplier matrix corresponding to the idempotency constraint.

Prior to an PAO calculation, an initial guess for the transformation matrix has to be made. If no information from prior calculations is available, we obtain a transformation matrix by diagonalizing the atom-centered block-diagonals of H .

Given a transformation matrix B , the next step is to compute a variational density matrix, either from converged MO's or by solving directly for the PAO density matrix without any explicit construction of orbitals. If we transform the Hamiltonian matrix and the overlap matrix into the PAO basis, $\tilde{H} = B^T H B$ and $\tilde{S} = B^T S B$, Eq. (5.6) can be rewritten and we have to minimize,

$$\Omega = \text{Tr}[\tilde{P}\tilde{H}] + \text{Tr}[\tilde{\Lambda}(\tilde{P} - \tilde{P}\tilde{S}\tilde{P})], \quad (5.7)$$

which leads to a generalized eigenvalue problem, $\tilde{H}\tilde{C} = \tilde{S}\tilde{C}\tilde{\Lambda}$. The alternative way, to solve directly for the PAO density matrix using linear scaling methods, is discussed in section 5.3. It is important to notice that both procedures are only in the minimal basis.

In the optimization cycle for the PAO basis the derivative of Ω , as defined in Eq. (5.6), with respect to the degrees of freedom of B , namely the non-zero matrix elements of $A_{I,ij}$ are needed. In the following we drop the additional atom index I . Using the chain rule the gradient splits into two pieces,

$$\frac{\partial \Omega}{\partial A_{ij}} = \sum_{st} \frac{\partial \Omega}{\partial U_{st}} \frac{\partial U_{st}}{\partial A_{ij}} . \quad (5.8)$$

Inserting $B = NU$ into Eq. (5.6), the evaluation of the first part leads to,

$$\frac{\partial \Omega}{\partial U_{st}} = 2 [(N^T H B \tilde{P})_{st} - (N^T S B \tilde{P} \tilde{H} \tilde{P})_{st}] , \quad (5.9)$$

where the Lagrangian multiplier matrix has been eliminated from Eq. (5.9) by using the convergence of the PAO density matrix and the idem-potency constraint. The convergence of the PAO density matrix $\frac{\partial \Omega}{\partial P} = 0$ eliminates terms from the unknown implicit dependence of the PAO density matrix with respect to the unitary transformation, which would otherwise enter Eq. (5.9) through $\frac{\partial \Omega}{\partial P} \frac{\partial \tilde{P}}{\partial U}$.

The calculation of $\partial U_{st} / \partial A_{ij}$ is less straightforward. We have to calculate the derivative of a matrix function $U = \exp(A)$ with respect to A . This can be done by writing the matrix function using a complex contour integral [25],

$$\begin{aligned} \frac{\partial U}{\partial A_{ij}} &= \frac{\partial \exp(A)}{\partial A_{ij}} \\ &= \frac{1}{2\pi i} \frac{\partial}{\partial A_{ij}} \oint \exp(z) (z \underline{1} - A)^{-1} dz \\ &= \frac{1}{2\pi i} \oint \exp(z) (z \underline{1} - A)^{-1} (1^{ij} - 1^{ji}) (z \underline{1} - A)^{-1} dz \\ &= R^\dagger \frac{1}{2\pi i} \oint \exp(z) (z \underline{1} - \Lambda)^{-1} R (1^{ij} - 1^{ji}) R^\dagger (z \underline{1} - \Lambda)^{-1} dz R . \end{aligned} \quad (5.10)$$

$\underline{1}$ denotes the identity matrix, $(1^{ij})_{kl} = \delta_{ki} \delta_{lj}$, R is the eigenvector matrix of A with eigenvalues λ_k and $\Lambda_{kl} = \lambda_k \delta_{kl}$. Carrying out the integration over z , one obtains,

$$\frac{1}{2\pi i} \oint \frac{\exp(z)}{(z - \lambda_k)(z - \lambda_l)} dz = \begin{cases} e^{\lambda_k}, & \lambda_k = \lambda_l . \\ \frac{e^{\lambda_k} - e^{\lambda_l}}{\lambda_k - \lambda_l}, & \lambda_k \neq \lambda_l . \end{cases} \quad (5.11)$$

Combining Eqs. (5.9-5.11) and performing some algebraic transformations, Eq. (5.8) becomes,

$$\frac{\partial \Omega}{\partial A_{ij}} = (R^\dagger \{ R M^T R^\dagger, D \} R)_{ji} - (R^\dagger \{ R M^T R^\dagger, D \} R)_{ij} . \quad (5.12)$$

where D_{kl} is the matrix defined by Eq. (5.11), $M = \partial\Omega/\partial U$ is computed from Eq. (5.9) and $\{X, Y\}$ denotes a component-wise matrix multiplication between X and Y .

Using the results above we have now a scheme to solve for the PAO's by iterating the following steps:

1. Choose an initial guess for B .
2. Compute the Hamiltonian matrix and the overlap matrix in the minimal basis, $\tilde{H} = B^T H B$ and $\tilde{S} = B^T S B$.
3. Determine a variational density matrix, i.e. $\frac{\partial\Omega}{\partial P} = 0$, by minimizing Eq. (5.7) by either one of the following methods
 - (a) Solve $\tilde{H}\tilde{C} = \tilde{S}\tilde{C}\tilde{\Lambda}$ and compute $\tilde{P} = \tilde{C}\tilde{C}^T$.
 - (b) Use a linear scaling density update method as described in section 5.3.
4. Calculate the gradient, $\frac{\partial\Omega}{\partial A_I}$, according to Eq. 5.12.
5. Update A_I , by e.g. a conjugate gradient method, and compute the transformation matrix via, $B_I = N_I U_I = N_I \exp(A_I)$.
6. If convergence is achieved compute the density matrix in the extended basis, $P = B\tilde{P}B^T$, otherwise continue with step 2.

The scheme described above contains the solution of a Roothaan-type equation in step 3. However, this has to be done only in the PAO basis, reducing the work considerably both in a diagonalization based method as also in linear scaling methods.

As an example we have plotted in figure 5.1 the PAO's of water centered at the oxygen atom. While the p-type PAO basis functions appear rather unaltered, the s-type PAO basis function is strongly polarized towards the hydrogen atoms.

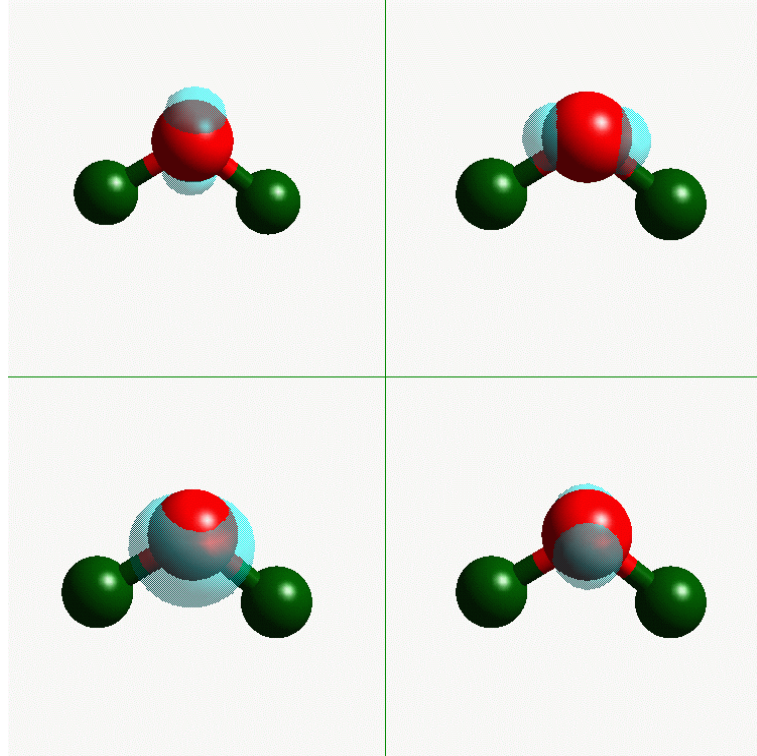


Figure 5.1: Polarized atomic orbitals of water.

5.3 Combination with Linear Scaling Methods

5.3.1 Sparsity of the Density Matrix

$O(N)$ algorithms used in this study are all based on the density matrix. The methods alone do not yield linear scaling. However, in large molecules we can take advantage of matrix sparsity, i.e., the fact that many elements of a matrix are zero or below a certain threshold. For sparse systems, where the number of significant elements scales linearly with the size of the system, the cost of all sparse matrix operations will scale linearly with the size of the system.

The sparsity of the density matrix, depends on several factors. We can write the density matrix as,

$$P_{ij} = \sum_{kl} S_{ik}^{-1} Q_{kl} S_{lj}^{-1} , \quad (5.13)$$

$$Q_{kl} = \iint dr dr' \phi_k(r) P(r, r') \phi_l(r') , \quad S_{ij} = \int dr \phi_i(r) \phi_j(r) ,$$

where $P(r, r')$ denotes the one-particle density operator and Q_{kl} is its matrix representation with respect to the atomic basis functions. The appearance of the inverse of the overlap matrix in Eq. (5.13) is due to the fact that we use nonorthogonal basis sets. From Eq. (5.13) it is clear that the sparsity of P depends on the decay properties of $P(r, r')$ as well as on the sparsity of S^{-1} .

Theoretical models of periodic solids suggest that the locality of $P(r, r')$ is related to the band gap $\Delta\epsilon$. For the case of an insulator, the one-particle density operator decays asymptotically as an exponential [77–79],

$$P(r, r') \approx \exp\left(-\sqrt{\Delta\epsilon}|r - r'|\right) . \quad (5.14)$$

While the decay behavior of $P(r, r')$ is a fundamental property of the respective quantum mechanical system, this is certainly not true for the sparsity of S^{-1} . The determining quantities are kind and size of the basis set we have chosen for our calculation. In the following we will concentrate on Gaussian-type basis functions, since they are almost exclusively used in ab-initio quantum chemistry.

Extended basis sets are often needed in DFT calculations to achieve converged results. These extended basis sets, especially if they include diffuse functions may often lead to near linear dependencies. Procedures used to eliminate the dependencies can not be applied in the linear scaling context as they destroy the local character of the basis. For a more systematic treatment of the consequences of extended basis sets, it is useful to consider the following definition. A well conditioned overlap matrix is one containing no linear or near linear dependencies. Its condition number, the ratio of its largest and smallest eigenvalue, is small. Using this definition, it can be shown [80], that a localized and well conditioned overlap matrix leads to a similarly localized inverse and finally via Eq. (5.13) to a localized density matrix. Unfortunately the reverse conclusion is also true, ill conditioned overlap matrices caused by large nonorthogonal basis sets destroy the sparsity of the density matrix. Since this is the crucial point determining whether a $O(N)$ method is faster than traditional electronic structure calculations, linear scaling methods are not used when large nonorthogonal basis sets are required.

It is for this type of calculations that the PAO method provides a possible solution. Recall from the last section, that the density update procedure is in terms of the PAO minimal basis. Since the size of the PAO basis is fixed, independent from the size of the underlying extended basis set, we may hope that the corresponding PAO overlap matrix is stable with respect to an increase of the extended basis set size. To investigate this supposition we have set up four test systems, the polyglycine chain 10-glycine, the water cluster $(\text{H}_2\text{O})_{30}$, the fullerene C_{60} and the steroid testosterone $\text{C}_{19}\text{H}_{28}\text{O}_2$. Gaussian-type valence basis sets are applied and increased in size from a minimal basis (MIN), to a double-zeta basis (DZ), a double-zeta basis with one set of polarization function (DZSP) and a triple-zeta basis with two sets of polarization functions (TZDP). All calculations were performed with a DFT program using pseudopotentials to eliminate the core electrons from the calculation. More computational details can be found in section 5.4. Table 5.1 shows the logarithm of the condition number of the PAO overlap matrix as a function of the basis set size. As a reference, we have also computed the quantities from the overlap matrices of the complete basis. For all molecules and basis sets tested the condition number of the PAO overlap matrix is almost independent from the size of the extended basis set. This is in contrast to the findings from the reference condition numbers, where an increase of several orders of magnitude is found.

To confirm the above mentioned connection between the condition number and the decay behavior of the density matrix we perform an additional investigation for the 10-glycine chain. We computed for each C-atom the density matrix element having the largest absolute value with respect to the central C-atom. In figure 5.2 a logarithmic plot for the minimal and the TZDP basis sets are shown to be in agreement with the findings of Table 5.1. The smallest condition number of the TZDP-PAO systems corresponds to the fastest decrease of the PAO density matrix and vice versa.

Having motivated the usefulness of a combined PAO- $\mathcal{O}(N)$ method we proceed with a detailed description of its implementation.

Basis	MIN	DZ	DZSP	TZDP
10-glycine				
SCF	2.83	9.39	9.60	10.87
PAO-SCF	—	2.08	2.15	2.12
$(\text{H}_2\text{O})_{30}$				
SCF	2.26	6.18	6.63	9.01
PAO-SCF	—	1.70	1.87	1.85
C_{60}				
SCF	3.37	13.25	13.40	14.56
PAO-SCF	—	2.49	2.49	2.41
$\text{C}_{19}\text{H}_{28}\text{O}_2$				
SCF	3.24	10.58	10.90	12.71
PAO-SCF	—	2.38	2.46	2.48

Table 5.1: Logarithm of the condition number of the overlap matrix for different basis sets computed from full SCF and PAO-SCF calculations.

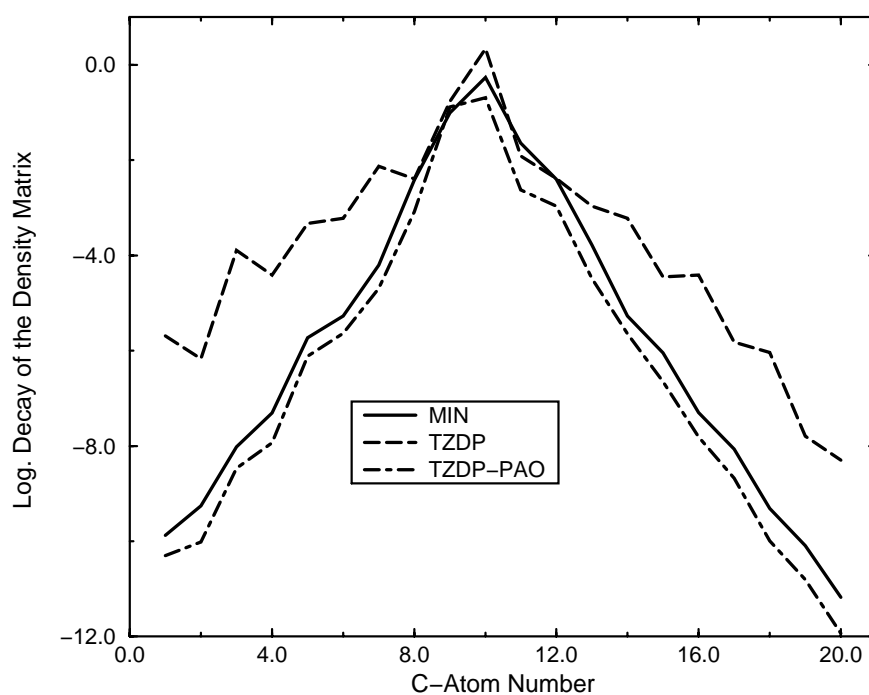


Figure 5.2: Logarithmic decay of the density matrix for the 10-glycine chain. Only matrix elements corresponding to C-atoms are considered. The results are taken from a full SCF calculation with a MIN/TZDP basis and from a PAO-SCF calculation with a TZDP basis.

5.3.2 Implementation of a Combined PAO-O(N) Method

The O(N) methods we have implemented, namely the Chebyshev polynomial expansion method (CPE) [73], the conjugate gradient density matrix search (DMS) [74] and the canonical purification of the density matrix (CP) [75], were developed originally for tight-binding models which usually assume an orthogonal basis. To handle nonorthogonal basis sets modifications are required. Two main approaches may be used for this purpose. The first is to solve for the density matrix directly in terms of the nonorthogonal basis. These procedures has been described in the literature [39,82,83,75] and are not investigated here. The approach used in our work is to transform the Hamiltonian matrix into an orthonormal basis, solving for the density matrix and back transforming to the original basis. The small condition number of the PAO overlap matrix makes this procedure especially attractive.

A number of choices for transformations to an orthonormal basis are possible. We obtain the transformation matrix from the Cholesky decomposition of the overlap matrix [50]. For symmetric positive definite matrices, such as the overlap matrix, the decomposition and the required transformations can be expressed as,

$$S = U^T U ,$$

$$H_{ortho} = U^{-T} H U^{-1} , \quad P = U^{-1} P_{ortho} U^{-T} , \quad (5.15)$$

where U is an upper triangular matrix. The computational cost of the Cholesky decomposition scales linearly with system size in the large molecule limit [84]. The inverse of U is never explicitly computed, we solve Eqs. (5.15) with respect to H and P_{ortho} and treat the corresponding set of triangular equations. This can be done in linear scaling time using sparse matrix techniques.

Having transformed the density update problem first to a PAO basis (see sec. 5.2) and finally to an orthonormal basis we can apply the O(N) methods mentioned above. Since technical details can be found in the literature [73–75,39,85], we restrict ourselves to a brief description of the algorithms.

The DMS scheme used in this work is based on minimizing the following functional

with respect to the density matrix [84],

$$\Omega = \text{Tr}[(3P^2 - 2P^3)H] + \mu(\text{Tr}[P] - n_{el}) . \quad (5.16)$$

Here, μ is a Lagrangian multiplier introduced to enforce the correct number of electrons n_{el} . The idem-potency constraint on the density matrix is implicitly imposed by the use of the McWeeny purification transformation: $3P^2 - 2P^3$. It is straightforward to compute the gradient,

$$\frac{\partial \Omega}{\partial P} = 3PH + 3HP - 2P^2H - 2PHP - 2HP^2 + \mu \mathbf{1} . \quad (5.17)$$

μ is calculated at every step from the condition, $\text{Tr}[\frac{\partial \Omega}{\partial P}] = 0$. Since the gradient and hence the search direction is traceless, any finite step preserves the number of electrons.

CP is an iterative scheme to compute the density matrix involving an initial guess, a purification formula and an energy expression. The initial guess density matrix is obtained from,

$$P_0 = \frac{\lambda}{2}(\mu \mathbf{1} - H) + \frac{1}{2} \mathbf{1} , \quad (5.18)$$

$$\mu = \frac{\text{Tr}[H]}{n} , \quad \lambda = \min\left(\frac{n_{el}}{H_{max} - \mu}, \frac{n - n_{el}}{\mu - H_{min}}\right) ,$$

where n denotes the number of basis functions. $H_{max/min}$ are upper and lower bounds on the spectrum of H . We find these values using a linear scaling Lanczos algorithm.

The purification formula is as follows,

$$P_{n+1} = \begin{cases} ((1 - 2c_n)P_n + (1 + c_n)P_n^2 - P_n^3)/(1 - c_n), & c_n \leq \frac{1}{2} , \\ ((1 + c_n)P_n^2 - P_n^3)/c_n, & c_n > \frac{1}{2} , \end{cases} \quad (5.19)$$

$$c_n = \frac{\text{Tr}[P_n^2 - P_n^3]}{\text{Tr}[P_n - P_n^2]} .$$

The purification transformation is repeated until the energy expression, $E_n = \text{Tr}[P_n H]$, is minimized. It can be shown, that the purification algorithm converges monotonically and quadratically to the correct density matrix [75].

Whereas DMS and CP require an iterative process, CPE calculates the density matrix directly using its representation as a matrix functional of the Hamiltonian matrix,

$$P = f(H) , \tag{5.20}$$

$$f(\epsilon) = \frac{1}{1 + \exp(\beta(\epsilon - \mu))} .$$

f is the Fermi distribution and β denotes the inverse of the (fictitious) temperature. The matrix functional is expanded in terms of Chebyshev polynomials [50],

$$P = f(H) = \sum_{i=0}^{n_p} c_i T_i(H) , \tag{5.21}$$

where n_p is the order of the expansion. The polynomials $T_i(H)$ satisfy the recursion relations,

$$T_{i+1}(H) = 2HT_i(H) - T_{i-1}(H) , \tag{5.22}$$

with $T_0(H) = \mathbf{1}$ and $T_1(H) = H$. The expansion coefficients c_i of the Chebyshev expansion can easily be determined [50]. Since Chebyshev polynomials require a domain between minus one and one, H has to be scaled such that its eigenvalues are between these values. The extreme eigenvalues needed to scale H are found using a linear scaling Lanczos algorithm. From Eq. (5.20) it is clear, that we have to compute in addition the chemical potential in order to locate the decay region. This is done in the following way. The Chebyshev polynomials are calculated without any reference to the chemical potential. Thus, they can be used with different expansion coefficients depending on μ . Using a bisection scheme, we chose μ such that the correct number of electrons is enforced. For computational convenience it is advantageous to take the complementary error function instead of the Fermi distribution since it decays faster to zero away from the chemical potential μ . [39]

We finish the section with a brief description of the storage format we have chosen to manipulate the sparse matrices. We use the variable block row (VBR) format which is a generalization of the compressed sparse row (CSR) format [86]. The VBR format takes into consideration that the matrices we have to manipulate are blocked with different block sizes. To maintain a desired level of accuracy during

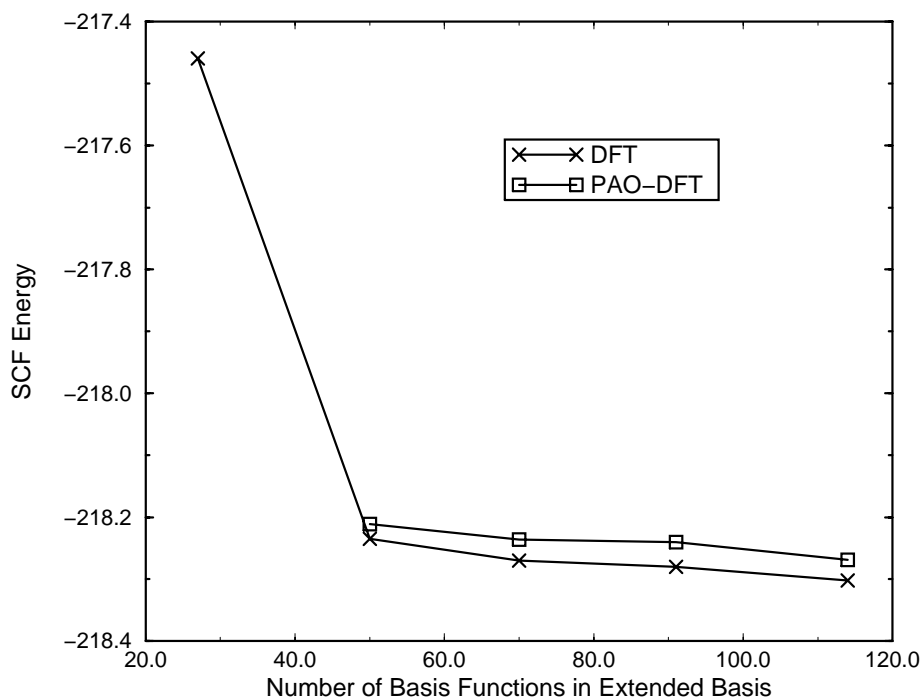


Figure 5.3: Absolute SCF energies [Hartrees] of DFT and PAO-DFT theory for various basis sets. *All-electron* calculations of C_3H_7F using the exchange-correlation functional BLYP are shown. The following basis sets are applied (from left to right): STO-6G, 6-31G, 6-31G*, 6-31G** and 6-31++G**.

the calculation, the VBR matrices are allowed to grow or shrink and every matrix is allowed to have its own form. Thus, before performing an algebraic matrix operation, we determine the form of the resulting matrix. Consequently, periodic screening of matrix elements is required to avoid excessive growth of the number of nonzero matrix elements. Using a thresholding criterion, the additional small elements can be eliminated.

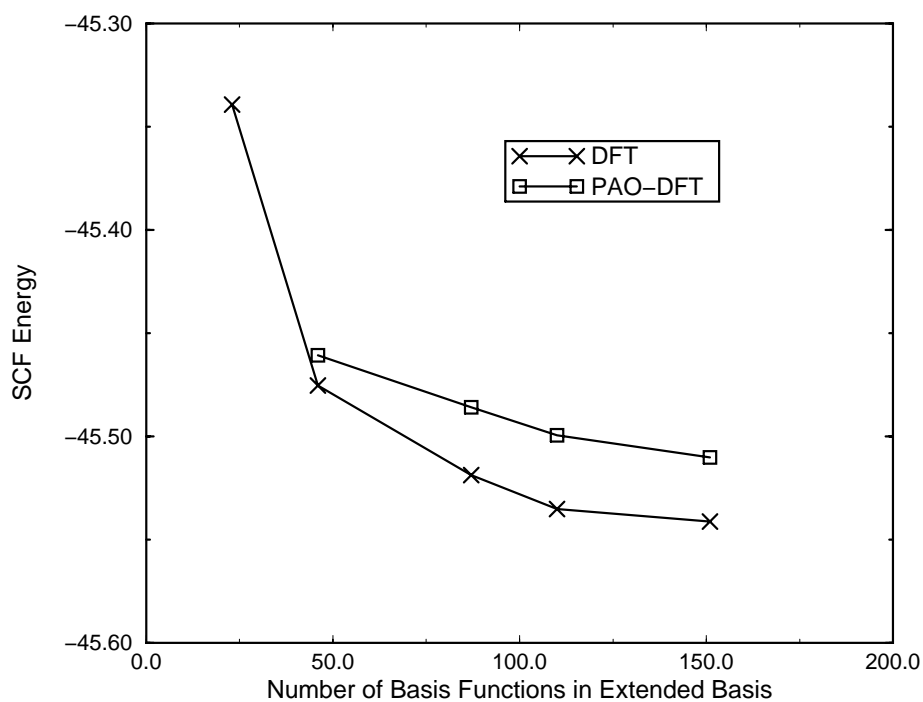


Figure 5.4: Absolute SCF energies [Hartrees] of DFT and PAO-DFT theory for various basis sets. *Pseudopotential* calculations of C_3H_7F using the exchange-correlation functional BLYP are shown. The following basis sets are applied (from left to right): MIN, DZ, DZSP, TZSP and TZDP.

5.4 Results and Discussion

The PAO method has been implemented in the GAPW method, which allows pseudopotential and all-electron calculations to be performed (see section 1.3). The program is suited for periodic and for finite systems (see section 3).

First, to demonstrate the concerted acting of the optimized minimal PAO basis and the larger set of atomic orbitals it is formed from, we look at the convergence of the PAO's in terms of the absolute energy as the extended basis set size is increased. For that, we performed a series of pseudopotential and all-electron DFT calculations on the molecule 1-fluoropropane (C_3H_7F). From figure 5.3 and 5.4, one can see that, although the PAO's are still a minimal basis, the PAO energy is much improved in energy with respect to the conventional minimal basis set energy (STO-6G/MIN). In fact, the PAO energy of the largest extended basis set used, namely the 6-31++G** and the TZDP basis, is almost as low in energy as the double zeta basis sets with one set of polarization functions (6-31G*/DZSP). Furthermore, the PAO energies nearly parallel the full SCF energy as basis set size is increased from double zeta onwards, suggesting that the relative error in energy may be even lower. It is not surprising, that the same trends obtained with DFT calculations, apply also to Hartree-Fock calculations. [70].

Next, we performed geometry optimizations on a variety of different systems. Tables 5.2 and 5.3 show a comparison of structural parameters of small molecules obtained from full DFT and PAO-DFT calculations. Pseudopotential and all-electron calculations using the exchange-correlation functional BLYP are shown. The computations were done with a double zeta basis with one set of polarization functions (6-31G*/DZSP). In addition, a minimal basis (STO-6G/MIN) was used for comparison. The root-mean-square (RMS) error of the PAO bond distances relative to the full SCF distances is 0.015 rÅ using pseudopotentials and 0.010 rÅ in the case of all-electron calculations. This deviation is a substantial improvement over the corresponding RMS difference for the minimal basis, which is in both cases 0.074 rÅ. The RMS error of bond angles is 0.60° and 1.02°, compared to considerably larger

		Pseudopotential Calculation			All-Electron Calculation		
Molecule	Parameter	DZSP	PAO-DZSP	MIN	6-31G*	PAO-6-31G*	STO-6G
H ₂	r(HH)	0.755	0.755	0.697	0.748	0.748	0.733
CH ₄	r(CH)	1.103	1.102	1.103	1.101	1.100	1.101
C ₂ H ₂	r(CC)	1.222	1.224	1.258	1.215	1.215	1.211
	r(CH)	1.076	1.079	1.094	1.073	1.073	1.086
C ₂ H ₄	r(CC)	1.346	1.351	1.394	1.341	1.347	1.343
	r(CH)	1.097	1.097	1.101	1.095	1.094	1.102
	∠(CCH)	121.6	121.5	121.7	121.9	121.6	122.0
	∠(HCH)	116.7	117.0	116.6	116.2	116.7	116.0
C ₂ H ₆	r(CC)	1.537	1.559	1.613	1.542	1.563	1.559
	r(CH)	1.105	1.104	1.103	1.104	1.102	1.105
	∠(CCH)	111.4	111.4	110.6	111.4	111.1	110.7
	∠(HCH)	107.5	107.5	108.3	107.5	107.8	108.2
SiH ₄	r(SiH)	1.478	1.481	1.443	1.496	1.496	1.449
H ₂ O	r(OH)	0.979	0.975	1.033	0.980	0.978	1.044
	∠(HOH)	103.7	105.1	104.9	102.9	104.8	96.02

Table 5.2: Structural parameters [$\mathbf{rA},^\circ$] of small molecules obtained from full SCF and PAO-SCF calculations. Pseudopotential and all-electron calculations using the exchange-correlation functional BLYP are shown.

		Pseudopotential Calculation			All-Electron Calculation		
Molecule	Parameter	DZSP	PAO-DZSP	MIN	6-31G*	PAO-6-31G*	STO-6G
H ₂ S	r(SH)	1.354	1.359	1.371	1.361	1.361	1.367
	∠(HSH)	91.85	92.40	93.79	92.35	91.92	90.77
HCN	r(CH)	1.082	1.083	1.112	1.078	1.077	1.096
	r(CN)	1.169	1.169	1.267	1.169	1.169	1.209
NH ₃	r(NH)	1.029	1.025	1.041	1.030	1.024	1.076
	∠(HNH)	106.5	107.3	120.0	104.9	107.1	99.42
PH ₃	r(PH)	1.426	1.430	1.411	1.437	1.436	1.418
	∠(HPH)	93.13	93.56	95.94	92.77	92.72	91.16
N ₂	r(NN)	1.122	1.120	1.295	1.118	1.117	1.205
CO	r(CO)	1.147	1.146	1.299	1.150	1.149	1.206
CO ₂	r(CO)	1.181	1.183	1.321	1.184	1.187	1.243
SF ₆	r(SF)	1.624	1.686	-	1.628	1.667	-
B ₂ H ₆	r(BB)	1.788	1.802	1.864	1.783	1.785	1.824
	r ₁ (BH)	1.193	1.194	1.165	1.197	1.196	1.173
	r ₂ (BH)	1.329	1.335	1.370	1.324	1.325	1.348
	∠(HBH)	122.2	122.4	120.4	122.3	122.0	122.6
	∠(BHB)	84.60	84.85	85.77	84.64	84.72	85.14

Table 5.3: Continuation of table 5.2. Structural parameters [r Å, $^\circ$] of small molecules obtained from full DFT and PAO-DFT calculations. Pseudopotential and all-electron calculations using the exchange-correlation functional BLYP are shown. In the case of SF₆, the sulfur minimal basis consists of six basis functions.

Bond	PAO-DZSP	DZSP	Reference	Bond	PAO-DZSP	DZSP	Reference
1	1.361	1.353	1.359	2	1.463	1.448	1.457
3	1.399	1.392	1.400	4	1.510	1.485	1.499
5	1.441	1.440	1.456	6	1.382	1.365	1.373
7	1.402	1.380	1.381	8	1.461	1.455	1.470
9	1.420	1.408	1.421	10	1.395	1.390	1.399
11	1.517	1.511	1.531	12	1.548	1.525	1.541
13	1.550	1.532	1.555	14	1.507	1.497	1.526
15	1.391	1.372	1.379	16	1.377	1.358	1.365
17	1.387	1.382	1.395	18	1.536	1.505	1.535
19	1.508	1.500	1.528	20	1.599	1.583	1.594
21	1.227	1.225	1.222	22	1.465	1.453	1.468
23	1.423	1.419	1.430	24	1.413	1.407	1.419
25	1.409	1.396	1.402	26	1.503	1.484	1.499
27	1.446	1.443	1.462	28	1.413	1.389	1.396
29	1.353	1.399	1.343	30	1.407	1.400	1.410
31	1.411	1.402	1.414	32	1.448	1.447	1.463
33	1.507	1.489	1.507	34	1.545	1.524	1.544
35	1.390	1.382	1.390	36	1.509	1.487	1.502

Table 5.4: Bond lengths [rÅ] of chlorophyll a (see figure 5.5) obtained from full SCF and PAO-SCF calculations. Pseudopotential calculations using LDA are shown. The reference bond lengths are taken from all-electron calculations using the exchange-correlation functional BP [87].

Bond	PAO-DZSP	DZSP	Reference	Bond	PAO-DZSP	DZSP	Reference
37	1.448	1.446	1.461	38	1.386	1.368	1.373
39	1.397	1.377	1.382	40	1.418	1.407	1.417
41	1.402	1.395	1.406	42	2.018	2.016	2.046
43	2.045	2.051	2.082	44	2.010	2.003	2.031
45	2.123	2.134	2.165	46	1.547	1.527	1.550
47	1.543	1.523	1.544	48	1.508	1.500	1.522
49	1.229	1.227	1.219	50	1.364	1.345	1.357
51	1.487	1.459	1.463	52	1.498	1.482	1.500
53	1.357	1.355	1.358	54	1.520	1.500	1.512
55	1.518	1.503	1.519	56	1.551	1.532	1.547
57	1.540	1.521	1.539	58	1.546	1.531	1.551
59	1.542	1.523	1.540	60	1.537	1.524	1.548
61	1.535	1.518	1.543	62	1.545	1.525	1.539
63	1.542	1.527	1.551	64	1.548	1.528	1.540
65	1.547	1.532	1.547	66	1.539	1.520	1.543
67	1.537	1.519	1.540	68	1.539	1.525	1.551
69	1.539	1.520	1.540	70	1.544	1.524	1.540

Table 5.5: Continuation of table 5.4. Bond lengths [rÅ] of chlorophyll a (see figure 5.5) obtained from full DFT and PAO-DFT calculations. Pseudopotential calculations using LDA are shown. The reference bond lengths are taken from all-electron calculations using the exchange-correlation functional BP [87].

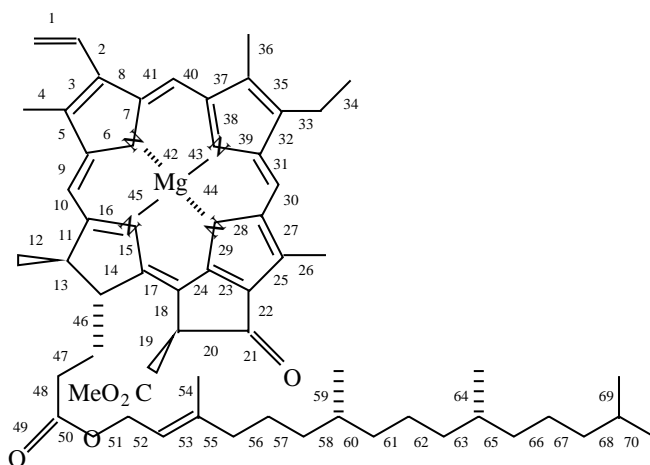


Figure 5.5: Chlorophyll a

Crystal	DZSP		PAO-DZSP		MIN	
	a [rÅ]	E [Hartrees]	a [rÅ]	E [Hartrees]	a [rÅ]	E [Hartrees]
C (D)	3.56	-366.896	3.58	-366.328	3.64	-365.259
Si (D)	5.44	-254.572	5.54	-253.979	5.63	-253.040
BN (ZB)	3.61	-414.768	3.62	-414.153	3.70	-413.285
SiC (ZB)	4.35	-311.366	4.42	-310.019	4.48	-309.242
NaCl (SC)	5.68	-2021.25	5.58	-2020.54	5.43	-2017.65

Table 5.6: Lattice constants and absolute SCF energies of various crystals. Full SCF and PAO-SCF calculations using pseudopotentials and the exchange-correlation functional BP are shown. The lattice type is added in parenthesis: D denotes diamond, ZB zincblende, and SC sodium chloride.

deviations of 4.67° and 3.06° .

As a more ambitious example we have studied the molecular structure of chlorophyll a ($\text{MgC}_{55}\text{H}_{72}\text{N}_4\text{O}_5$; see figure 5.5) playing a central role in many biological system. Tables 5.4 and 5.5 reports our results for the bond distances obtained from pseudopotential calculations using LDA with a DZSP basis. The RMS error of the PAO bond distances relative to the full SCF distances is 0.017 rA. To stress the very high quality of the result, we have added in tables 5.4 and 5.5 the bond distances from a DFT all-electron calculation using the exchange-correlation functional BP [87]. Here, the RMS error of the bond distances relative to the full SCF distances is 0.018 rA.

There are also cases where the PAO method is not performing that well. In table 5.6, we have summarized the lattice constants of various crystals obtained from pseudopotential calculations with a DZSP basis using the exchange-correlation functional BP. Here the RMS error is 0.08 rA. Although the deviation is an improvement over the corresponding RMS difference for the minimal basis (0.18 rA), the result is not satisfactory. This is even more astonishing, since the absolute energies seem to be reasonably converged. This certainly needs further investigations.

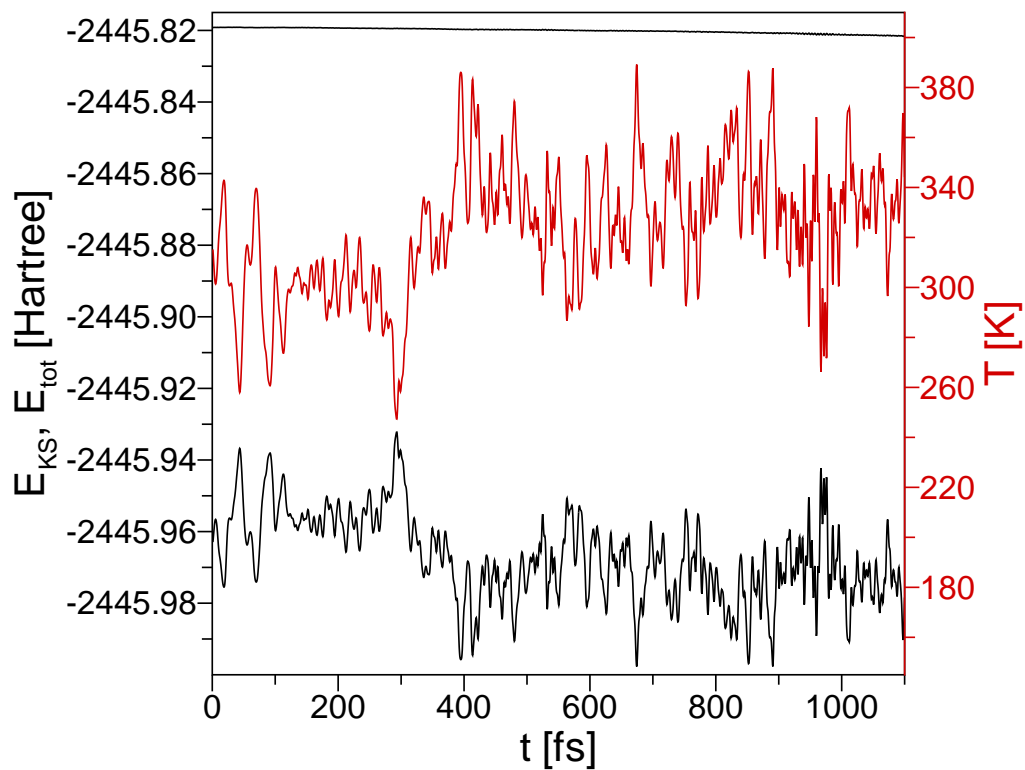


Figure 5.6: Time evolution of the total energy E_{tot} (upper curve), the temperature T (middle curve) and the Kohn-Sham energy E_{KS} (lower curve) of 32 H_2O molecules during the MD simulation. All-electron calculations with a TZDP basis set using the exchange-correlation functional BLYP are shown.

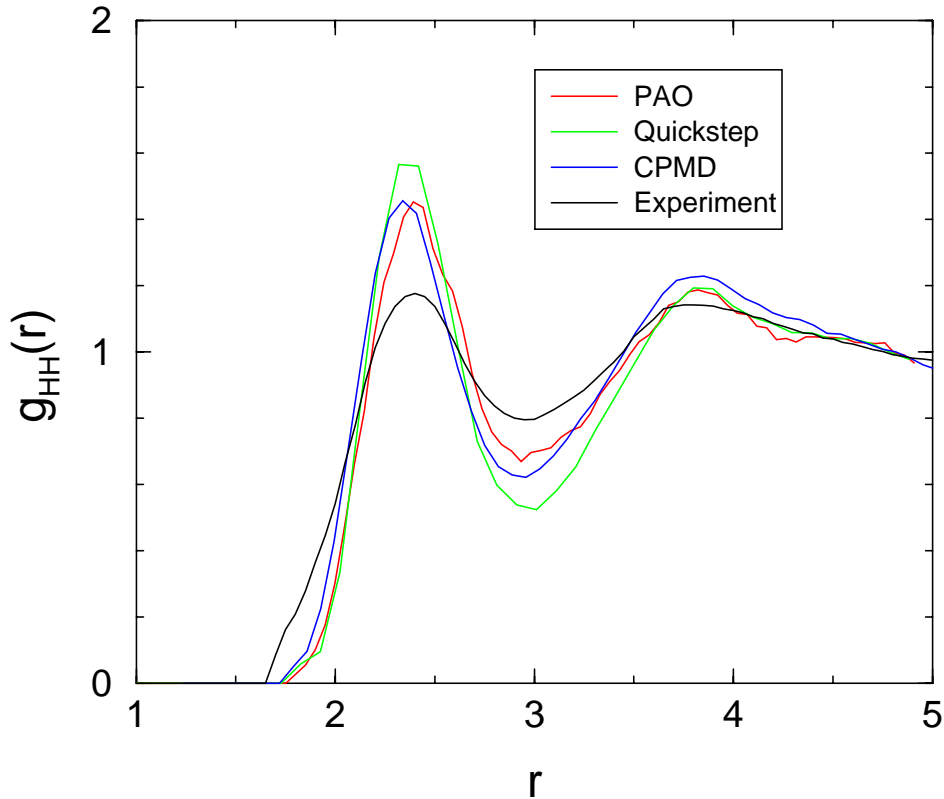


Figure 5.7: H-H pair correlation function of 32 water molecules obtained from our PAO-MD (blue line) compared to the full MD simulation using the same program from ref. [21] (red line), to a Car-Parrinello MD from ref. [88] (black line) and to the experimental neutron scattering from ref. [89] (green line).

As a final test of the applicability of the PAO method to different fields of electronic structure calculations, we present an ab-initio all-electron PAO-MD run for liquid water. A system of 32 H_2O molecules contained in a simple box of length 9.87 \AA was simulated imposing periodic boundary conditions. This model system was also employed by earlier full SCF-MD studies using the same program [21], and a Car-Parrinello molecular dynamics (CPMD) simulation of liquid heavy water (D_2O) [88]. All simulations were done with the exchange-correlation functional BLYP. In contrast to the work of Ref. [21], where a DZ basis and a DZSP basis set is used for the hydrogen and the oxygen atom respectively, we added a p-type polarization function to the hydrogen basis set (DZSP) and a triple-zeta basis with d-type polarization

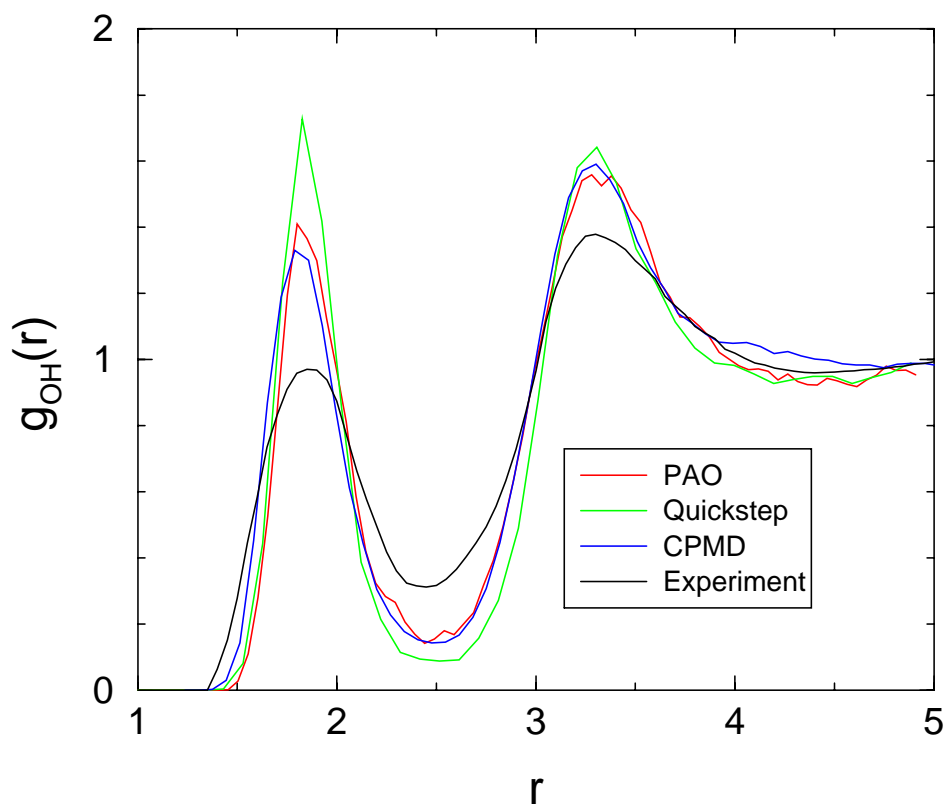


Figure 5.8: O-H pair correlation function of 32 water molecules obtained from our PAO-MD (blue line) compared to the full MD simulation using the same program from ref. [21] (red line), to a Car-Parrinello MD from ref. [88] (black line) and to the experimental neutron scattering from ref. [89] (green line).

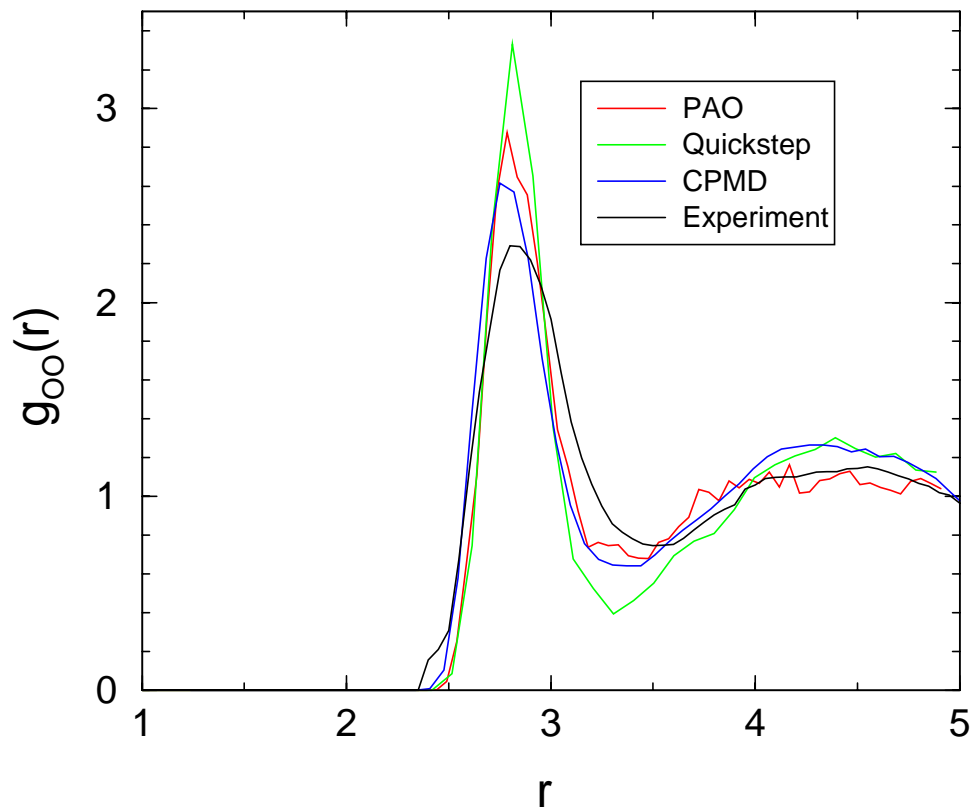


Figure 5.9: O-O pair correlation function of 32 water molecules obtained from our PAO-MD (blue line) compared to the full MD simulation using the same program from ref. [21] (red line), to a Car-Parrinello MD from ref. [88] (black line) and to the experimental neutron scattering from ref. [89] (green line).

functions (TZSP) was applied for oxygen.

The time evolution of the total energy, the Kohn-Sham energy and the temperature of the 32 H₂O molecules are plotted in figure 5.6. The average temperature of the sample was 328.8 K. The total energy shows a slight drift, but this deficiency can be removed by converging the PAO's to a higher accuracy. In figure 5.6, PAO convergence is assumed if the relative value of the energy functional, defined in Eq. 5.6, is decreasing by less than 10⁻¹². The latter might already seem an extraordinarily small convergence threshold. However, experience shows that extremely accurate converged PAO's are required to perfectly conserve the total energy.

Radial distribution functions $g_{HH}(r)$, $g_{OH}(r)$ and $g_{OO}(r)$ are shown in Figs. 5.7, 5.8 and 5.9, respectively. The maxima and the minima of the $g_{ij}(r)$ are properly reproduced compared to the experimental results from neutron scattering [89], but all the simulated distribution functions appear to be overstructured relative to the experimental result. It is important to notice that the distribution functions of our PAO-MD simulation with the enlarged basis set is much closer to the experimental results compared to the original full SCF-MD run with the smaller basis set.

Finally, we have implemented a combined PAO-O(N) scheme as described in section 5.3. The linear scaling methods we use are the Chebyshev polynomial expansion method (CPE) [73], the conjugate gradient density matrix search (DMS) [74], and the canonical purification of the density matrix (CP) [75].

If we keep all the matrix elements during a PAO-O(N) calculation, all O(N) methods above have to reproduce the results of diagonalization exactly. Consequently, Table 5.7 reports the forces, acting on a randomly chosen atom, and the absolute SCF energies of a Benzene and a NH₃ molecule with a neglect threshold equals zero. Besides numerical noise all results are equal, confirming a correct implementation.

Benchmark calculations on the unbranched alkane molecules n-C₃₀H₆₂, n-C₆₀H₁₂₂ and n-C₉₀H₁₈₂ were carried out using pseudopotentials, a DZSP basis and LDA. All timings reported are for a standard single processor workstation (IBM 397). The thresholding criterion for neglecting matrix elements is chosen such that the absolute

Benzene				
Method	Energy in [a.u.]	Forces in [a.u.] acting on C-Atom one		
		F_x	F_y	F_z
D	-37.31225	-0.02591	0.03405	-0.00012
CPE	-37.31224	-0.02591	0.03405	-0.00012
CP	-37.31225	-0.02591	0.03405	-0.00012
DMM	-37.31225	-0.02591	0.03405	-0.00012
NH ₃				
Method	Energy in [a.u.]	Forces in [a.u.] acting on H-Atom three		
		F_x	F_y	F_z
D	-11.67487	0.00741	0.00427	0.00000
CPE	-11.67481	0.00742	0.00428	0.00002
CP	-11.67487	0.00741	0.00427	0.00000
DMM	-11.67487	0.00739	0.00426	-0.00001

Table 5.7: Forces, acting on a randomly chosen atom, and absolute SCF energies of Benzene and NH₃. The results are obtained from PAO-DFT calculations using diagonalization (D), Chebyshev polynomial expansion (CPE), canonical purification of the density matrix (CP) and conjugate gradient density matrix search (DMS).

Method	CPU time in [s]	Energy in [a.u.]	Forces [a.u.]: C-Atom ten		
			F_x	F_y	F_z
$C_{30}H_{62}$					
D	0.38	-206.6892	0.0014	0.0364	0.0003
CPE	5.72	-206.6892	0.0013	0.0364	0.0003
CP	2.72	-206.6891	0.0014	0.0364	0.0003
DMS	5.36	-206.6892	0.0014	0.0365	0.0003
$C_{60}H_{122}$					
D	2.66	-412.2175	0.0006	0.0367	-0.0060
CPE	11.58	-412.2176	0.0006	0.0364	-0.0060
CP	6.32	-412.2174	0.0006	0.0366	-0.0060
DMS	13.58	-412.2174	0.0006	0.0368	-0.0060
$C_{90}H_{182}$					
D	30.87	-617.7448	0.0003	0.0297	-0.0205
CPE	17.91	-617.7444	0.0003	0.0297	-0.0205
CP	9.77	-617.7438	0.0002	0.0297	-0.0204
DMS	21.84	-617.7445	0.0003	0.0298	-0.0206

Table 5.8: Forces, acting on a randomly chosen atom, and absolute SCF energies of linear alkane chains. In addition the CPU time requirements of a density matrix update step are given (see figure 5.10). The results are obtained from PAO-DFT calculations using diagonalization (D), Chebyshev polynomial expansion (CPE), canonical purification of the density matrix (CP) and conjugate gradient density matrix search (DMS).

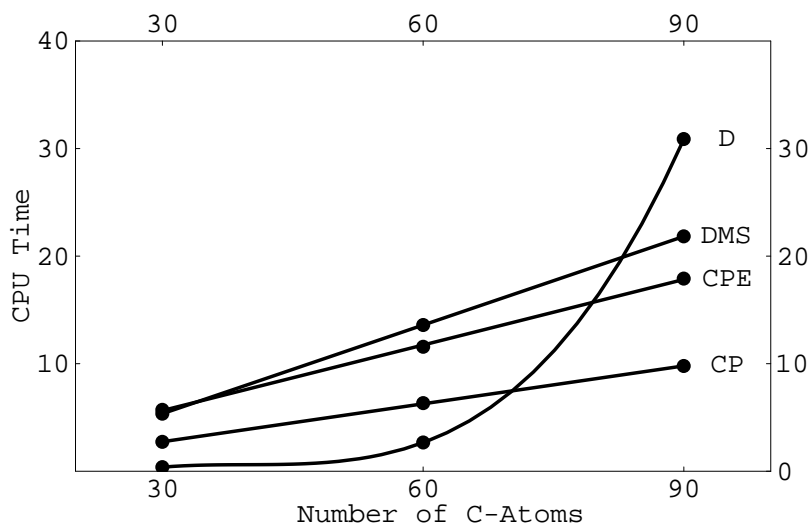


Figure 5.10: CPU time requirements [s] of a density matrix update step for the linear alkane chains $C_{30}H_{62}$, $C_{60}H_{122}$ and $C_{90}H_{182}$ using diagonalization (D), Chebyshev polynomial expansion (CPE), canonical purification of the density matrix (CP) and conjugate gradient density matrix search (DMS). All density update procedures are in terms of the PAO minimal basis.

SCF energies yield microhartree accuracy relative to the results of diagonalization (D). The appropriate choice of the neglect threshold is confirmed by the results given in table 5.8. Here, the forces, acting on a randomly chosen atom, and the absolute SCF energies of the linear alkane chains are reported. In addition the CPU time requirements of a density matrix update step in terms of the PAO minimal basis are given. The CPU times are plotted in figure 5.10. Linear scaling is achieved using each of these $O(N)$ methods. In the present implementation, the fastest method for replacing diagonalization in large scale calculations is CP. Here, the crossover in CPU time compared to diagonalization occurs at roughly 70 C-atoms. It should be noted at this point, that the sparse matrix routines used are not optimized with respect to computational efficiency and use of machine architecture whereas standard diagonalization are well optimized. Removing this deficiency should move the crossover with diagonalization to smaller systems than presented here. However, to apply the PAO- $O(N)$ method to study extended systems currently intractable

with large nonorthogonal basis sets, a parallelized implementation of the program is required. We proceed in this direction, encouraged by the present results.

5.5 Conclusions

We have presented the theoretical and algorithmic principles to construct minimal basis sets optimized in the molecular environment. These so-called polarized atomic orbitals (PAO's) permit the density matrix update procedure of self-consistent field theory to be performed in a small basis, while retaining a considerable amount of the flexibility of larger basis sets. Our test calculations show, that electronic structure calculation at the PAO level greatly improve all investigated properties ranging from absolute energies, to bond distances and bond angles relative to a rigid conventional minimal basis set. The simulation of the liquid water system shows that an all-electron ab-initio MD is feasible using the PAO method. A comparison of the measured radial distribution functions gives good agreement with the results of earlier MD studies. However, the PAO's have to be converged to extremely high accuracy to conserve the total energy during a MD run.

A detailed study of the condition number of the PAO overlap matrix and the decay properties of the PAO density matrix show the potential of the PAO method combined with linear scaling algorithms. We find that a combined PAO- $O(N)$ method results in a numerical stable procedure, even if large nonorthogonal basis sets are applied. With implementations of the Chebyshev polynomial expansion method, the conjugate gradient density matrix search, and the canonical purification of the density matrix, linear scaling was demonstrated for unbranched alkane chains. In the present implementation canonical purification is found to be most efficient.

References

- [1] M. P. Allen and D. J. Tildesley, *Computer Simulations of Liquids*, Oxford University Press, Oxford, (1987).
- [2] A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry*, McGraw-Hill, New York, (1982).
- [3] J. C. Slater and G. F. Koster, *Phys. Rev.*, **94**, 1498, (1954).
- [4] C. M. Goringe, D. R. Bowler, and E. Hernández, *Rep. Prog. Phys.*, **60**, 1447, (1997).
- [5] J. A. Poble and D. L. Beveridge, *Approximate Molecular Orbital Theory*, Academic Press, New York, (1987).
- [6] W. Thiel, *Adv. Chem. Phys.*, **93**, 703, (1996).
- [7] R. G. Parr and W. Yang, *Density-Functional Theory of Atoms and Molecules*, Oxford University Press, Oxford, (1989).
- [8] P. Hohenberg and W. Kohn, *Phys. Rev. B*, **136**, 864, (1964).
- [9] W. Kohn and L. J. Sham, *Phys. Rev. A*, **140**, 1133, (1965).
- [10] D. M. Ceperley and B. J. Alder, *Phys. Rev. Lett.*, **45**, 566, (1980).
- [11] A. D. Becke, *Phys. Rev. A*, **38**, 3098, (1988).
- [12] C. Lee, W. Yang, R. G. Parr, *Phys. Rev. B*, **37**, 785, (1988).
- [13] J. P. Perdew, *Phys. Rev. B*, **33**, 8822, (1986).
- [14] B. G. Johnson, P. M. W. Gill and J. A. Poble, *J. Chem. Phys.*, **98**, 5612, (1993).
- [15] R. N. Barnett and U. Landmann, *Phys. Rev. B*, **48**, 2081, (1993).

- [16] F. Sim, A. St-Amant, I. Papai and D. R. Salahub, *J. Am. Chem. Soc.*, **114**, 4391, (1992).
- [17] M. Sprik, *J. Chem. Phys.*, **95**, 2283, (1991).
- [18] M. T. Yin and M. L. Cohen, *Phys. Rev. B*, **26**, 5668, (1982).
- [19] G. Lippert, J. Hutter and M. Parrinello, *Mol. Phys.*, **92**, 477, (1997).
- [20] G. Lippert, J. Hutter and M. Parrinello, *Theor. Chim. Acta*, **103**, 124, (1999).
- [21] M. Krack and M. Parrinello, *Phys. Chem. Chem. Phys.*, **2**, 2105, (2000).
- [22] P. E. Blöchl, *Phys. Rev. B*, **50**, 17953, (1994).
- [23] Y. C. Zheng and J. Almlöf, *Chem. Phys. Lett.*, **214**, 397, (1993).
- [24] Y. C. Zheng and J. Almlöf, *J. Mol. Struct.*, **388**, 277, (1996).
- [25] F. R. Gantmacher, *Matrizenrechnung*, VEB Deutscher Verlag der Wissenschaften, Berlin, (1970).
- [26] V. I. Lebedev, *Zh. Vychisl. Mat. Mat. Fiz.*, **15**, 48, (1975).
- [27] V. I. Lebedev, *Zh. Vychisl. Mat. Mat. Fiz.*, **16**, 293, (1976).
- [28] S. Goedecker, M. Teter and J. Hutter, *Phys. Rev. B*, **54**, 1703, (1996).
- [29] R. W. Hockney, *Computer Simulation Using Particles*, McGraw-Hill, New York, (1981).
- [30] R. N. Barnett and U. Landmann, *Phys. Rev. B*, **48**, 2081, (1993).
- [31] G. J. Martyna and M. E. Tuckerman, *J. Chem. Phys.*, **110**, 2810, (1999).
- [32] P. E. Blöchl, *J. Chem. Phys.*, **103**, 7422, (1995).
- [33] P. Pulay, *Molec. Phys.*, **17**, 197, (1969).

- [34] Gaussian 94, Revision B.2, M. J. Frisch, G. W. Trucks, H. B. Schlegel, P. M. W. Gill, B. G. Johnson, M. A. Robb, J. R. Cheeseman, T. Keith, G. A. Petersson, J. A. Montgomery, K. Raghavachari, M. A. Al-Laham, V. G. Zakrzewski, J. V. Ortiz, J. B. Foresman, C. Y. Peng, P. Y. Ayala, W. Chen, M. W. Wong, J. L. Andres, E. S. Replogle, R. Gomperts, R. L. Martin, D. J. Fox, J. S. Binkley, D. J. Defrees, J. Baker, J. P. Stewart, M. Head-Gordon, C. Gonzalez, and J. A. Pople, Gaussian, Inc., Pittsburgh PA, (1995).
- [35] J. B. Hendrickson, D. J. Cram and G. S. Hammond, *Organic Chemistry*, McGraw-Hill, Tokyo, (1970).
- [36] G. H. Wannier, *Phys. Rev.* **52**, 191, (1937).
- [37] D. Vanderbilt and R. D. King-Smith, *Phys. Rev. B* **47**, 1651, (1993).
- [38] R. Resta, *Rev. Mod. Phys.* **66**, 899, (1994).
- [39] S. Goedecker, *Rev. of Mod. Phys.* **71**, 1085, (1999).
- [40] S. F. Boys, *Rev. Mod. Phys.* **32**, 296, (1960).
- [41] J. Pipek and P. G. Mezey, *J. Chem. Phys.* **90**, 4916, (1989).
- [42] C. Edmiston and K. Ruedenberg, *Rev. Mod. Phys.* **35**, 457, (1963).
- [43] W. von Niessen, *J. Chem. Phys.* **56**, 4290, (1972).
- [44] N. Marzari and D. Vanderbilt, *Phys. Rev. B* **56**, 12847, (1997).
- [45] I. Souza and R. M. Martin, *Phys. Rev. Lett.* **81**, 4452, (1998).
- [46] P. L. Silvestrelli, N. Marzari, D. Vanderbilt and M. Parrinello, *Solid State Com.* **107**, 7, (1998).
- [47] P. L. Silvestrelli, *Phys. Rev. B* **59**, 9703, (1999).
- [48] R. Resta, *Phys. Rev. Lett.* **80**, 1800, (1998).
- [49] R. Resta, *Phys. Rev. Lett.* **82**, 370, (1999).

- [50] W. H. Press, S. A. Teukolsky, W. T. Vetterling and B. P. Flannery, *Numerical Recipes*, Cambridge University Press, Cambridge, (1992).
- [51] P. Pulay, Chem. Phys. Lett. **73**, 393, (1980).
- [52] M. Parrinello and A. Rahman, Phys. Rev. Lett. **45**, 1196, (1980).
- [53] J. M. Lighthill, *Introduction to Fourier Analysis and Generalized Functions*, Cambridge University Press, Cambridge, (1958).
- [54] J. Hutter *et al.*, computer code CPMD, MPI für Festkörperforschung and IBM Zürich Research Laboratory, (1995-1999).
- [55] The primitive lattice vectors for the various symmetries considered are below (in Bohr). The cell volume for all cases is $V \approx 8000 \text{ Bohr}^3$.
- simple cubic: $\mathbf{a}_1 = (20.0, 0.0, 0.0)$, $\mathbf{a}_2 = (0.0, 20.0, 0.0)$, $\mathbf{a}_3 = (0.0, 0.0, 20.0)$,
orthorhombic: $\mathbf{a}_1 = (20.0, 0.0, 0.0)$, $\mathbf{a}_2 = (0.0, 22.0, 0.0)$, $\mathbf{a}_3 = (0.0, 0.0, 18.2)$,
fcc: $\mathbf{a}_1 = (-15.9, 0.0, 15.9)$, $\mathbf{a}_2 = (0.0, 15.9, 15.9)$, $\mathbf{a}_3 = (-15.9, 15.9, 0.0)$,
bcc: $\mathbf{a}_1 = (12.6, 12.6, 12.6)$, $\mathbf{a}_2 = (-12.6, 12.6, 12.6)$, $\mathbf{a}_3 = (-12.6, -12.6, 12.6)$,
hexagonal: $\mathbf{a}_1 = (20.0, 0.0, 0.0)$, $\mathbf{a}_2 = (-10.0, 17.3, 0.0)$, $\mathbf{a}_3 = (0.0, 0.0, 23.1)$,
triclinic: $\mathbf{a}_1 = (20.0, 0.0, 0.0)$, $\mathbf{a}_2 = (7.2, 22.9, 0.0)$, $\mathbf{a}_3 = (1.8, 3.2, 17.6)$.
- [56] J. Nocedal, Math. Comp. **35**, 773 (1980).
- [57] A direct comparison between the methods relying on Eq. (4.30) and the OR method, which consists of $n(n-1)/2$ rotations among all pairs of orbitals, is not proper. Certainly both methods scale $O(n^3)$, but the effort to perform one step in the OR scheme is about three times more expensive in CPU time, compared to the other methods.
- [58] J. J. Mortensen and M. Parrinello, Phys. Rev. B, submitted.
- [59] D. Sebastiani and M. Parrinello, J. Phys. Chem., submitted.
- [60] C. J. Mundy, J. Hutter and M. Parrinello, J. Am. Chem. Soc., submitted.

- [61] D. Aktah and I. Frank, *Science*, submitted.
- [62] T. W. Ebbesen, *Phys. Today*, **49**, 26, (1996).
- [63] P. M. Ajayan and T. W. Ebbesen, *Rep. Prog. Phys.*, **60**, 1025, (1997).
- [64] A. Rubio, *Condens. Matter News*, **6**, 6, (1997).
- [65] N. Hamada, S. Sawada and A. Oshiyama, *Phys. Rev. Lett.*, **68**, 1579, (1992).
- [66] R. Saito, M. Fujita G. Dresselhaus and M. S. Dresselhaus, *Appl. Phys. Lett.*, **60**, 2204, (1992).
- [67] X. Blase, L. X. Benedict, E. L. Shirley and S. G. Louie, *Phys. Rev. Lett.*, **72**, 1878, (1994).
- [68] C. A. White, B. G. Johnson, P. M. W. Gill and M. Head-Gordon, *Chem. Phys. Lett.*, **253**, 268, (1996).
- [69] M. C. Strain, G. E. Scuseria and M. J. Frisch, *Science*, **271**, 51, (1996).
- [70] M. S. Lee and M. Head-Gordon, *J. Chem. Phys.*, **107**, 9085, (1997).
- [71] G. Galli and M. Parrinello, *Phys. Rev. Lett.*, **69**, 3547 (1992).
- [72] J. J. P. Stewart, *Int. J. Quant. Chem.*, **58**, 133, (1996)
- [73] S. Goedecker and M. Teter, *Phys. Rev. B*, **51**, 9455 (1995).
- [74] X. P. Li, R. W. Nunes and D. Vanderbilt, *Phys. Rev. B*, **47**, 10891 (1993).
- [75] A. H. R. Palser and D. E. Manolopoulos, *Phys. Rev. B*, **58**, 12704 (1998).
- [76] R. W. Nunes and D. Vanderbilt, *Phys. Rev. B*, **50**, 17611 (1994).
- [77] W. Kohn, *Phys. Rev.*, **115**, 809 (1959).
- [78] W. Kohn, *Phys. Rev. B*, **7**, 4388 (1973).
- [79] W. Kohn, *Chem. Phys. Lett.*, **208**, 167 (1993).

- [80] P. E. Maslen, C. Ochsenfeld, C. A. White, M. S. Lee and M. Head-Gordon, J. Phys. Chem. A, **102**, 2215 (1998).
- [81] G. Lippert, J. Hutter, P. Ballone and M. Parrinello, J. Phys. Chem., **100**, 6231 (1996).
- [82] C. A. White, P. Maslen, M. S. Lee and M. Head-Gordon, Chem. Phys. Lett., **276**, 133, (1997).
- [83] U. Stephan and D. Drabold, Phys. Rev. B, **57**, 6391, (1998).
- [84] J. M. Millam and G. E. Scuseria, J. Chem. Phys., **106**, 5569, (1997).
- [85] A. D. Daniels and G. E. Scuseria, J. Chem. Phys., **110**, 1321, (1999).
- [86] Y. Saad, *SPARSKIT: A basic tool kit for sparse matrix computations*, Version 2, (1994).
- [87] D. Sundholm, Chem. Phys. Lett., **317** 545 (2000)
- [88] M. Sprik, J. Hutter and M. Parrinello, J. Chem. Phys., **105**, 1142, (1996).
- [89] A. K. Soper, F. Bruni, P. O. Widmark, R. Lindh and K. Karlström, J. Chem. Phys., **106**, 4597, (1997).

List of abbreviations

BLYP: Becke-Lee-Yang-Parr

BF: Boys-Forster

BO: Bloch orbitals

BP: Becke-Perdew

CG: conjugate gradient

CI: configuration interaction

CP: canonical purification

CPE: Chebyshev polynomial expansion

CPMD: Car-Parrinello molecular dynamics

CSR: compressed sparse row

DIIS: direct inversion in the iterative subspace

DFT: density functional theory

DM: density matrix

DMS: density matrix search

DZ: double-zeta

DZSP: double-zeta with one set of polarization function

GAPW: Gaussian augmented plane wave

GC: gradient correction

GTO: Gaussian type orbital

HF: Hartree-Fock

KS: Kohn-Sham

LDA: local density approximation

MD: molecular dynamics

MIN: minimal basis

MO: molecular orbitals

OR: orbital rotations

PAO: polarized atomic orbitals

PM: Pipek-Mezey

PR: preconditioner

PW: plane wave

R: Resta

RMS: root-mean-square

SCF: self consistent field

SD: steepest descent

SE: semiempirical

SMV: Silvestrelli-Mazari-Vanderbilt

TB: tight-binding

TZDP: triple-zeta with two sets of polarization functions

VBR: variable block row

WF: Wannier function

XC: exchange-correlation

Acknowledgements

First of all, I am particularly grateful to Prof. Dr. Michele Parrinello, who cordially accommodated me in his department during the last three years. Not only did he provide essential advice in all questions concerning physics, but also constant interest in my personal progress.

This thesis would not have been possible without the detailed knowledge of science which my advisor Prof. Dr. Jürg Hutter shared with me. From him I learned that theoretical physics does not only consist in inventing formulae, but also in their realization.

Special thanks also go to my various collaborators, who supported this work either with illuminating questions or with pertinent answers. In particular, I want to thank Dr. Stefan Goedecker, Dr. Gerald Lippert, Dr. Matthias Krack, Dr. Valeria Meriggalli, Dr. Chris Mundy, Dr. Daniel Sebastiani, Dr. Aldo Romero, Ata Roudgar, Dr. Walter Silvestri, and Dr. Georg Stapper.

I was happy to share many interesting times with the other group and institute members, especially during many football matches, roughly 2000 coffee breaks, and the numerous nights in the few good pubs of Stuttgart.

Last but not least, I am particularly grateful to my parents who supported me all of my life.

Curriculum Vitae

Gerd Berghold, geboren am 14.8.1968 in Stuttgart

1975–1979: Grundschule Wendlingen am Neckar

1979–1988: Robert-Bosch-Gymnasium Wendlingen am Neckar; Abitur mit den
Leistungskursen Chemie und Deutsch

1988–1997: Studium der Chemie an der Eberhard-Karls-Universität Tübingen
und an der Technischen Universität Berlin. Abschluß: Diplom

1993–2000: Studium der Mathematik an der Freien Universität Berlin und an
der Universität Stuttgart. Abschluß: Diplom

1997–2000: Promotion bei Michele Parrinello am Max-Planck-Institut für Fest-
körperforschung in Stuttgart

Seit 2000: Risk Consultant bei Arthur Andersen

Stuttgart, Juli 2001